

Sadri Hassani

# Mathematical Physics

## A Modern Introduction to Its Foundations

With 152 Figures



Springer

ODTÜ KÜTÜPHANESİ  
M. E. T. U. LIBRARY

Sadri Hassani  
Department of Physics  
Illinois State University  
Normal, IL 61790  
USA  
hassani@entropy.phy.ilstu.edu

QC 20  
H394  
c.2

QC20 H394

METU LIBRARY

c.2

Mathematical physics : a modern



0020227569

*To my wife Sarah*

*and to my children*

*Dane Arash and Daisy Bitá*

**336417**

Library of Congress Cataloging-in-Publication Data  
Hassani, Sadri.

Mathematical physics : a modern introduction its foundations /  
Sadri Hassani.

p. cm.

Includes bibliographical references and index.

ISBN 0-387-98579-4 (alk. paper)

1. Mathematical physics. I. Title.

QC20.H394 1998

530.15—dc21

98-24738

Printed on acid-free paper.

© 1999 Springer-Verlag New York, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Production managed by Karina Mikhli; manufacturing supervised by Thomas King.

Photocomposed copy prepared from the author's TeX files.

Printed and bound by Hamilton Printing Co., Rensselaer, NY.

Printed in the United States of America.

9 8 7 6 5 4 3 (Corrected third printing, 2002)

ISBN 0-387-98579-4

SPIN 10854281

Springer-Verlag New York Berlin Heidelberg

A member of BertelsmannSpringer Science+Business Media GmbH

---

# Preface

“Ich kann es nun einmal nicht lassen, in diesem Drama von Mathematik und Physik—die sich im Dunkeln befruchten, aber von Angesicht zu Angesicht so gerne einander verkennen und verleugnen—die Rolle des (wie ich genügsam erfuhr, oft unerwünschten) *Boten* zu spielen.”

Hermann Weyl

It is said that mathematics is the language of Nature. If so, then physics is its poetry. Nature started to whisper into our ears when Egyptians and Babylonians were compelled to invent and use mathematics in their day-to-day activities. The faint geometric and arithmetical pidgin of over four thousand years ago, suitable for rudimentary conversations with nature as applied to simple landscaping, has turned into a sophisticated language in which the heart of matter is articulated.

The interplay between mathematics and physics needs no emphasis. What may need to be emphasized is that mathematics is not merely a tool with which the presentation of physics is facilitated, but the only medium in which physics can survive. Just as language is the means by which humans can express their thoughts and without which they lose their unique identity, mathematics is the only language through which physics can express itself and without which it loses its identity. And just as language is perfected due to its constant usage, mathematics develops in the most dramatic way because of its usage in physics. The quotation by Weyl above, an approximation to whose translation is “*In this drama of mathematics and physics—which fertilize each other in the dark, but which prefer to deny and misconstrue each other face to face—I cannot, however, resist playing the role of a messenger, albeit, as I have abundantly learned, often an unwelcome one.*”

is a perfect description of the natural intimacy between what mathematicians and physicists do, and the unnatural estrangement between the two camps. Some of the most beautiful mathematics has been motivated by physics (differential equations by Newtonian mechanics, differential geometry by general relativity, and operator theory by quantum mechanics), and some of the most fundamental physics has been expressed in the most beautiful poetry of mathematics (mechanics in symplectic geometry, and fundamental forces in Lie group theory).

I do not want to give the impression that mathematics and physics cannot develop independently. On the contrary, it is precisely the independence of each discipline that reinforces not only itself, but the other discipline as well—just as the study of the grammar of a language improves its usage and vice versa. However, the most effective means by which the two camps can accomplish great success is through an intense dialogue. Fortunately, with the advent of gauge and string theories of particle physics, such a dialogue has been reestablished between physics and mathematics after a relatively long lull.

## **Level and Philosophy of Presentation**

This is a book for physics students interested in the mathematics they use. It is also a book for mathematics students who wish to see some of the abstract ideas with which they are familiar come alive in an applied setting. The level of presentation is that of an advanced undergraduate or beginning graduate course (or sequence of courses) traditionally called “Mathematical Methods of Physics” or some variation thereof. Unlike most existing mathematical physics books intended for the same audience, which are usually lexicographic collections of facts about the diagonalization of matrices, tensor analysis, Legendre polynomials, contour integration, etc., with little emphasis on formal and systematic development of topics, this book attempts to strike a balance between formalism and application, between the abstract and the concrete.

I have tried to include as much of the essential formalism as is necessary to render the book optimally coherent and self-contained. This entails stating and proving a large number of theorems, propositions, lemmas, and corollaries. The benefit of such an approach is that the student will recognize clearly both the power and the limitation of a mathematical idea used in physics. There is a tendency on the part of the novice to universalize the mathematical methods and ideas encountered in physics courses because the limitations of these methods and ideas are not clearly pointed out.

There is a great deal of freedom in the topics and the level of presentation that instructors can choose from this book. My experience has shown that Parts I, II, III, Chapter 12, selected sections of Chapter 13, and selected sections or examples of Chapter 19 (or a large subset of all this) will be a reasonable course content for advanced undergraduates. If one adds Chapters 14 and 20, as well as selected topics from Chapters 21 and 22, one can design a course suitable for first-year graduate

students. By judicious choice of topics from Parts VII and VIII, the instructor can bring the content of the course to a more modern setting. Depending on the sophistication of the students, this can be done either in the first year or the second year of graduate school.

## Features

To better understand theorems, propositions, and so forth, students need to see them in action. There are over 350 worked-out examples and over 850 problems (many with detailed hints) in this book, providing a vast arena in which students can watch the formalism unfold. The philosophy underlying this abundance can be summarized as “An example is worth a thousand words of explanation.” Thus, whenever a statement is intrinsically vague or hard to grasp, worked-out examples and/or problems with hints are provided to clarify it. The inclusion of such a large number of examples is the means by which the balance between formalism and application has been achieved. However, although applications are essential in understanding mathematical physics, they are only one side of the coin. The theorems, propositions, lemmas, and corollaries, being highly condensed versions of knowledge, are equally important.

A conspicuous feature of the book, which is not emphasized in other comparable books, is the attempt to exhibit—as much as it is useful and applicable—interrelationships among various topics covered. Thus, the underlying theme of a vector space (which, in my opinion, is the most primitive concept at this level of presentation) recurs throughout the book and alerts the reader to the connection between various seemingly unrelated topics.

Another useful feature is the presentation of the historical setting in which men and women of mathematics and physics worked. I have gone against the trend of the “ahistoricism” of mathematicians and physicists by summarizing the life stories of the people behind the ideas. Many a time, the anecdotes and the historical circumstances in which a mathematical or physical idea takes form can go a long way toward helping us understand and appreciate the idea, especially if the interaction among—and the contributions of—all those having a share in the creation of the idea is pointed out, and the historical continuity of the development of the idea is emphasized.

To facilitate reference to them, all mathematical statements (definitions, theorems, propositions, lemmas, corollaries, and examples) have been numbered consecutively within each section and are preceded by the section number. For example, **4.2.9 Definition** indicates the ninth mathematical statement (which happens to be a definition) in Section 4.2. The end of a proof is marked by an empty square  $\square$ , and that of an example by a filled square  $\blacksquare$ , placed at the right margin of each.

Finally, a comprehensive index, a large number of marginal notes, and many explanatory underbraced and overbraced comments in equations facilitate the use

and comprehension of the book. In this respect, the book is also useful as a reference.

## Organization and Topical Coverage

Aside from Chapter 0, which is a collection of purely mathematical concepts, the book is divided into eight parts. Part I, consisting of the first four chapters, is devoted to a thorough study of finite-dimensional vector spaces and linear operators defined on them. As the unifying theme of the book, vector spaces demand careful analysis, and Part I provides this in the more accessible setting of finite dimension in a language that is conveniently generalized to the more relevant infinite dimensions, the subject of the next part.

Following a brief discussion of the technical difficulties associated with infinity, Part II is devoted to the two main infinite-dimensional vector spaces of mathematical physics: the classical orthogonal polynomials, and Fourier series and transform.

Complex variables appear in Part III. Chapter 9 deals with basic properties of complex functions, complex series, and their convergence. Chapter 10 discusses the calculus of residues and its application to the evaluation of definite integrals. Chapter 11 deals with more advanced topics such as multivalued functions, analytic continuation, and the method of steepest descent.

Part IV treats mainly ordinary differential equations. Chapter 12 shows how ordinary differential equations of second order arise in physical problems, and Chapter 13 consists of a formal discussion of these differential equations as well as methods of solving them numerically. Chapter 14 brings in the power of complex analysis to a treatment of the hypergeometric differential equation. The last chapter of this part deals with the solution of differential equations using integral transforms.

Part V starts with a formal chapter on the theory of operator and their spectral decomposition in Chapter 16. Chapter 17 focuses on a specific type of operator, namely the integral operators and their corresponding integral equations. The formalism and applications of Sturm-Liouville theory appear in Chapters 18 and 19, respectively.

The entire Part VI is devoted to a discussion of Green's functions. Chapter 20 introduces these functions for ordinary differential equations, while Chapters 21 and 22 discuss the Green's functions in an  $m$ -dimensional Euclidean space. Some of the derivations in these last two chapters are new and, as far as I know, unavailable anywhere else.

Parts VII and VIII contain a thorough discussion of Lie groups and their applications. The concept of group is introduced in Chapter 23. The theory of group representation, with an eye on its application in quantum mechanics, is discussed in the next chapter. Chapters 25 and 26 concentrate on tensor algebra and tensor analysis on manifolds. In Part VIII, the concepts of group and manifold are

brought together in the context of Lie groups. Chapter 27 discusses Lie groups and their algebras as well as their representations, with special emphasis on their application in physics. Chapter 28 is on differential geometry including a brief introduction to general relativity. Lie's original motivation for constructing the groups that bear his name is discussed in Chapter 29 in the context of a systematic treatment of differential equations using their symmetry groups. The book ends in a chapter that blends many of the ideas developed throughout the previous parts in order to treat variational problems and their symmetries. It also provides a most fitting example of the claim made at the beginning of this preface and one of the most beautiful results of mathematical physics: Noether's theorem on the relation between symmetries and conservation laws.

## Acknowledgments

It gives me great pleasure to thank all those who contributed to the making of this book. George Rutherford was kind enough to volunteer for the difficult task of condensing hundreds of pages of biography into tens of extremely informative pages. Without his help this unique and valuable feature of the book would have been next to impossible to achieve. I thank him wholeheartedly. Rainer Grobe and Qichang Su helped me with my rusty computational skills. (R. G. also helped me with my rusty German!) Many colleagues outside my department gave valuable comments and stimulating words of encouragement on the earlier version of the book. I would like to record my appreciation to Neil Rasband for reading part of the manuscript and commenting on it. Special thanks go to Tom von Foerster, senior editor of physics and mathematics at Springer-Verlag, not only for his patience and support, but also for the extreme care he took in reading the entire manuscript and giving me invaluable advice as a result. Needless to say, the ultimate responsibility for the content of the book rests on me. Last but not least, I thank my wife, Sarah, my son, Dane, and my daughter, Daisy, for the time taken away from them while I was writing the book, and for their support during the long and arduous writing process.

Many excellent textbooks, too numerous to cite individually here, have influenced the writing of this book. The following, however, are noteworthy for both their excellence and the amount of their influence:

Birkhoff, G., and G.-C. Rota, *Ordinary Differential Equations*, 3rd ed., New York, Wiley, 1978.

Bishop, R., and S. Goldberg, *Tensor Analysis on Manifolds*, New York, Dover, 1980.

Dennery, P., and A. Krzywicki, *Mathematics for Physicists*, New York, Harper & Row, 1967.

Halmos, P., *Finite-Dimensional Vector Spaces*, 2nd ed., Princeton, Van Nostrand, 1958.

Hamermesh, M. *Group Theory and its Application to Physical Problems*, Dover, New York, 1989.

Olver, P. *Application of Lie Groups to Differential Equations*, New York, Springer-Verlag, 1986.

Unless otherwise indicated, all biographical sketches have been taken from the following three sources:

Gillispie, C., ed., *Dictionary of Scientific Biography*, Charles Scribner's, New York, 1970.

Simmons, G. *Calculus Gems*, New York, McGraw-Hill, 1992.

History of Mathematics archive at [www-groups.dcs.st-and.ac.uk:80](http://www-groups.dcs.st-and.ac.uk:80).

I would greatly appreciate any comments and suggestions for improvements. Although extreme care was taken to correct all the misprints, the mere volume of the book makes it very likely that I have missed some (perhaps many) of them. I shall be most grateful to those readers kind enough to bring to my attention any remaining mistakes, typographical or otherwise. Please feel free to contact me.

Sadri Hassani  
Campus Box 4560  
Department of Physics  
Illinois State University  
Normal, IL 61790-4560, USA  
e-mail: [hassani@entropy.phyilstu.edu](mailto:hassani@entropy.phyilstu.edu)

It is my pleasure to thank all those readers who pointed out typographical mistakes and suggested a few clarifying changes. With the exception of a couple that required substantial revision, I have incorporated all the corrections and suggestions in this second printing.



---

# Note to the Reader

Mathematics and physics are like the game of chess (or, for that matter, like any game)—you will learn only by “playing” them. No amount of reading about the game will make you a master. In this book you will find a large number of examples and problems. Go through as many examples as possible, and try to reproduce them. Pay particular attention to sentences like “The reader may check . . . ” or “It is straightforward to show . . . ” These are red flags warning you that for a good understanding of the material at hand, you need to provide the missing steps. The problems often fill in missing steps as well; and in this respect they are essential for a thorough understanding of the book. Do not get discouraged if you cannot get to the solution of a problem at your first attempt. If you start from the beginning and think about each problem hard enough, you *will* get to the solution, and you will see that the subsequent problems will not be as difficult.

The extensive index makes the specific topics about which you may be interested to learn easily accessible. Often the marginal notes will help you easily locate the index entry you are after.

I have included a large collection of biographical sketches of mathematical physicists of the past. These are truly inspiring stories, and I encourage you to read them. They let you see that even under excruciating circumstances, the human mind can work miracles. You will discover how these remarkable individuals overcame the political, social, and economic conditions of their time to let us get a faint glimpse of the truth. They are our true heroes.



---

# Contents

<b>Preface</b>	<b>v</b>
<b>Note to the Reader</b>	<b>xi</b>
<b>List of Symbols</b>	<b>xxi</b>
<b>0 Mathematical Preliminaries</b>	<b>1</b>
0.1 Sets . . . . .	1
0.2 Maps . . . . .	4
0.3 Metric Spaces . . . . .	7
0.4 Cardinality . . . . .	10
0.5 Mathematical Induction . . . . .	12
0.6 Problems . . . . .	14
<b>I Finite-Dimensional Vector Spaces</b>	<b>17</b>
<b>1 Vectors and Transformations</b>	<b>19</b>
1.1 Vector Spaces . . . . .	19
1.2 Inner Product . . . . .	23
1.3 Linear Transformations . . . . .	32
1.4 Algebras . . . . .	41
1.5 Problems . . . . .	44
<b>2 Operator Algebra</b>	<b>49</b>
2.1 Algebra of $\mathcal{L}(V)$ . . . . .	49

2.2	Derivatives of Functions of Operators . . . . .	56
2.3	Conjugation of Operators . . . . .	61
2.4	Hermitian and Unitary Operators . . . . .	63
2.5	Projection Operators . . . . .	67
2.6	Operators in Numerical Analysis . . . . .	70
2.7	Problems . . . . .	76
<b>3</b>	<b>Matrices: Operator Representations</b>	<b>82</b>
3.1	Matrices . . . . .	82
3.2	Operations on Matrices . . . . .	87
3.3	Orthonormal Bases . . . . .	89
3.4	Change of Basis and Similarity Transformation . . . . .	91
3.5	The Determinant . . . . .	93
3.6	The Trace . . . . .	101
3.7	Problems . . . . .	103
<b>4</b>	<b>Spectral Decomposition</b>	<b>109</b>
4.1	Direct Sums . . . . .	109
4.2	Invariant Subspaces . . . . .	112
4.3	Eigenvalues and Eigenvectors . . . . .	114
4.4	Spectral Decomposition . . . . .	117
4.5	Functions of Operators . . . . .	125
4.6	Polar Decomposition . . . . .	129
4.7	Real Vector Spaces . . . . .	130
4.8	Problems . . . . .	138
<b>II</b>	<b>Infinite-Dimensional Vector Spaces</b>	<b>143</b>
<b>5</b>	<b>Hilbert Spaces</b>	<b>145</b>
5.1	The Question of Convergence . . . . .	145
5.2	The Space of Square-Integrable Functions . . . . .	150
5.3	Problems . . . . .	157
<b>6</b>	<b>Generalized Functions</b>	<b>159</b>
6.1	Continuous Index . . . . .	159
6.2	Generalized Functions . . . . .	165
6.3	Problems . . . . .	169
<b>7</b>	<b>Classical Orthogonal Polynomials</b>	<b>172</b>
7.1	General Properties . . . . .	172
7.2	Classification . . . . .	175
7.3	Recurrence Relations . . . . .	176
7.4	Examples of Classical Orthogonal Polynomials . . . . .	179

7.5	Expansion in Terms of Orthogonal Polynomials . . . . .	186
7.6	Generating Functions . . . . .	190
7.7	Problems . . . . .	190
<b>8</b>	<b>Fourier Analysis</b>	<b>196</b>
8.1	Fourier Series . . . . .	196
8.2	The Fourier Transform . . . . .	208
8.3	Problems . . . . .	220
<b>III</b>	<b>Complex Analysis</b>	<b>225</b>
<b>9</b>	<b>Complex Calculus</b>	<b>227</b>
9.1	Complex Functions . . . . .	227
9.2	Analytic Functions . . . . .	228
9.3	Conformal Maps . . . . .	236
9.4	Integration of Complex Functions . . . . .	241
9.5	Derivatives as Integrals . . . . .	248
9.6	Taylor and Laurent Series . . . . .	252
9.7	Problems . . . . .	263
<b>10</b>	<b>Calculus of Residues</b>	<b>270</b>
10.1	Residues . . . . .	270
10.2	Classification of Isolated Singularities . . . . .	273
10.3	Evaluation of Definite Integrals . . . . .	275
10.4	Problems . . . . .	290
<b>11</b>	<b>Complex Analysis: Advanced Topics</b>	<b>293</b>
11.1	Meromorphic Functions . . . . .	293
11.2	Multivalued Functions . . . . .	295
11.3	Analytic Continuation . . . . .	302
11.4	The Gamma and Beta Functions . . . . .	309
11.5	Method of Steepest Descent . . . . .	312
11.6	Problems . . . . .	319
<b>IV</b>	<b>Differential Equations</b>	<b>325</b>
<b>12</b>	<b>Separation of Variables in Spherical Coordinates</b>	<b>327</b>
12.1	PDEs of Mathematical Physics . . . . .	327
12.2	Separation of the Angular Part of the Laplacian . . . . .	331
12.3	Construction of Eigenvalues of $L^2$ . . . . .	334
12.4	Eigenvectors of $L^2$ : Spherical Harmonics . . . . .	338
12.5	Problems . . . . .	346

<b>13</b>	<b>Second-Order Linear Differential Equations</b>	<b>348</b>
13.1	General Properties of ODEs . . . . .	349
13.2	Existence and Uniqueness for First-Order DEs . . . . .	350
13.3	General Properties of SOLDEs . . . . .	352
13.4	The Wronskian . . . . .	355
13.5	Adjoint Differential Operators . . . . .	364
13.6	Power-Series Solutions of SOLDEs . . . . .	367
13.7	SOLDEs with Constant Coefficients . . . . .	376
13.8	The WKB Method . . . . .	380
13.9	Numerical Solutions of DEs . . . . .	383
13.10	Problems . . . . .	394
<b>14</b>	<b>Complex Analysis of SOLDEs</b>	<b>400</b>
14.1	Analytic Properties of Complex DEs . . . . .	401
14.2	Complex SOLDEs . . . . .	404
14.3	Fuchsian Differential Equations . . . . .	410
14.4	The Hypergeometric Function . . . . .	413
14.5	Confluent Hypergeometric Functions . . . . .	419
14.6	Problems . . . . .	426
<b>15</b>	<b>Integral Transforms and Differential Equations</b>	<b>433</b>
15.1	Integral Representation of the Hypergeometric Function . . . . .	434
15.2	Integral Representation of the Confluent Hypergeometric Function . . . . .	437
15.3	Integral Representation of Bessel Functions . . . . .	438
15.4	Asymptotic Behavior of Bessel Functions . . . . .	443
15.5	Problems . . . . .	445
<b>V</b>	<b>Operators on Hilbert Spaces</b>	<b>449</b>
<b>16</b>	<b>An Introduction to Operator Theory</b>	<b>451</b>
16.1	From Abstract to Integral and Differential Operators . . . . .	451
16.2	Bounded Operators in Hilbert Spaces . . . . .	453
16.3	Spectra of Linear Operators . . . . .	457
16.4	Compact Sets . . . . .	458
16.5	Compact Operators . . . . .	464
16.6	Spectrum of Compact Operators . . . . .	467
16.7	Spectral Theorem for Compact Operators . . . . .	473
16.8	Resolvents . . . . .	480
16.9	Problems . . . . .	485
<b>17</b>	<b>Integral Equations</b>	<b>488</b>
17.1	Classification . . . . .	488

17.2	Fredholm Integral Equations . . . . .	494
17.3	Problems . . . . .	505
<b>18</b>	<b>Sturm–Liouville Systems: Formalism</b>	<b>507</b>
18.1	Unbounded Operators with Compact Resolvent . . . . .	507
18.2	Sturm–Liouville Systems and SOLDEs . . . . .	513
18.3	Other Properties of Sturm–Liouville Systems . . . . .	517
18.4	Problems . . . . .	522
<b>19</b>	<b>Sturm–Liouville Systems: Examples</b>	<b>524</b>
19.1	Expansions in Terms of Eigenfunctions . . . . .	524
19.2	Separation in Cartesian Coordinates . . . . .	526
19.3	Separation in Cylindrical Coordinates . . . . .	535
19.4	Separation in Spherical Coordinates . . . . .	540
19.5	Problems . . . . .	545
<b>VI</b>	<b>Green’s Functions</b>	<b>551</b>
<b>20</b>	<b>Green’s Functions in One Dimension</b>	<b>553</b>
20.1	Calculation of Some Green’s Functions . . . . .	554
20.2	Formal Considerations . . . . .	557
20.3	Green’s Functions for SOLDOs . . . . .	565
20.4	Eigenfunction Expansion of Green’s Functions . . . . .	577
20.5	Problems . . . . .	580
<b>21</b>	<b>Multidimensional Green’s Functions: Formalism</b>	<b>583</b>
21.1	Properties of Partial Differential Equations . . . . .	584
21.2	Multidimensional GFs and Delta Functions . . . . .	592
21.3	Formal Development . . . . .	596
21.4	Integral Equations and GFs . . . . .	600
21.5	Perturbation Theory . . . . .	603
21.6	Problems . . . . .	610
<b>22</b>	<b>Multidimensional Green’s Functions: Applications</b>	<b>613</b>
22.1	Elliptic Equations . . . . .	613
22.2	Parabolic Equations . . . . .	621
22.3	Hyperbolic Equations . . . . .	626
22.4	The Fourier Transform Technique . . . . .	628
22.5	The Eigenfunction Expansion Technique . . . . .	636
22.6	Problems . . . . .	641

<b>VII</b>	<b>Groups and Manifolds</b>	<b>649</b>
<b>23</b>	<b>Group Theory</b>	<b>651</b>
23.1	Groups . . . . .	652
23.2	Subgroups . . . . .	656
23.3	Group Action . . . . .	663
23.4	The Symmetric Group $S_n$ . . . . .	664
23.5	Problems . . . . .	669
<b>24</b>	<b>Group Representation Theory</b>	<b>673</b>
24.1	Definitions and Examples . . . . .	673
24.2	Orthogonality Properties . . . . .	680
24.3	Analysis of Representations . . . . .	685
24.4	Group Algebra . . . . .	687
24.5	Relationship of Characters to Those of a Subgroup . . . . .	692
24.6	Irreducible Basis Functions . . . . .	695
24.7	Tensor Product of Representations . . . . .	699
24.8	Representations of the Symmetric Group . . . . .	707
24.9	Problems . . . . .	723
<b>25</b>	<b>Algebra of Tensors</b>	<b>728</b>
25.1	Multilinear Mappings . . . . .	729
25.2	Symmetries of Tensors . . . . .	736
25.3	Exterior Algebra . . . . .	739
25.4	Inner Product Revisited . . . . .	749
25.5	The Hodge Star Operator . . . . .	756
25.6	Problems . . . . .	758
<b>26</b>	<b>Analysis of Tensors</b>	<b>763</b>
26.1	Differentiable Manifolds . . . . .	763
26.2	Curves and Tangent Vectors . . . . .	770
26.3	Differential of a Map . . . . .	776
26.4	Tensor Fields on Manifolds . . . . .	780
26.5	Exterior Calculus . . . . .	791
26.6	Symplectic Geometry . . . . .	801
26.7	Problems . . . . .	808
<b>VIII</b>	<b>Lie Groups and Their Applications</b>	<b>813</b>
<b>27</b>	<b>Lie Groups and Lie Algebras</b>	<b>815</b>
27.1	Lie Groups and Their Algebras . . . . .	815
27.2	An Outline of Lie Algebra Theory . . . . .	833
27.3	Representation of Compact Lie Groups . . . . .	845



---

27.4	Representation of the General Linear Group . . . . .	856
27.5	Representation of Lie Algebras . . . . .	859
27.6	Problems . . . . .	876
<b>28</b>	<b>Differential Geometry</b>	<b>882</b>
28.1	Vector Fields and Curvature . . . . .	883
28.2	Riemannian Manifolds . . . . .	887
28.3	Covariant Derivative and Geodesics . . . . .	897
28.4	Isometries and Killing Vector Fields . . . . .	908
28.5	Geodesic Deviation and Curvature . . . . .	913
28.6	General Theory of Relativity . . . . .	918
28.7	Problems . . . . .	932
<b>29</b>	<b>Lie Groups and Differential Equations</b>	<b>936</b>
29.1	Symmetries of Algebraic Equations . . . . .	936
29.2	Symmetry Groups of Differential Equations . . . . .	941
29.3	The Central Theorems . . . . .	951
29.4	Application to Some Known PDEs . . . . .	956
29.5	Application to ODEs . . . . .	964
29.6	Problems . . . . .	970
<b>30</b>	<b>Calculus of Variations, Symmetries, and Conservation Laws</b>	<b>973</b>
30.1	The Calculus of Variations . . . . .	973
30.2	Symmetry Groups of Variational Problems . . . . .	988
30.3	Conservation Laws and Noether's Theorem . . . . .	992
30.4	Application to Classical Field Theory . . . . .	997
30.5	Problems . . . . .	1000
	<b>Bibliography</b>	<b>1003</b>
	<b>Index</b>	<b>1007</b>



---

# List of Symbols

$\in, (\notin)$	“belongs to”, (“does not belong to”)
$\mathbb{Z}$	Set of integers
$\mathbb{R}$	Set of real numbers
$\mathbb{R}^+$	Set of positive real numbers
$\mathbb{C}$	Set of complex numbers
$\mathbb{N}$	Set of nonnegative integers
$\mathbb{Q}$	Set of rational numbers
$\sim A$	Complement of the set $A$
$A \times B$	Set of ordered pairs $(a, b)$ with $a \in A$ and $b \in B$
$A^n$	$\{(a_1, a_2, \dots, a_n)   a_i \in A\}$
$\cup, (\cap)$	Union, (Intersection)
$A \equiv B$	$A$ is equivalent to $B$
$x \mapsto f(x)$	$x$ is mapped to $f(x)$ via the map $f$
$\forall$	for all (values of)
$\exists$	There exists (a value of)
$[[a]]$	Equivalence class to which $a$ belongs
$g \circ f$	Composition of maps $f$ and $g$
<i>iff</i>	if and only if
$\mathcal{C}^k(a, b)$	Set of functions on $(a, b)$ with continuous derivatives up to order $k$
$\mathbb{C}^n$ (or $\mathbb{R}^n$ )	Set of complex (or real) $n$ -tuples
$\mathcal{P}^{\mathbb{C}}[t]$	Set of polynomials in $t$ with complex coefficients
$\mathcal{P}^{\mathbb{R}}[t]$	Set of polynomials in $t$ with real coefficients
$\mathcal{P}_n^{\mathbb{C}}[t]$	Set of polynomials with complex coefficients of degree $n$ or less
$\mathbb{C}^\infty$	Set of all complex sequences $\{\alpha_i\}_{i=1}^\infty$ such that $\sum_{i=1}^\infty  \alpha_i ^2 < \infty$
$\langle a   b \rangle$	Inner product of $ a\rangle$ and $ b\rangle$
$\ a\ $	Norm (length) of the vector $ a\rangle$

$\mathcal{L}(\mathcal{V})$	Set of endomorphisms (linear operators) on vector space $\mathcal{V}$
$[\mathbf{S}, \mathbf{T}]$	Commutator of operators $\mathbf{S}$ and $\mathbf{T}$
$\mathbf{T}^\dagger$	Adjoint (hermitian conjugate) of operator $\mathbf{T}$
$A^t$ , or $\bar{A}$	Transpose of matrix $A$
$\mathcal{U} \oplus \mathcal{V}$	Direct sum of vector spaces $\mathcal{U}$ and $\mathcal{V}$
$\delta(x - x_0)$	Dirac delta function nonvanishing only at $x = x_0$
$\text{Res}[f(z_0)]$	Residue of $f$ at point $z_0$
DE, ODE, PDE	Differential equation, Ordinary DE, Partial DE
SOLDE	Second order linear (ordinary) differential equation
$GL(\mathcal{V})$	Set of all invertible operators on vector space $\mathcal{V}$
$GL(n, \mathbb{C})$	Set of all $n \times n$ complex matrices of nonzero determinant
$SL(n, \mathbb{C})$	Set of all $n \times n$ complex matrices of unit determinant
$\tau_1 \otimes \tau_2$	Tensor product of $\tau_1$ and $\tau_2$
$\mathbf{A} \wedge \mathbf{B}$	Exterior (wedge) product of skew-symmetric tensors $\mathbf{A}$ and $\mathbf{B}$
$\Lambda^p(\mathcal{V})$	Set of all skew-symmetric tensors of type $(p, 0)$ on $\mathcal{V}$

# Mathematical Preliminaries

This introductory chapter gathers together some of the most basic tools and notions that are used throughout the book. It also introduces some common vocabulary and notations used in modern mathematical physics literature. Readers familiar with such concepts as sets, maps, equivalence relations, and metric spaces may wish to skip this chapter.

## 0.1 Sets

concept of set  
elaborated

Modern mathematics starts with the basic (and undefinable) concept of **set**. We think of a set as a structureless family, or collection, of objects. We speak, for example, of the set of students in a college, of men in a city, of women working for a corporation, of vectors in space, of points in a plane, or of events in the continuum of space-time. Each member  $a$  of a set  $A$  is called an **element** of that set. This relation is denoted by  $a \in A$  (read “ $a$  is an element of  $A$ ” or “ $a$  belongs to  $A$ ”), and its negation by  $a \notin A$ . Sometimes  $a$  is called a **point** of the set  $A$  to emphasize a geometric connotation.

A set is usually designated by enumeration of its elements between braces. For example,  $\{2, 4, 6, 8\}$  represents the set consisting of the first four even natural numbers;  $\{0, \pm 1, \pm 2, \pm 3, \dots\}$  is the set of all integers;  $\{1, x, x^2, x^3, \dots\}$  is the set of all nonnegative powers of  $x$ ; and  $\{1, i, -1, -i\}$  is the set of the four complex fourth roots of unity. In many cases, a set is defined by a (mathematical) statement that holds for all of its elements. Such a set is generally denoted by  $\{x | P(x)\}$  and read “the set of all  $x$ ’s such that  $P(x)$  is true.” The foregoing examples of sets can be written alternatively as follows:

$$\{n \mid n \text{ is even and } 1 < n < 9\}$$

$$\{\pm n \mid n \text{ is a natural number}\}$$

$$\{y \mid y = x^n \text{ and } n \text{ is a natural number}\}$$

$$\{z \mid z^4 = 1 \text{ and } z \text{ is a complex number}\}$$

In a frequently used shorthand notation, the last two sets can be abbreviated as  $\{x^n \mid n \geq 0 \text{ and } n \text{ is an integer}\}$  and  $\{z \in \mathbb{C} \mid z^4 = 1\}$ . Similarly, the unit circle can be denoted by  $\{z \mid |z| = 1\}$ , the closed interval  $[a, b]$  as  $\{x \mid a \leq x \leq b\}$ , the open interval  $(a, b)$  as  $\{x \mid a < x < b\}$ , and the set of all nonnegative powers of  $x$  as  $\{x^n\}_{n=0}^{\infty}$ . This last notation will be used frequently in this book. A set with a single element is called a **singleton**.

singleton

If  $a \in A$  whenever  $a \in B$ , we say that  $B$  is a **subset** of  $A$  and write  $B \subset A$  or  $A \supset B$ . If  $B \subset A$  and  $A \subset B$ , then  $A = B$ . If  $B \subset A$  and  $A \neq B$ , then  $B$  is called a **proper subset** of  $A$ . The set defined by  $\{a \mid a \neq a\}$  is called the **empty set** and is denoted by  $\emptyset$ . Clearly,  $\emptyset$  contains no elements and is a subset of any arbitrary set. The collection of all subsets (including  $\emptyset$ ) of a set  $A$  is denoted by  $2^A$ . The reason for this notation is that the number of subsets of a set containing  $n$  elements is  $2^n$  (Problem 0.1). If  $A$  and  $B$  are sets, their **union**, denoted by  $A \cup B$ , is the set containing all elements that belong to  $A$  or  $B$  or both. The **intersection** of the sets  $A$  and  $B$ , denoted by  $A \cap B$ , is the set containing all elements belonging to both  $A$  and  $B$ . If  $\{B_\alpha\}_{\alpha \in I}$  is a collection of sets,<sup>1</sup> we denote their union by  $\cup_\alpha B_\alpha$  and their intersection by  $\cap_\alpha B_\alpha$ .

(proper) subset  
empty setunion, intersection,  
complement

universal set

In any application of set theory there is an underlying **universal set** whose subsets are the objects of study. This universal set is usually clear from the context. For example, in the study of the properties of integers, the set of integers, denoted by  $\mathbb{Z}$ , is the universal set. The set of reals,  $\mathbb{R}$ , is the universal set in real analysis, and the set of complex numbers,  $\mathbb{C}$ , is the universal set in complex analysis. With a universal set  $X$  in mind, one can write  $X \sim A$  instead of  $\sim A$ . The **complement** of a set  $A$  is denoted by  $\sim A$  and defined as

$$\sim A \equiv \{a \mid a \notin A\}.$$

The complement of  $B$  in  $A$  (or their difference) is

$$A \sim B \equiv \{a \mid a \in A \text{ and } a \notin B\}.$$

Cartesian product  
ordered pairs

From two given sets  $A$  and  $B$ , it is possible to form the **Cartesian product** of  $A$  and  $B$ , denoted by  $A \times B$ , which is the set of **ordered pairs**  $(a, b)$ , where  $a \in A$  and  $b \in B$ . This is expressed in set-theoretic notation as

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}.$$

<sup>1</sup>Here  $I$  is an index set—or a counting set—with its typical element denoted by  $\alpha$ . In most cases,  $I$  is the set of (nonnegative) integers, but, in principle, it can be any set, for example, the set of real numbers.

We can generalize this to an arbitrary number of sets. If  $A_1, A_2, \dots, A_n$  are sets, then the Cartesian product of these sets is

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \dots, a_n) \mid a_i \in A_i\},$$

which is a set of ordered  $n$ -tuples. If  $A_1 = A_2 = \cdots = A_n = A$ , then we write  $A^n$  instead of  $A \times A \times \cdots \times A$ , and

$$A^n = \{(a_1, a_2, \dots, a_n) \mid a_i \in A\}.$$

The most familiar example of a Cartesian product occurs when  $A = \mathbb{R}$ . Then  $\mathbb{R}^2$  is the set of pairs  $(x_1, x_2)$  with  $x_1, x_2 \in \mathbb{R}$ . This is simply the points in the Euclidean plane. Similarly,  $\mathbb{R}^3$  is the set of triplets  $(x_1, x_2, x_3)$ , or the points in space, and  $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R}\}$  is the set of real  $n$ -tuples.

### 0.1.1 Equivalence Relations

There are many instances in which the elements of a set are naturally grouped together. For example, all vector potentials that differ by the gradient of a scalar function can be grouped together because they all give the same magnetic field. Similarly, all quantum state functions (of unit "length") that differ by a multiplicative complex number of unit length can be grouped together because they all represent the same physical state. The abstraction of these ideas is summarized in the following definition.

relation and  
equivalence relation

**0.1.1. Definition.** Let  $A$  be a set. A **relation** on  $A$  is a comparison test between ordered pairs of elements of  $A$ . If the pair  $(a, b) \in A \times A$  pass this test, we write  $a \triangleright b$  and read "a is related to b." An **equivalence relation** on  $A$  is a relation that has the following properties:

$$a \triangleright a \quad \forall a \in A, \quad (\text{reflexivity})$$

$$a \triangleright b \Rightarrow b \triangleright a \quad a, b \in A, \quad (\text{symmetry})$$

$$a \triangleright b, b \triangleright c \Rightarrow a \triangleright c \quad a, b, c \in A, \quad (\text{transitivity})$$

When  $a \triangleright b$ , we say that "a is equivalent to b." The set  $\llbracket a \rrbracket = \{b \in A \mid b \triangleright a\}$  of all elements that are equivalent to  $a$  is called the **equivalence class** of  $a$ .

equivalence class

The reader may verify the following property of equivalence relations.

**0.1.2. Proposition.** If  $\triangleright$  is an equivalence relation on  $A$  and  $a, b \in A$ , then either  $\llbracket a \rrbracket \cap \llbracket b \rrbracket = \emptyset$  or  $\llbracket a \rrbracket = \llbracket b \rrbracket$ .

Therefore,  $a' \in \llbracket a \rrbracket$  implies that  $\llbracket a' \rrbracket = \llbracket a \rrbracket$ . In other words, any element of an equivalence class can be chosen to be a **representative** of that class. Because of the symmetry of equivalence relations, sometimes we denote them by  $\bowtie$ .

representative of an  
equivalence class

**0.1.3. Example.** Let  $A$  be the set of human beings. Let  $a \triangleright b$  be interpreted as “ $a$  is older than  $b$ .” Then clearly,  $\triangleright$  is a relation but not an equivalence relation. On the other hand, if we interpret  $a \triangleright b$  as “ $a$  and  $b$  have the same paternal grandfather,” then  $\triangleright$  is an equivalence relation, as the reader may check. The equivalence class of  $a$  is the set of all grandchildren of  $a$ ’s paternal grandfather.

Let  $V$  be the set of vector potentials. Write  $A \triangleright A'$  if  $A - A' = \nabla f$  for some function  $f$ . The reader may verify that  $\triangleright$  is an equivalence relation, and that  $\llbracket A \rrbracket$  is the set of all vector potentials giving rise to the same magnetic field.

Let the underlying set be  $\mathbb{Z} \times (\mathbb{Z} \sim \{0\})$ . Say “ $(a, b)$  is related to  $(c, d)$ ” if  $ad = bc$ . Then this relation is an equivalence relation. Furthermore,  $\llbracket (a, b) \rrbracket$  can be identified as the ratio  $a/b$ . ■

**0.1.4. Definition.** Let  $A$  be a set and  $\{B_\alpha\}$  a collection of subsets of  $A$ . We say that  $\{B_\alpha\}$  is a *partition* of  $A$ , or  $\{B_\alpha\}$  *partitions*  $A$ , if the  $B_\alpha$ ’s are disjoint, i.e., have no element in common, and  $\cup_\alpha B_\alpha = A$ .

Now consider the collection  $\{\llbracket a \rrbracket \mid a \in A\}$  of all equivalence classes of  $A$ . These classes are disjoint, and evidently their union covers all of  $A$ . Therefore, the collection of equivalence classes of  $A$  is a partition of  $A$ . This collection is denoted by  $A / \triangleright$  and is called the **quotient set** of  $A$  under the equivalence relation  $\triangleright$ .

**0.1.5. Example.** Let the underlying set be  $\mathbb{R}^3$ . Define an equivalence relation on  $\mathbb{R}^3$  by saying that  $P_1 \in \mathbb{R}^3$  and  $P_2 \in \mathbb{R}^3$  are equivalent if they lie on the same line passing through the origin. Then  $\mathbb{R}^3 / \triangleright$  is the set of all lines in space passing through the origin. If we choose the unit vector with positive third coordinate along a given line as the representative of that line, then  $\mathbb{R}^3 / \triangleright$  can be identified with the upper unit hemisphere.<sup>2</sup>  $\mathbb{R}^3 / \triangleright$  is called the **projective space** associated with  $\mathbb{R}^3$ .

On the set  $\mathbb{Z}$  of integers define a relation by writing  $m \triangleright n$  for  $m, n \in \mathbb{Z}$  if  $m - n$  is divisible by  $k$ , where  $k$  is a fixed integer. Then  $\triangleright$  is not only a relation, but an equivalence relation. In this case, we have

$$\mathbb{Z} / \triangleright = \{\llbracket 0 \rrbracket, \llbracket 1 \rrbracket, \dots, \llbracket k - 1 \rrbracket\},$$

as the reader is urged to verify.

For the equivalence relation defined on  $\mathbb{Z} \times \mathbb{Z}$  of Example 0.1.3, the set  $\mathbb{Z} \times \mathbb{Z} / \triangleright$  can be identified with  $\mathbb{Q}$ , the set of rational numbers. ■

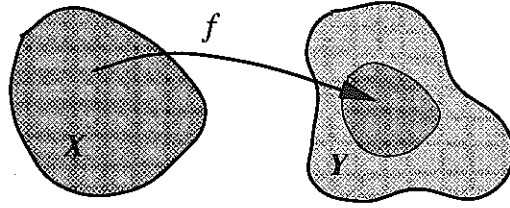
## 0.2 Maps

To communicate between sets, one introduces the concept of a map. A **map**  $f$  from a set  $X$  to a set  $Y$ , denoted by  $f : X \rightarrow Y$  or  $X \xrightarrow{f} Y$ , is a correspondence between elements of  $X$  and those of  $Y$  in which all the elements of  $X$  participate,

map, domain,  
codomain, image

<sup>2</sup>Furthermore, we need to identify any two points on the edge of the hemisphere which lie on the same diameter.





**Figure 1** The map  $f$  maps all of the set  $X$  onto a subset of  $Y$ . The shaded area in  $Y$  is  $f(X)$ , the range of  $f$ .

and each element of  $X$  corresponds to only one element of  $Y$  (see Figure 1). If  $y \in Y$  is the element that corresponds to  $x \in X$  via the map  $f$ , we write

$$y = f(x) \quad \text{or} \quad x \mapsto f(x) \quad \text{or} \quad x \xrightarrow{f} y$$

and call  $f(x)$  the **image** of  $x$  under  $f$ . Thus, by the definition of map,  $x \in X$  can have only one image. The set  $X$  is called the **domain**, and  $Y$  the **codomain** or the **target space**. Two maps  $f : X \rightarrow Y$  and  $g : X \rightarrow Y$  are said to be equal if  $f(x) = g(x)$  for all  $x \in X$ .

function

**0.2.1. Box.** A map whose codomain is the set of real numbers  $\mathbb{R}$  or the set of complex numbers  $\mathbb{C}$  is commonly called a **function**.

A special map that applies to all sets  $A$  is  $\text{id}_A : A \rightarrow A$ , called the **identity map** of  $A$ , and defined by

$$\text{id}_A(a) = a \quad \forall a \in A.$$

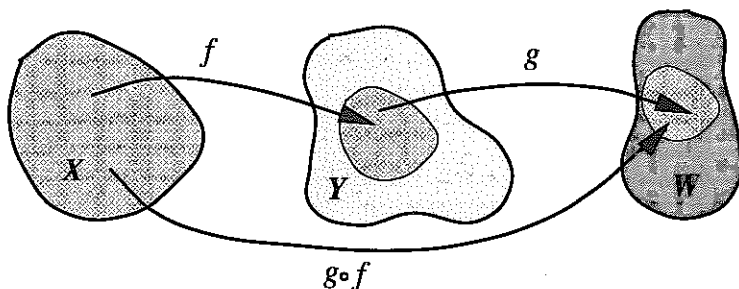
The **graph**  $\Gamma_f$  of a map  $f : A \rightarrow B$  is a subset of  $A \times B$  defined by

$$\Gamma_f = \{(a, f(a)) \mid a \in A\} \subset A \times B.$$

graph of a map

This general definition reduces to the ordinary graphs encountered in algebra and calculus where  $A = B = \mathbb{R}$  and  $A \times B$  is the  $xy$ -plane. If  $A$  is a subset of  $X$ , we call  $f(A) = \{f(x) \mid x \in A\}$  the **image** of  $A$ . Similarly, if  $B \subset f(X)$ , we call  $f^{-1}(B) = \{x \in X \mid f(x) \in B\}$  the **inverse image**, or **preimage**, of  $B$ . In words,  $f^{-1}(B)$  consists of all elements in  $X$  whose images are in  $B \subset Y$ . If  $B$  consists of a single element  $b$ , then  $f^{-1}(b) = \{x \in X \mid f(x) = b\}$  consists of all elements of  $X$  that are mapped to  $b$ . Note that it is possible for many points of  $X$  to have the same image in  $Y$ . The subset  $f(X)$  of the codomain of a map  $f$  is called the **range** of  $f$  (see Figure 1).

preimage



**Figure 2** The composition of two maps is another map.

If  $f : X \rightarrow Y$  and  $g : Y \rightarrow W$ , then the mapping  $h : X \rightarrow W$  given by  $h(x) = g(f(x))$  is called the **composition** of  $f$  and  $g$ , and is denoted by  $h = g \circ f$  (see Figure 2).<sup>3</sup> It is easy to verify that

$$f \circ \text{id}_X = f = \text{id}_Y \circ f$$

If  $f(x_1) = f(x_2)$  implies that  $x_1 = x_2$ , we call  $f$  **injective**, or *one-to-one* (denoted 1-1). For an injective map only one element of  $X$  corresponds to an element of  $Y$ . If  $f(X) = Y$ , the mapping is said to be **surjective**, or *onto*. A map that is both injective and surjective is said to be **bijective**, or to be a **one-to-one correspondence**. Two sets that are in one-to-one correspondence, have, by definition, the same number of elements. If  $f : X \rightarrow Y$  is a bijection from  $X$  onto  $Y$ , then for each  $y \in Y$  there is one and only one element  $x$  in  $X$  for which  $f(x) = y$ . Thus, there is a mapping  $f^{-1} : Y \rightarrow X$  given by  $f^{-1}(y) = x$ , where  $x$  is the unique element such that  $f(x) = y$ . This mapping is called the **inverse** of  $f$ . The inverse of  $f$  is also identified as the map that satisfies  $f \circ f^{-1} = \text{id}_Y$  and  $f^{-1} \circ f = \text{id}_X$ . For example, one can easily verify that  $\ln^{-1} = \exp$  and  $\exp^{-1} = \ln$ , because  $\ln(e^x) = x$  and  $e^{\ln x} = x$ .

Given a map  $f : X \rightarrow Y$ , we can define a relation  $\bowtie$  on  $X$  by saying  $x_1 \bowtie x_2$  if  $f(x_1) = f(x_2)$ . The reader may check that this is in fact an *equivalence* relation. The equivalence classes are subsets of  $X$  all of whose elements map to the same point in  $Y$ . In fact,  $\llbracket x \rrbracket = f^{-1}(f(x))$ . Corresponding to  $f$ , there is a map  $\tilde{f} : X/\bowtie \rightarrow Y$  given by  $\tilde{f}(\llbracket x \rrbracket) = f(x)$ . This map is injective because if  $\tilde{f}(\llbracket x_1 \rrbracket) = \tilde{f}(\llbracket x_2 \rrbracket)$ , then  $f(x_1) = f(x_2)$ , so  $x_1$  and  $x_2$  belong to the same equivalence class; therefore,  $\llbracket x_1 \rrbracket = \llbracket x_2 \rrbracket$ . It follows that  $\tilde{f} : X/\bowtie \rightarrow f(X)$  is bijective.

If  $f$  and  $g$  are both bijections with inverses  $f^{-1}$  and  $g^{-1}$ , respectively, then  $g \circ f$  also has an inverse, and verifying that  $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$  is straightforward.

<sup>3</sup>Note the importance of the order in which the composition is written. The reverse order may not even exist.

**0.2.2. Example.** As an example of the preimage of a set, consider the sine and cosine functions. Then it should be clear that

$$\sin^{-1} 0 = \{n\pi\}_{n=-\infty}^{\infty}, \quad \cos^{-1} 0 = \left\{\frac{\pi}{2} + n\pi\right\}_{n=-\infty}^{\infty}$$

Similarly,  $\sin^{-1}[0, \frac{1}{2}]$  consists of all the intervals on the  $x$ -axis marked by heavy line segments in Figure 3, i.e., all the points whose sine lies between 0 and  $\frac{1}{2}$ . ■

injectivity and surjectivity depend on the domain and codomain

As examples of maps, we consider functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  studied in calculus. The two functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $g : \mathbb{R} \rightarrow (-1, +1)$  given, respectively, by  $f(x) = x^3$  and  $g(x) = \tanh x$  are bijective. The latter function, by the way, shows that there are as many points in the whole real line as there are in the interval  $(-1, +1)$ . If we denote the set of positive real numbers by  $\mathbb{R}^+$ , then the function  $f : \mathbb{R} \rightarrow \mathbb{R}^+$  given by  $f(x) = x^2$  is surjective but not injective (both  $x$  and  $-x$  map to  $x^2$ ). The function  $g : \mathbb{R}^+ \rightarrow \mathbb{R}$  given by the same rule,  $g(x) = x^2$ , is injective but not surjective. On the other hand,  $h : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  again given by  $h(x) = x^2$  is bijective, but  $u : \mathbb{R} \rightarrow \mathbb{R}$  given by the same rule is neither injective nor surjective.

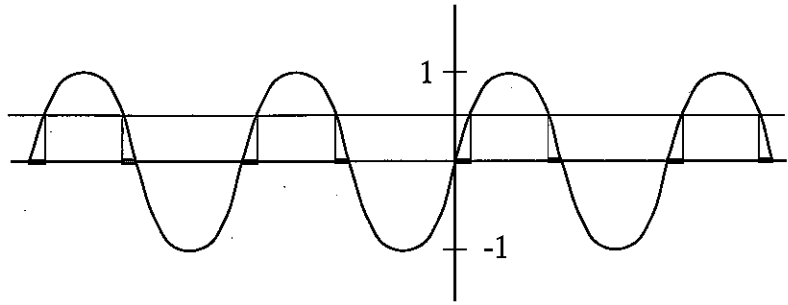
Let  $\mathcal{M}^{n \times n}$  denote the set of  $n \times n$  real matrices. Define a function  $\det : \mathcal{M}^{n \times n} \rightarrow \mathbb{R}$  by  $\det(A) = \det A$ , where  $\det A$  is the determinant of  $A$  for  $A \in \mathcal{M}^{n \times n}$ . This function is clearly surjective (why?) but not injective. The set of all matrices whose determinant is 1 is  $\det^{-1}(1)$ . Such matrices occur frequently in physical applications.

Another example of interest is  $f : \mathbb{C} \rightarrow \mathbb{R}$  given by  $f(z) = |z|$ . This function is also neither injective nor surjective. Here  $f^{-1}(1)$  is the **unit circle**, the circle of radius 1 in the complex plane.

The domain of a map can be a Cartesian product of a set, as in  $f : X \times X \rightarrow Y$ . Two specific cases are worthy of mention. The first is when  $Y = \mathbb{R}$ . An example of this case is the dot product on vectors. Thus, if  $X$  is the set of vectors in space, we can define  $f(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$ . The second case is when  $Y = X$ . Then  $f$  is called a **binary operation** on  $X$ , whereby an element in  $X$  is associated with two elements in  $X$ . For instance, let  $X = \mathbb{Z}$ , the set of all integers; then the function  $f : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$  defined by  $f(m, n) = mn$  is the binary operation of multiplication of integers. Similarly,  $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  given by  $g(x, y) = x + y$  is the binary operation of addition of real numbers.

### 0.3 Metric Spaces

Although sets are at the root of modern mathematics, they are only of formal and abstract interest by themselves. To make sets useful, it is necessary to introduce some structures on them. There are two general procedures for the implementation of such structures. These are the abstractions of the two major branches of mathematics—algebra and analysis.



**Figure 3** The union of all the intervals on the  $x$ -axis marked by heavy line segments is  $\sin^{-1}[0, \frac{1}{2}]$ .

We can turn a set into an algebraic structure by introducing a binary operation on it. For example, a vector space consists, among other things, of the binary operation of vector addition. A group is, among other things, a set together with the binary operation of “multiplication”. There are many other examples of algebraic systems, and they constitute the rich subject of algebra.

When analysis, the other branch of mathematics, is abstracted using the concept of sets, it leads to topology, in which the concept of continuity plays a central role. This is also a rich subject with far-reaching implications and applications. We shall not go into any details of these two areas of mathematics. Although some algebraic systems will be discussed and the ideas of limit and continuity will be used in the sequel, this will be done in an intuitive fashion, by introducing and employing the concepts when they are needed. On the other hand, some general concepts will be introduced when they require minimum prerequisites. One of these is a metric space:

**0.3.1. Definition.** A *metric space* is a set  $X$  together with a real-valued function  $d : X \times X \rightarrow \mathbb{R}$  such that

metric space defined

- (a)  $d(x, y) \geq 0 \quad \forall x, y$ , and  $d(x, y) = 0$  iff  $x = y$ .
- (b)  $d(x, y) = d(y, x)$ . (symmetry)
- (c)  $d(x, y) \leq d(x, z) + d(z, y)$ . (the triangle inequality)

It is worthwhile to point out that  $X$  is a completely arbitrary set and needs no other structure. In this respect Definition 0.3.1 is very broad and encompasses many different situations, as the following examples will show. Before examining the examples, note that the function  $d$  defined above is the abstraction of the notion of distance: (a) says that the distance between any two points is always nonnegative and is zero only if the two points coincide; (b) says that the distance between two points does not change if the two points are interchanged; (c) states the known fact

that the sum of the lengths of two sides of a triangle is always greater than or equal to the length of the third side. Now consider these examples:

1. Let  $X = \mathbb{Q}$ , the set of rational numbers, and define  $d(x, y) = |x - y|$ .
2. Let  $X = \mathbb{R}$ , and again define  $d(x, y) = |x - y|$ .
3. Let  $X$  consist of the points on the surface of a sphere. We can define two distance functions on  $X$ . Let  $d_1(P, Q)$  be the length of the chord joining  $P$  and  $Q$  on the sphere. We can also define another metric,  $d_2(P, Q)$ , as the length of the arc of the great circle passing through points  $P$  and  $Q$  on the surface of the sphere. It is not hard to convince oneself that  $d_1$  and  $d_2$  satisfy all the properties of a metric function.
4. Let  $\mathcal{C}^0[a, b]$  denote the set of continuous real-valued functions on the closed interval  $[a, b]$ . We can define  $d(f, g) = \int_a^b |f(x) - g(x)| dx$  for  $f, g \in \mathcal{C}^0(a, b)$ .
5. Let  $\mathcal{C}_B(a, b)$  denote the set of *bounded* continuous real-valued functions on the closed interval  $[a, b]$ . We then define

$$d(f, g) = \max_{x \in [a, b]} \{|f(x) - g(x)|\}$$

for  $f, g \in \mathcal{C}_B(a, b)$ . This notation says: Take the absolute value of the difference in  $f$  and  $g$  at all  $x$  in the interval  $[a, b]$  and then pick the maximum of all these values.

The metric function creates a natural setting in which to test the “closeness” of points in a metric space. One occasion on which the idea of closeness becomes essential is in the study of a sequence. A **sequence** is a mapping  $s : \mathbb{N} \rightarrow X$  from the set of natural numbers  $\mathbb{N}$  into the metric space  $X$ . Such a mapping associates with a positive integer  $n$  a point  $s(n)$  of the metric space  $X$ . It is customary to write  $s_n$  (or  $x_n$  to match the symbol  $X$ ) instead of  $s(n)$  and to enumerate the values of the function by writing  $\{x_n\}_{n=1}^{\infty}$ .

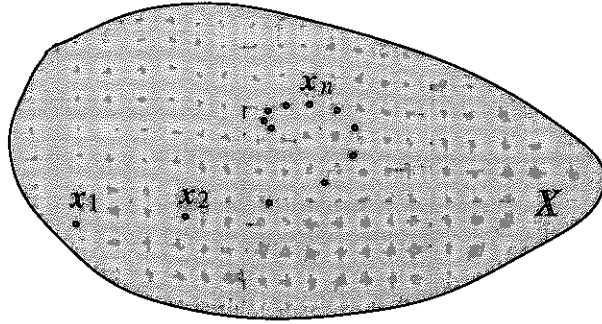
Knowledge of the behavior of a sequence for large values of  $n$  is of fundamental importance. In particular, it is important to know whether a sequence approaches a finite value as  $n$  increases.

**0.3.2. Box.** Suppose that for some  $x$  and for any positive real number  $\epsilon$ , there exists a natural number  $N$  such that  $d(x_n, x) < \epsilon$  whenever  $n > N$ . Then we say that the sequence  $\{x_n\}_{n=1}^{\infty}$  **converges** to  $x$  and write  $\lim_{n \rightarrow \infty} d(x_n, x) = 0$  or  $d(x_n, x) \rightarrow 0$  or simply  $x_n \rightarrow x$ .

It may not be possible to test directly for the convergence of a given sequence because this requires a knowledge of the limit point  $x$ . However, it is possible to

sequence defined

convergence defined



**Figure 4** The distance between the elements of a Cauchy sequence gets smaller and smaller.

do the next best thing—to see whether the points of the sequence get closer and closer as  $n$  gets larger and larger. A **Cauchy sequence** is a sequence for which  $\lim_{m,n \rightarrow \infty} d(x_m, x_n) = 0$ , as shown in Figure 4. We can test directly whether or not a sequence is Cauchy. However, the fact that a sequence is Cauchy does not guarantee that it converges. For example, let the metric space be the set of rational numbers  $\mathbb{Q}$  with the metric function  $d(x, y) = |x - y|$ , and consider the sequence  $\{x_n\}_{n=1}^{\infty}$  where  $x_n = \sum_{k=1}^n (-1)^{k+1}/k$ . It is clear that  $x_n$  is a rational number for any  $n$ . Also, to show that  $|x_m - x_n| \rightarrow 0$  is an exercise in calculus. Thus, the sequence is Cauchy. However, it is probably known to the reader that  $\lim_{n \rightarrow \infty} x_n = \ln 2$ , which is not a rational number.

A metric space in which every Cauchy sequence converges is called a **complete metric space**. Complete metric spaces play a crucial role in modern analysis. The preceding example shows that  $\mathbb{Q}$  is not a complete metric space. However, if the limit points of all Cauchy sequences are added to  $\mathbb{Q}$ , the resulting space becomes complete. This complete space is, of course, the real number system  $\mathbb{R}$ . It turns out that any incomplete metric space can be “enlarged” to a complete metric space.

## 0.4 Cardinality

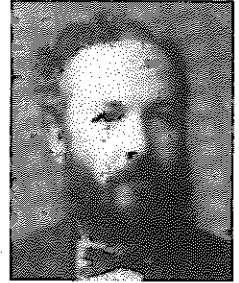
The process of counting is a one-to-one comparison of one set with another. If two sets are in one-to-one correspondence, they are said to have the same **cardinality**. Two sets with the same cardinality essentially have the same “number” of elements. The set  $F_n = \{1, 2, \dots, n\}$  is finite and has cardinality  $n$ . Any set from which there is a bijection onto  $F_n$  is said to be finite with  $n$  elements.

---

Although some steps had been taken before him in the direction of a definitive theory of sets, the creator of the theory of sets is considered to be **Georg Cantor** (1845–1918), who was born in Russia of Danish-Jewish parentage but moved to Germany with his parents.

His father urged him to study engineering, and Cantor entered the University of Berlin in 1863 with that intention. There he came under the influence of Weierstrass and turned to pure mathematics. He became Privatdozent at Halle in 1869 and professor in 1879. When he was twenty-nine he published his first revolutionary paper on the theory of infinite sets in the *Journal für Mathematik*. Although some of its propositions were deemed faulty by the older mathematicians, its overall originality and brilliance attracted attention. He continued to publish papers on the theory of sets and on transfinite numbers until 1897.

One of Cantor's main concerns was to differentiate among infinite sets by "size" and, like Bolzano before him, he decided that one-to-one correspondence should be the basic principle. In his correspondence with Dedekind in 1873, Cantor posed the question of whether the set of real numbers can be put into one-to-one correspondence with the integers, and some weeks later he answered in the negative. He gave two proofs. The first is more complicated than the second, which is the one most often used today. In 1874 Cantor occupied himself with the equivalence of the points of a line and the points of  $\mathbb{R}^n$  and sought to prove that a one-to-one correspondence between these two sets was impossible. Three years later he proved that there is such a correspondence. He wrote to Dedekind, "I see it but I do not believe it." He later showed that given any set, it is always possible to create a new set, the set of subsets of the given set, whose cardinal number is larger than that of the given set. If  $\aleph_0$  is the cardinal number of the set of subsets is denoted by  $2^{\aleph_0}$ . Cantor proved that  $2^{\aleph_0} = c$ , where  $c$  is the cardinal number of the continuum; i.e., the set of real numbers.



Cantor's work, which resolved age-old problems and reversed much previous thought, could hardly be expected to receive immediate acceptance. His ideas on transfinite ordinal and cardinal numbers aroused the hostility of the powerful Leopold Kronecker, who attacked Cantor's theory savagely over more than a decade, repeatedly preventing Cantor from obtaining a more prominent appointment in Berlin. Though Kronecker died in 1891, his attacks left mathematicians suspicious of Cantor's work. Poincaré referred to set theory as an interesting "pathological case." He also predicted that "Later generations will regard [Cantor's] *Mengenlehre* as a disease from which one has recovered." At one time Cantor suffered a nervous breakdown, but resumed work in 1887.

Many prominent mathematicians, however, were impressed by the uses to which the new theory had already been put in analysis, measure theory, and topology. Hilbert spread Cantor's ideas in Germany, and in 1926 said, "No one shall expel us from the paradise which Cantor created for us." He praised Cantor's transfinite arithmetic as "the most astonishing product of mathematical thought, one of the most beautiful realizations of human activity in the domain of the purely intelligible." Bertrand Russell described Cantor's work as "probably the greatest of which the age can boast." The subsequent utility of Cantor's work in formalizing mathematics—a movement largely led by Hilbert—seems at odds with Cantor's Platonic view that the greater importance of his work was in its implications for metaphysics and theology. That his work could be so seamlessly diverted from the goals intended by its creator is strong testimony to its objectivity and craftsmanship.

countably infinite Now consider the set of natural numbers  $\mathbb{N} = \{1, 2, 3, \dots\}$ . If there exists a bijection between a set  $A$  and  $\mathbb{N}$ , then  $A$  is said to be **countably infinite**. Some examples of countably infinite sets are the set of all integers, the set of even natural numbers, the set of odd natural numbers, the set of all prime numbers, and the set of energy levels of the bound states of a hydrogen atom.

It may seem surprising that a subset (such as the set of all even numbers) can be put into one-to-one correspondence with the full set (the set of all natural numbers); however, this is a property shared by all *infinite* sets. In fact, sometimes infinite sets are *defined* as those sets that are in one-to-one correspondence with at least one of their proper subsets. It is also astonishing to discover that there are as many rational numbers as there are natural numbers. After all, there are infinitely many rational numbers just in the interval  $(0, 1)$ —or between any two distinct real numbers.

uncountable sets Sets that are neither finite nor countably infinite are said to be **uncountable**. In some sense they are “more infinite” than any countable set. Examples of uncountable sets are the points in the interval  $(-1, +1)$ , the real numbers, the points in a plane, and the points in space. It can be shown that these sets have the same cardinality: There are as many points in three-dimensional space—the whole universe—as there are in the interval  $(-1, +1)$  or in any other finite interval.

Cantor set constructed Cardinality is a very intricate mathematical notion with many surprising results. Consider the interval  $[0, 1]$ . Remove the open interval  $(\frac{1}{3}, \frac{2}{3})$  from its middle. This means that the points  $\frac{1}{3}$  and  $\frac{2}{3}$  will not be removed. From the remaining portion,  $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ , remove the two middle thirds; the remaining portion will then be

$$[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$$

(see Figure 5). Do this indefinitely. What is the cardinality of the remaining set, which is called the **Cantor set**? Intuitively we expect hardly anything to be left. We might persuade ourselves into accepting the fact that the number of points remaining is at most infinite but countable. The surprising fact is that the cardinality is that of the continuum! Thus, after removal of infinitely many middle thirds, the set that remains has as many points as the original set!

## 0.5 Mathematical Induction

Many a time it is desirable to make a mathematical statement that is true for all natural numbers. For example, we may want to establish a formula involving an integer parameter that will hold for all positive integers. One encounters this situation when, after experimenting with the first few positive integers, one recognizes a pattern and discovers a formula, and wants to make sure that the formula holds for all natural numbers. For this purpose, one uses **mathematical induction**. The essence of mathematical induction is stated as follows:

induction principle



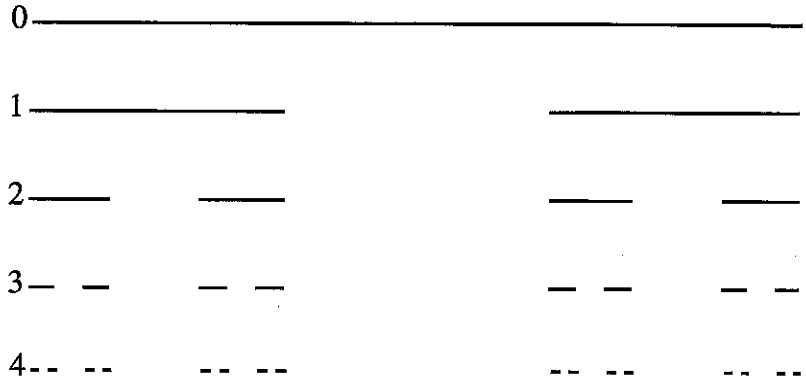


Figure 5 The Cantor set after one, two, three, and four “dissections.”

**0.5.1. Box.** Suppose that there is associated with each natural number (positive integer)  $n$  a statement  $S_n$ . Then  $S_n$  is true for every positive integer provided the following two conditions hold:

1.  $S_1$  is true.
2. If  $S_m$  is true for some given positive integer  $m$ , then  $S_{m+1}$  is also true.

We illustrate the use of mathematical induction by proving the **binomial theorem**

$$\begin{aligned}
 (a + b)^m &= \sum_{k=0}^m \binom{m}{k} a^{m-k} b^k = \sum_{k=0}^m \frac{m!}{k!(m-k)!} a^{m-k} b^k \\
 &= a^m + ma^{m-1}b + \frac{m(m-1)}{2!} a^{m-2}b^2 + \dots + mab^{m-1} + b^m,
 \end{aligned} \tag{1}$$

where we have used the notation

$$\binom{m}{k} \equiv \frac{m!}{k!(m-k)!}. \tag{2}$$

The mathematical statement  $S_m$  is Equation (1). We note that  $S_1$  is trivially true. Now we assume that  $S_m$  is true and show that  $S_{m+1}$  is also true. This means starting with Equation (1) and showing that

$$(a + b)^{m+1} = \sum_{k=0}^{m+1} \binom{m+1}{k} a^{m+1-k} b^k.$$

Then the induction principle ensures that the statement (equation) holds for all positive integers.

Multiply both sides of Equation (1) by  $a + b$  to obtain

$$(a + b)^{m+1} = \sum_{k=0}^m \binom{m}{k} a^{m-k+1} b^k + \sum_{k=0}^m \binom{m}{k} a^{m-k} b^{k+1}.$$

Now separate the  $k = 0$  term from the first sum and the  $k = m$  term from the second sum:

$$\begin{aligned} (a + b)^{m+1} &= a^{m+1} + \sum_{k=1}^m \binom{m}{k} a^{m-k+1} b^k + \underbrace{\sum_{k=0}^{m-1} \binom{m}{k} a^{m-k} b^{k+1}}_{\text{let } k = j - 1 \text{ in this sum}} + b^{m+1} \\ &= a^{m+1} + \sum_{k=1}^m \binom{m}{k} a^{m-k+1} b^k + \sum_{j=1}^m \binom{m}{j-1} a^{m-j+1} b^j + b^{m+1}. \end{aligned}$$

The second sum in the last line involves  $j$ . Since this is a dummy index, we can substitute any symbol we please. The choice  $k$  is especially useful because then we can unite the two summations. This gives

$$(a + b)^{m+1} = a^{m+1} + \sum_{k=1}^m \left\{ \binom{m}{k} + \binom{m}{k-1} \right\} a^{m-k+1} b^k + b^{m+1}.$$

If we now use

$$\binom{m+1}{k} = \binom{m}{k} + \binom{m}{k-1},$$

which the reader can easily verify, we finally obtain

$$\begin{aligned} (a + b)^{m+1} &= a^{m+1} + \sum_{k=1}^m \binom{m+1}{k} a^{m-k+1} b^k + b^{m+1} \\ &= \sum_{k=0}^{m+1} \binom{m+1}{k} a^{m-k+1} b^k. \end{aligned}$$

Mathematical induction is also used in *defining* quantities involving integers. Such definitions are called **inductive definitions**. For example, inductive definition is used in defining powers:  $a^1 = a$  and  $a^m = a^{m-1}a$ .

## 0.6 Problems

0.1. Show that the number of subsets of a set containing  $n$  elements is  $2^n$ .

**0.2.** Let  $A$ ,  $B$ , and  $C$  be sets in a universal set  $U$ . Show that

- $A \subset B$  and  $B \subset C$  implies  $A \subset C$ .
- $A \subset B$  iff  $A \cap B = A$  iff  $A \cup B = B$ .
- $A \subset B$  and  $B \subset C$  implies  $(A \cup B) \subset C$ .
- $A \cup B = (A \sim B) \cup (A \cap B) \cup (B \sim A)$ .

Hint: To show the equality of two sets, show that each set is a subset of the other.

**0.3.** For each  $n \in \mathbb{N}$ , let

$$I_n = \left\{ x \mid |x - 1| < n \text{ and } |x + 1| > \frac{1}{n} \right\}.$$

Find  $\cup_n I_n$  and  $\cap_n I_n$ .

**0.4.** Show that  $a' \in \llbracket a \rrbracket$  implies that  $\llbracket a' \rrbracket = \llbracket a \rrbracket$ .

**0.5.** Can you define a binary operation of “multiplication” on the set of vectors in space? What about vectors in the plane?

**0.6.** Show that  $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$  when  $f$  and  $g$  are both bijections.

**0.7.** Take any two open intervals  $(a, b)$  and  $(c, d)$ , and show that there are as many points in the first as there are in the second, regardless of the size of the intervals. Hint: Find a (linear) algebraic relation between points of the two intervals.

**0.8.** Use mathematical induction to derive the **Leibniz rule** for differentiating a product:

Leibniz rule

$$\frac{d^n}{dx^n}(f \cdot g) = \sum_{k=0}^n \binom{n}{k} \frac{d^k f}{dx^k} \frac{d^{n-k} g}{dx^{n-k}}.$$

**0.9.** Use mathematical induction to derive the following results:

$$\sum_{k=0}^n r^k = \frac{r^{n+1} - 1}{r - 1}, \quad \sum_{k=0}^n k = \frac{n(n+1)}{2}.$$

## Additional Reading

- Halmos, P. *Naive Set Theory*, Springer-Verlag, 1974. A classic text on intuitive (as opposed to axiomatic) set theory covering all the topics discussed in this chapter and much more.
- Kelley, J. *General Topology*, Springer-Verlag, 1985. The introductory chapter of this classic reference is a detailed introduction to set theory and mappings.
- Simmons, G. *Introduction to Topology and Modern Analysis*, Krieger, 1983. The first chapter of this book covers not only set theory and mappings, but also the Cantor set and the fact that integers are as abundant as rational numbers.



Part I \_\_\_\_\_

# Finite-Dimensional Vector Spaces



# Vectors and Transformations

Two- and three-dimensional vectors—undoubtedly familiar objects to the reader—can easily be generalized to higher dimensions. Representing vectors by their components, one can conceive of vectors having  $N$  components. This is the most immediate generalization of vectors in the plane and in space, and such vectors are called  $N$ -dimensional *Cartesian* vectors. Cartesian vectors are limited in two respects: Their components are real, and their dimensionality is finite. Some applications in physics require the removal of one or both of these limitations. It is therefore convenient to study vectors stripped of any dimensionality or reality of components. Such properties become consequences of more fundamental definitions. Although we will be concentrating on finite-dimensional vector spaces in this part of the book, many of the concepts and examples introduced here apply to infinite-dimensional spaces as well.

## 1.1 Vector Spaces

Let us begin with the definition of an abstract (complex) vector space.<sup>1</sup>

**1.1.1. Definition.** A *vector space*  $\mathcal{V}$  over  $\mathbb{C}$  is a set of objects denoted by  $|a\rangle$ ,  $|b\rangle$ ,  $|x\rangle$ , and so on, called *vectors*, with the following properties:<sup>2</sup>

1. To every pair of vectors  $|a\rangle$  and  $|b\rangle$  in  $\mathcal{V}$  there corresponds a vector  $|a\rangle + |b\rangle$ , also in  $\mathcal{V}$ , called the *sum* of  $|a\rangle$  and  $|b\rangle$ , such that

$$(a) \quad |a\rangle + |b\rangle = |b\rangle + |a\rangle,$$

---

<sup>1</sup>Keep in mind that  $\mathbb{C}$  is the set of complex numbers and  $\mathbb{R}$  the set of reals.

<sup>2</sup>The *bra*,  $\langle |$ , and *ket*,  $| \rangle$ , notation for vectors, invented by Dirac, is very useful when dealing with complex vector spaces. However, it is somewhat clumsy for certain topics such as norm and metrics and will therefore be abandoned in those discussions.

$$(b) |a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle,$$

(c) There exists a unique vector  $|0\rangle \in \mathcal{V}$ , called the zero vector, such that  $|a\rangle + |0\rangle = |a\rangle$  for every vector  $|a\rangle$ ,

(d) To every vector  $|a\rangle \in \mathcal{V}$  there corresponds a unique vector  $-|a\rangle$  (also in  $\mathcal{V}$ ) such that  $|a\rangle + (-|a\rangle) = |0\rangle$ .

scalars are numbers

2. To every complex number<sup>3</sup>  $\alpha$ —also called a *scalar*—and every vector  $|a\rangle$  there corresponds a vector  $\alpha|a\rangle$  in  $\mathcal{V}$  such that

$$(a) \alpha(\beta|a\rangle) = (\alpha\beta)|a\rangle,$$

$$(b) 1|a\rangle = |a\rangle.$$

3. Multiplication involving vectors and scalars is distributive:

$$(a) \alpha(|a\rangle + |b\rangle) = \alpha|a\rangle + \alpha|b\rangle.$$

$$(b) (\alpha + \beta)|a\rangle = \alpha|a\rangle + \beta|a\rangle.$$

complex vs. real  
vector space

concept of field  
summarized

The vector space defined above is also called a **complex vector space**. It is possible to replace  $\mathbb{C}$  with  $\mathbb{R}$ —the set of real numbers—in which case the resulting space will be called a **real vector space**. Real and complex numbers are prototypes of a mathematical structure called **field**. A field is a set of objects with two binary operations called addition and multiplication. Each operation distributes with respect to the other, and each operation has an identity. The identity for addition is denoted by 0 and is called *additive identity*. The identity for multiplication is denoted by 1 and is called *multiplicative identity*. Furthermore, every element has an additive inverse, and every element except the additive identity has a multiplicative inverse.

### 1.1.2. Example. SOME VECTOR SPACES

1.  $\mathbb{R}$  is a vector space over the field of real numbers.
2.  $\mathbb{C}$  is a vector space over the field of real numbers.
3.  $\mathbb{C}$  is a vector space over the complex numbers.
4. Let  $\mathcal{V} = \mathbb{R}$  and let the field of scalars be  $\mathbb{C}$ . This is *not* a vector space, because property 2 of Definition 1.1.1 is not satisfied: A complex number times a real number is not a real number and therefore does not belong to  $\mathcal{V}$ .
5. The set of “arrows” in the plane (or in space) forms a vector space over  $\mathbb{R}$  under the parallelogram law of addition of planar (or spatial) vectors.

<sup>3</sup>Complex numbers, particularly when they are treated as *variables*, are usually denoted by  $z$ , and we shall adhere to this convention in Part III. However, in the discussion of vector spaces, we have found it more convenient to use lower case Greek letters to denote complex numbers as scalars.



6. Let  $\mathcal{P}^{\mathbb{C}}[t]$  be the set of all polynomials with complex coefficients in a variable  $t$ . Then  $\mathcal{P}^{\mathbb{C}}[t]$  is a vector space under the ordinary addition of polynomials and the multiplication of a polynomial by a complex number. In this case the zero vector is the zero polynomial.
7. For a given positive integer  $n$ , let  $\mathcal{P}_n^{\mathbb{C}}[t]$  be the set of all polynomials with complex coefficients of degree less than or equal to  $n$ . Again it is easy to verify that  $\mathcal{P}_n^{\mathbb{C}}[t]$  is a vector space under the usual addition of polynomials and their multiplication by complex scalars. In particular, the sum of two polynomials of degree less than or equal to  $n$  is also a polynomial of degree less than or equal to  $n$ , and multiplying a polynomial with complex coefficients by a complex number gives another polynomial of the same type. Here the zero polynomial is the zero vector.
8. The set  $\mathcal{P}_n^{\mathbb{R}}[t]$  of polynomials of degree less than or equal to  $n$  with real coefficients is a vector space over the reals, but it is *not* a vector space over the complex numbers.
9. Let  $\mathbb{C}^n$  consist of all complex  $n$ -tuples such as  $|a\rangle = (\alpha_1, \alpha_2, \dots, \alpha_n)$  and  $|b\rangle = (\beta_1, \beta_2, \dots, \beta_n)$ . Let  $\alpha$  be a complex number. Then we define

$$\begin{aligned} |a\rangle + |b\rangle &= (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n), \\ \alpha |a\rangle &= (\alpha\alpha_1, \alpha\alpha_2, \dots, \alpha\alpha_n), \\ |0\rangle &= (0, 0, \dots, 0), \\ -|a\rangle &= (-\alpha_1, -\alpha_2, \dots, -\alpha_n). \end{aligned}$$

It is easy to verify that  $\mathbb{C}^n$  is a vector space over the complex numbers. It is called the *n-dimensional complex coordinate space*.

10. The set of all real  $n$ -tuples  $\mathbb{R}^n$  is a vector space over the real numbers under the operations similar to that of  $\mathbb{C}^n$ . It is called the *n-dimensional real coordinate space*, or *Cartesian n-space*. It is not a vector space over the complex numbers.
11. The set of all complex matrices with  $m$  rows and  $n$  columns  $\mathcal{M}^{m \times n}$  is a vector space under the usual addition of matrices and multiplication by complex numbers. The zero vector is the  $m \times n$  matrix with all entries equal to zero.
12. Let  $\mathbb{C}^{\infty}$  be the set of all complex sequences  $|a\rangle = \{\alpha_i\}_{i=1}^{\infty}$  such that  $\sum_{i=1}^{\infty} |\alpha_i|^2 < \infty$ . One can show that with addition and scalar multiplication defined component-wise,  $\mathbb{C}^{\infty}$  is a vector space over the complex numbers.
13. The set of all complex-valued functions of a single real variable that are continuous in the real interval  $(a, b)$  is a vector space over the complex numbers.
14. The set  $\mathcal{C}^n(a, b)$  on  $(a, b)$  of all real-valued functions of a single real variable that possess continuous derivatives of all orders up to  $n$  forms a vector space over the reals.
15. The set  $\mathcal{C}^{\infty}(a, b)$  of all real-valued functions on  $(a, b)$  of a single real variable that possess derivatives of all orders forms a vector space over the reals. ■

It is clear from the example above that the existence of a vector space depends as much on the nature of the vectors as on the nature of the scalars.

**1.1.3. Definition.** The vectors  $|a_1\rangle, |a_2\rangle, \dots, |a_n\rangle$ , are said to be **linearly independent** if for  $\alpha_i \in \mathbb{C}$ , the relation  $\sum_{i=1}^n \alpha_i |a_i\rangle = 0$  implies  $\alpha_i = 0$  for all  $i$ . The sum  $\sum_{i=1}^n \alpha_i |a_i\rangle$  is called a **linear combination** of  $\{|a_i\rangle\}_{i=1}^n$ .

*n*-dimensional complex coordinate space

*n*-dimensional real coordinate space, or Cartesian *n*-space

linear independence defined linear combination of vectors

subspace **1.1.4. Definition.** A *subspace*  $\mathcal{W}$  of a vector space  $\mathcal{V}$  is a nonempty subset of  $\mathcal{V}$  with the property that if  $|a\rangle, |b\rangle \in \mathcal{W}$ , then  $\alpha |a\rangle + \beta |b\rangle$  also belongs to  $\mathcal{W}$  for all  $\alpha, \beta \in \mathbb{C}$ .

The intersection of two subspaces is also a subspace. A subspace is a vector space in its own right. The reader may verify that *the intersection of two subspaces is also a subspace*.

span of a subset of a vector space **1.1.5. Theorem.** If  $S$  is any nonempty set of vectors in a vector space  $\mathcal{V}$ , then the set  $\mathcal{W}_S$  of all linear combinations of vectors in  $S$  is a subspace of  $\mathcal{V}$ . We say that  $\mathcal{W}_S$  is the *span* of  $S$ , or that  $S$  spans  $\mathcal{W}_S$ , or that  $\mathcal{W}_S$  is spanned by  $S$ .  $\mathcal{W}_S$  is sometimes denoted by  $\text{Span}\{S\}$ .

The proof of Theorem 1.1.5 is left as Problem 1.8.

basis defined **1.1.6. Definition.** A *basis* of a vector space  $\mathcal{V}$  is a set  $B$  of linearly independent vectors that spans all of  $\mathcal{V}$ . A vector space that has a finite basis is called *finite-dimensional*; otherwise, it is *infinite-dimensional*.

We state the following without proof (see [Axle 96, page 31]):

**1.1.7. Theorem.** All bases of a given finite-dimensional vector space have the same number of linearly independent vectors. This number is called the *dimension* of the vector space. A vector space of dimension  $N$  is sometimes denoted by  $\mathcal{V}_N$ .

components of a vector in a basis If  $|a\rangle$  is a vector in an  $N$ -dimensional vector space  $\mathcal{V}$  and  $B = \{|a_i\rangle\}_{i=1}^N$  a basis in that space, then by the definition of a basis, there exists a unique (see Problem 1.4) set of scalars  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$  such that  $|a\rangle = \sum_{i=1}^N \alpha_i |a_i\rangle$ . The set  $\{\alpha_i\}_{i=1}^N$  is called the *components* of  $|a\rangle$  with respect to the basis  $B$ .

**1.1.8. Example.** The following are subspaces of some of the vector spaces considered in Example 1.1.2.

- The “space” of real numbers is a subspace of  $\mathbb{C}$  over the reals.
- $\mathbb{R}$  is *not* a subspace of  $\mathbb{C}$  over the complex numbers, because as explained in Example 1.1.2,  $\mathbb{R}$  cannot be a vector space over the complex numbers.
- The set of all vectors along a given line *going through the origin* is a subspace of arrows in the plane (or space) over  $\mathbb{R}$ .
- $\mathcal{P}_n^{\mathbb{C}}[t]$  is a subspace of  $\mathcal{P}^{\mathbb{C}}[t]$ .
- $\mathbb{C}^{n-1}$  is a subspace of  $\mathbb{C}^n$  when  $\mathbb{C}^{n-1}$  is identified with all complex  $n$ -tuples with zero last entry. In general,  $\mathbb{C}^m$  is a subspace of  $\mathbb{C}^n$  for  $m < n$  when  $\mathbb{C}^m$  is identified with all  $n$ -tuples whose last  $n - m$  elements are zero.
- $\mathcal{M}^{r \times s}$  is a subspace of  $\mathcal{M}^{m \times n}$  for  $r \leq m$  or  $s \leq n$ . Here, we identify an  $r \times s$  matrix with an  $m \times n$  matrix whose last  $m - r$  rows and  $n - s$  columns are all zero.
- $\mathcal{P}_m^{\mathbb{C}}[t]$  is a subspace of  $\mathcal{P}_n^{\mathbb{C}}[t]$  for  $m < n$ .
- $\mathcal{P}_m^r[t]$  is a subspace of  $\mathcal{P}_n^r[t]$  for  $m < n$ . Note that both  $\mathcal{P}_n^r[t]$  and  $\mathcal{P}_m^r[t]$  are vector spaces over the reals only.

- $\mathbb{R}^m$  is a subspace of  $\mathbb{R}^n$  for  $m < n$ . Therefore,  $\mathbb{R}^2$ , the plane, is a subspace of  $\mathbb{R}^3$ , the Euclidean space. Also,  $\mathbb{R}^1 \equiv \mathbb{R}$  is a subspace of both the plane  $\mathbb{R}^2$  and the Euclidean space  $\mathbb{R}^3$ . ■

**1.1.9. Example.** The following are bases for the vector spaces given in Example 1.1.2.

- The number 1 is a basis for  $\mathbb{R}$ , which is therefore one-dimensional.
- The numbers 1 and  $i = \sqrt{-1}$  are basis vectors for the vector space  $\mathbb{C}$  over  $\mathbb{R}$ . Thus, this space is two-dimensional.
- The number 1 is a basis for  $\mathbb{C}$  over  $\mathbb{C}$ , and the space is one-dimensional. Note that although the vectors are the same as in the preceding item, changing the nature of the scalars changes the dimensionality of the space.
- The set  $\{\hat{e}_x, \hat{e}_y, \hat{e}_z\}$  of the unit vectors in the directions of the three axes forms a basis in space. The space is three-dimensional.
- A basis of  $\mathcal{P}^c[t]$  can be formed by the monomials  $1, t, t^2, \dots$ . It is clear that this space is *infinite-dimensional*.
- A basis of  $\mathbb{C}^n$  is given by  $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n$ , where  $\hat{e}_j$  is an  $n$ -tuple that has a 1 at the  $j$ th position and zeros everywhere else. This basis is called the **standard basis** of  $\mathbb{C}^n$ . Clearly, the space has  $n$  dimensions.
- A basis of  $\mathcal{M}^{m \times n}$  is given by  $e_{11}, e_{12}, \dots, e_{ij}, \dots, e_{mn}$ , where  $e_{ij}$  is the  $m \times n$  matrix with zeros everywhere except at the intersection of the  $i$ th row and  $j$ th column, where it has a one.
- A set consisting of the monomials  $1, t, t^2, \dots, t^n$  forms a basis of  $\mathcal{P}_n^c[t]$ . Thus, this space is  $(n + 1)$ -dimensional.
- The standard basis of  $\mathbb{C}^n$  is a basis of  $\mathbb{R}^n$  as well. It is also called the *standard basis* of  $\mathbb{R}^n$ . Thus,  $\mathbb{R}^n$  is  $n$ -dimensional.
- If we assume that  $a < 0 < b$ , then the set of monomials  $1, x, x^2, \dots$  forms a basis for  $\mathcal{C}^\infty(a, b)$ , because, by Taylor's theorem, any function belonging to  $\mathcal{C}^\infty(a, b)$  can be expanded in an infinite power series about  $x = 0$ . Thus, this space is infinite-dimensional. ■

standard basis of  $\mathbb{C}^n$

Given a space  $\mathcal{V}$  with a basis  $B = \{[a_i]\}_{i=1}^n$ , the span of any  $m$  vectors ( $m < n$ ) of  $B$  is an  $m$ -dimensional subspace of  $\mathcal{V}$ .

## 1.2 Inner Product

A vector space, as given by Definition 1.1.1, is too general and structureless to be of much physical interest. One useful structure introduced on a vector space is a scalar product. Recall that the scalar (dot) product of vectors in the plane or in space is a rule that associates with two vectors  $\mathbf{a}$  and  $\mathbf{b}$ , a real number. This association, denoted symbolically by  $g : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ , with  $g(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$ , is symmetric:  $g(\mathbf{a}, \mathbf{b}) = g(\mathbf{b}, \mathbf{a})$ , is linear in the first (and by symmetry, in the second) factor:<sup>4</sup>

$$g(\alpha\mathbf{a} + \beta\mathbf{b}, \mathbf{c}) = \alpha g(\mathbf{a}, \mathbf{c}) + \beta g(\mathbf{b}, \mathbf{c}) \quad \text{or} \quad (\alpha\mathbf{a} + \beta\mathbf{b}) \cdot \mathbf{c} = \alpha\mathbf{a} \cdot \mathbf{c} + \beta\mathbf{b} \cdot \mathbf{c},$$

<sup>4</sup>A function that is linear in both of its arguments is called a **bilinear** function.

gives the “length” of a vector:  $|\mathbf{a}|^2 = g(\mathbf{a}, \mathbf{a}) = \mathbf{a} \cdot \mathbf{a} \geq 0$ , and ensures that the only vector with zero length<sup>5</sup> is the zero vector:  $g(\mathbf{a}, \mathbf{a}) = 0$  if and only if  $\mathbf{a} = \mathbf{0}$ .

We want to generalize these properties to abstract vector spaces for which the scalars are complex numbers. A verbatim generalization of the foregoing properties, however, leads to a contradiction. Using the linearity in both arguments and a nonzero  $|a\rangle$ , we obtain

$$g(i|a\rangle, i|a\rangle) = i^2 g(|a\rangle, |a\rangle) = -g(|a\rangle, |a\rangle). \quad (1.1)$$

Either the right-hand side (RHS) or left-hand side (LHS) of this equation must be negative! But this is inconsistent with the positivity of the “length” of a vector, which requires  $g(|a\rangle, |a\rangle)$  to be positive for *all* nonzero vectors, including  $i|a\rangle$ . The source of the problem is the linearity in *both* arguments. If we can change this property in such a way that one of the  $i$ 's in Equation (1.1) comes out complex-conjugated, the problem may go away. This requires linearity in one argument and complex-conjugate linearity in the other. Which argument is to be complex-conjugate linear is a matter of convention. We choose the first argument to be so.<sup>6</sup> We thus have

$$g(\alpha|a\rangle + \beta|b\rangle, |c\rangle) = \alpha^* g(|a\rangle, |c\rangle) + \beta^* g(|b\rangle, |c\rangle),$$

where  $\alpha^*$  denotes the complex conjugate. Consistency then requires us to change the symmetry property as well. In fact, we must demand that  $g(|a\rangle, |b\rangle) = (g(|b\rangle, |a\rangle))^*$ , from which the *reality* of  $g(|a\rangle, |a\rangle)$ —a necessary condition for its positivity—follows immediately.

The question of the existence of an inner product on a vector space is a deep problem in higher analysis. Generally, if an inner product exists, there may be many ways to introduce one on a vector space. However, as we shall see in Section 1.2.4, a *finite-dimensional* vector space always has an inner product and this inner product is unique.<sup>7</sup> So, for all practical purposes we can speak of *the* inner product on a finite-dimensional vector space, and as with the two- and three-dimensional cases, we can omit the letter  $g$  and use a notation that involves only the vectors. There are several such notations in use, but the one that will be employed in this book is the *Dirac bra(c)ket notation*, whereby  $g(|a\rangle, |b\rangle)$  is denoted by  $\langle a|b\rangle$ . Using this notation, we have

**1.2.1. Definition.** *The inner product of two vectors,  $|a\rangle$  and  $|b\rangle$ , in a vector space  $\mathcal{V}$  is a complex number,  $\langle a|b\rangle \in \mathbb{C}$ , such that*

1.  $\langle a|b\rangle = \langle b|a\rangle^*$
2.  $\langle a|(\beta|b\rangle + \gamma|c\rangle) = \beta\langle a|b\rangle + \gamma\langle a|c\rangle$

Dirac “bra,”  $\langle |$ , and “ket”  $| \rangle$ , notation is used for inner products.

inner product defined

<sup>5</sup>In our present discussion, we are avoiding situations in which a nonzero vector can have zero “length.” Such occasions arise in relativity, and we shall discuss them in Part VII.

<sup>6</sup>In some books, particularly in the mathematical literature, the second argument is chosen to be linear.

<sup>7</sup>This uniqueness holds up to a certain equivalence of inner products that we shall not get into here.

3.  $\langle a|a \rangle \geq 0$ , and  $\langle a|a \rangle = 0$  if and only if  $|a \rangle = |0 \rangle$ .

positive definite, or  
Riemannian inner  
product

The last relation is called the **positive definite** property of the inner product.<sup>8</sup> A positive definite inner product is also called a **Riemannian inner product**, otherwise it is called **pseudo-Riemannian**.

Note that linearity in the first argument is absent, because, as explained earlier, it would be inconsistent with the first property, which expresses the “symmetry” of the inner product. The extra operation of complex conjugation renders the true linearity in the second argument impossible. Because of this complex conjugation, the inner product on a complex vector space is not truly bilinear; it is commonly called **sesquilinear**.

sesquilinear

A shorthand notation will be useful when dealing with the inner product of a linear combination of vectors.

**1.2.2. Box.** We write the LHS of the second equation in the definition above as  $\langle a|\beta b + \gamma c \rangle$ .

This has the advantage of treating a linear combination as a single vector. The second property then states that if the complex scalars happen to be in a *ket*, they “split out” unaffected:

$$\langle a|\beta b + \gamma c \rangle = \beta \langle a|b \rangle + \gamma \langle a|c \rangle. \tag{1.2}$$

On the other hand, if the complex scalars happen to be in the first factor (the bra), then they should be conjugated when they are “split out”:

$$\langle \beta b + \gamma c|a \rangle = \beta^* \langle b|a \rangle + \gamma^* \langle c|a \rangle. \tag{1.3}$$

A vector space  $\mathcal{V}$  on which an inner product is defined is called an **inner product space**. As mentioned above, all finite-dimensional vector spaces can be turned into inner product spaces.

**1.2.3. Example.** In this example we introduce some of the most common inner products. The reader is urged to verify that in all cases, we indeed have an inner product.

- Let  $|a \rangle, |b \rangle \in \mathbb{C}^n$ , with  $|a \rangle = (\alpha_1, \alpha_2, \dots, \alpha_n)$  and  $|b \rangle = (\beta_1, \beta_2, \dots, \beta_n)$ , and define an inner product on  $\mathbb{C}^n$  as

$$\langle a|b \rangle \equiv \alpha_1^* \beta_1 + \alpha_2^* \beta_2 + \dots + \alpha_n^* \beta_n = \sum_{i=1}^n \alpha_i^* \beta_i.$$

That this product satisfies all the required properties of an inner product is easily checked. For example, if  $|b \rangle = |a \rangle$ , we obtain  $\langle a|a \rangle = |\alpha_1|^2 + |\alpha_2|^2 + \dots + |\alpha_n|^2$ , which is clearly nonnegative.

natural inner product  
for  $\mathbb{C}^n$

<sup>8</sup>The positive definiteness must be relaxed in the space-time of relativity theory, in which nonzero vectors can have zero “length.”

- Similarly, for  $|a\rangle, |b\rangle \in \mathbb{R}^n$  the same definition (without the complex conjugation) satisfies all the properties of an inner product.
- For  $|a\rangle, |b\rangle \in \mathbb{C}^\infty$  the natural inner product is defined as  $\langle a|b\rangle = \sum_{i=1}^\infty \alpha_i^* \beta_i$ . The question of the convergence of this sum is the subject of Problem 1.16.
- Let  $x(t), y(t) \in \mathcal{P}^c[t]$ , the space of all polynomials in  $t$  with complex coefficients. Define

weight function of an inner product defined in terms of integrals

$$\langle x|y\rangle \equiv \int_a^b w(t)x^*(t)y(t) dt, \tag{1.4}$$

where  $a$  and  $b$  are real numbers—or infinity—for which the integral exists, and  $w(t)$  is a real-valued, continuous function that is *always strictly positive* in the interval  $(a, b)$ . Then Equation (1.4) defines an inner product. Depending on the so-called **weight function**  $w(t)$ , there can be many different inner products defined on the infinite-dimensional space  $\mathcal{P}^c[t]$ .

natural inner product for complex functions

- Let  $f, g \in \mathbb{C}(a, b)$  and define their inner product by

$$\langle f|g\rangle \equiv \int_a^b w(x)f^*(x)g(x) dx.$$

It is easily shown that  $\langle f|g\rangle$  satisfies all the requirements of the inner product if, as in the previous case, the weight function  $w(x)$  is always positive in the interval  $(a, b)$ . This is called the *standard inner product* on  $\mathbb{C}(a, b)$ . ■

### 1.2.1 Orthogonality

The vectors of analytic geometry and calculus are often expressed in terms of unit vectors along the axes, i.e., vectors that are of unit length and perpendicular to one another. Such vectors are also important in abstract inner product spaces.

orthogonality defined  
orthonormal basis

**1.2.4. Definition.** Vectors  $|a\rangle, |b\rangle \in \mathcal{V}$  are *orthogonal* if  $\langle a|b\rangle = 0$ . A *normal vector*, or *normalized vector*,  $|e\rangle$  is one for which  $\langle e|e\rangle = 1$ . A basis  $B = \{|e_i\rangle\}_{i=1}^N$  in an  $N$ -dimensional vector space  $\mathcal{V}$  is an *orthonormal basis* if

$$\langle e_i|e_j\rangle = \delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \tag{1.5}$$

Kronecker delta

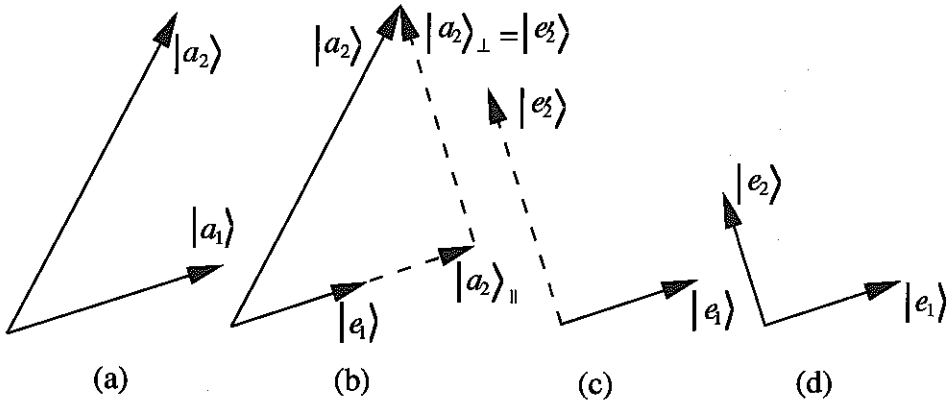
where  $\delta_{ij}$ , defined by the last equality, is called the **Kronecker delta**.

**1.2.5. Example.** Here are examples of orthonormal bases:

- The standard basis of  $\mathbb{R}^n$  (or  $\mathbb{C}^n$ )

$$|e_1\rangle = (1, 0, \dots, 0), |e_2\rangle = (0, 1, \dots, 0), \dots, |e_n\rangle = (0, 0, \dots, 1)$$

is orthonormal under the usual inner product of those spaces.



**Figure 1.1** The essence of the Gram–Schmidt process is neatly illustrated by the process in two dimensions. This figure, depicts the stages of the construction of two orthonormal vectors.

- Let  $|e_k\rangle = e^{ikx}/\sqrt{2\pi}$  be functions in  $C(0, 2\pi)$  with  $w(x) = 1$ . Then

$$\langle e_k | e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} e^{ikx} dx = 1,$$

and for  $l \neq k$ ,

$$\langle e_l | e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{-ilx} e^{ikx} dx = \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-l)x} dx = 0.$$

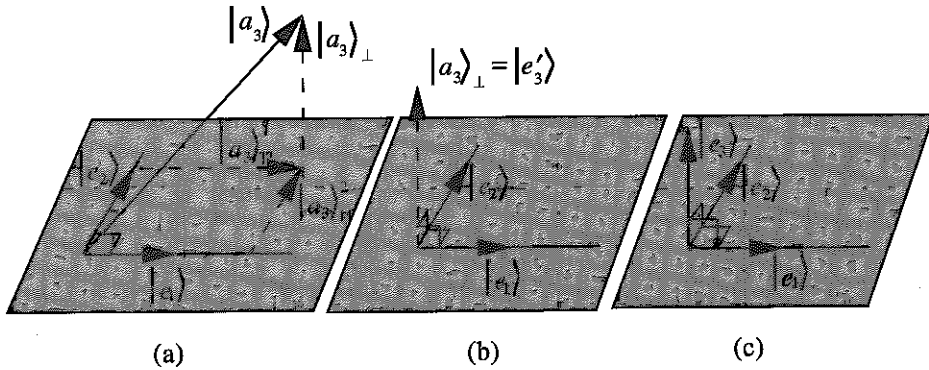
Thus,  $\langle e_l | e_k \rangle = \delta_{lk}$ . ■

### 1.2.2 The Gram–Schmidt Process

It is always possible to convert any basis in  $\mathcal{V}$  into an orthonormal basis. A process by which this may be accomplished is called **Gram–Schmidt orthonormalization**. Consider a basis  $B = \{|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle\}$ . We intend to take linear combinations of  $|a_i\rangle$  in such a way that the resulting vectors are orthonormal. First, we let  $|e_1\rangle = |a_1\rangle/\sqrt{\langle a_1 | a_1 \rangle}$  and note that  $\langle e_1 | e_1 \rangle = 1$ . If we subtract from  $|a_2\rangle$  its projection along  $|e_1\rangle$ , we obtain a vector that is orthogonal to  $|e_1\rangle$  (see Figure 1.1). Calling the resulting vector  $|e_2'\rangle$ , we have  $|e_2'\rangle = |a_2\rangle - \langle e_1 | a_2 \rangle |e_1\rangle$ , which can be written more symmetrically as  $|e_2'\rangle = |a_2\rangle - |e_1\rangle \langle e_1 | a_2 \rangle$ . Clearly, this vector is orthogonal to  $|e_1\rangle$ . In order to normalize  $|e_2'\rangle$ , we divide it by  $\sqrt{\langle e_2' | e_2' \rangle}$ . Then  $|e_2\rangle = |e_2'\rangle/\sqrt{\langle e_2' | e_2' \rangle}$  will be a normal vector orthogonal to  $|e_1\rangle$ . Subtracting from  $|a_3\rangle$  its projections along the first and second unit vectors obtained so far will give the vector

$$|e_3'\rangle = |a_3\rangle - |e_1\rangle \langle e_1 | a_3 \rangle - |e_2\rangle \langle e_2 | a_3 \rangle = |a_3\rangle - \sum_{i=1}^2 |e_i\rangle \langle e_i | a_3 \rangle,$$

The Gram–Schmidt process explained



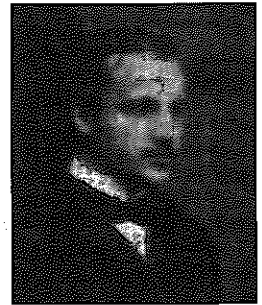
**Figure 1.2** Once the orthonormal vectors in the plane of two vectors are obtained, the third orthonormal vector is easily constructed.

which is orthogonal to both  $|e_1\rangle$  and  $|e_2\rangle$  (see Figure 1.2):

$$\langle e_1 | e'_3 \rangle = \langle e_1 | a_3 \rangle - \overbrace{\langle e_1 | e_1 \rangle}^{=1} \langle e_1 | a_3 \rangle - \overbrace{\langle e_1 | e_2 \rangle}^{=0} \langle e_2 | a_3 \rangle = 0.$$

Similarly,  $\langle e_2 | e'_3 \rangle = 0$ .

**Erhard Schmidt** (1876–1959) obtained his doctorate under the supervision of David Hilbert. His main interest was in integral equations and Hilbert spaces. He is the “Schmidt” of the **Gram–Schmidt orthogonalization process**, which takes a basis of a space and constructs an orthonormal one from it. (Laplace had presented a special case of this process long before Gram or Schmidt.)



In 1908 Schmidt worked on infinitely many equations in infinitely many unknowns, introducing various geometric notations and terms that are still in use for describing spaces of functions. Schmidt’s ideas were to lead to the geometry of Hilbert spaces. This was motivated by the study of integral equations (see Chapter 17) and an attempt at their abstraction.

Earlier, Hilbert regarded a function as given by its Fourier coefficients. These satisfy the condition that  $\sum_{k=1}^{\infty} a_k^2$  is finite. He introduced sequences of real numbers  $\{x_n\}$  such that  $\sum_{n=1}^{\infty} x_n^2$  is finite. Riesz and Fischer showed that there is a one-to-one correspondence between square-integrable functions and square-summable sequences of their Fourier coefficients. In 1907 Schmidt and Fréchet showed that a consistent theory could be obtained if the square-summable sequences were regarded as the coordinates of points in an infinite-dimensional space that is a generalization of  $n$ -dimensional Euclidean space. Thus *functions can be regarded as points of a space*, now called a **Hilbert space**.



In general, if we have calculated  $m$  orthonormal vectors  $|e_1\rangle, \dots, |e_m\rangle$ , with  $m < N$ , then we can find the next one using the following relations:

$$\begin{aligned} |e'_{m+1}\rangle &= |a_{m+1}\rangle - \sum_{i=1}^m |e_i\rangle \langle e_i | a_{m+1}\rangle, \\ |e_{m+1}\rangle &= \frac{|e'_{m+1}\rangle}{\sqrt{\langle e'_{m+1} | e'_{m+1}\rangle}}. \end{aligned} \quad (1.6)$$

Even though we have been discussing finite-dimensional vector spaces, the process of Equation (1.6) can continue for infinite-dimensions as well. The reader is asked to pay attention to the fact that, at each stage of the Gram–Schmidt process, one is taking linear combinations of the original vectors.

### 1.2.3 The Schwarz Inequality

Let us now consider an important inequality that is valid in both finite and infinite dimensions and whose restriction to two and three dimensions is equivalent to the fact that the cosine of the angle between two vectors is always less than one.

**1.2.6. Theorem.** For any pair of vectors  $|a\rangle, |b\rangle$  in an inner product space  $\mathcal{V}$ , the Schwarz inequality holds:  $\langle a|a\rangle \langle b|b\rangle \geq |\langle a|b\rangle|^2$ . Equality holds when  $|a\rangle$  is proportional to  $|b\rangle$ .

*Proof.* Let  $|c\rangle = |b\rangle - (\langle a|b\rangle / \langle a|a\rangle) |a\rangle$ , and note that  $\langle a|c\rangle = 0$ . Write  $|b\rangle = (\langle a|b\rangle / \langle a|a\rangle) |a\rangle + |c\rangle$  and take the inner product of  $|b\rangle$  with itself:

$$\langle b|b\rangle = \left| \frac{\langle a|b\rangle}{\langle a|a\rangle} \right|^2 \langle a|a\rangle + \langle c|c\rangle.$$

Since the last term is never negative, we have

$$\langle b|b\rangle \geq \frac{|\langle a|b\rangle|^2}{\langle a|a\rangle} \Rightarrow \langle a|a\rangle \langle b|b\rangle \geq |\langle a|b\rangle|^2.$$

Equality holds iff  $\langle c|c\rangle = 0$  or  $|c\rangle = 0$ . From the definition of  $|c\rangle$ , we conclude that  $|a\rangle$  and  $|b\rangle$  must be proportional.  $\square$

Notice the power of abstraction: We have derived the Schwarz inequality solely from the basic assumptions of inner product spaces independent of the specific nature of the inner product. Therefore, we do not have to prove the Schwarz inequality every time we encounter a new inner product space.

---

**Karl Herman Amandus Schwarz** (1843–1921) the son of an architect, was born in what is now Sobiecin, Poland. After gymnasium, Schwarz studied chemistry in Berlin for a time

before switching to mathematics, receiving his doctorate in 1864. He was greatly influenced by the reigning mathematicians in Germany at the time, especially Kummer and Weierstrass. The lecture notes that Schwarz took while attending Weierstrass's lectures on the integral calculus still exist. Schwarz received an initial appointment at Halle and later appointments in Zurich and Göttingen before being named as Weierstrass's successor at Berlin in 1892. These later years, filled with students and lectures, were not Schwarz's most productive, but his early papers assure his place in mathematics history.

Schwarz's favorite tool was geometry, which he soon turned to the study of analysis. He conclusively proved some of Riemann's results that had been previously (and justifiably) challenged. The primary result in question was the assertion that every simply connected region in the plane could be conformally mapped onto a circular area. From this effort came several well-known results now associated with Schwarz's name, including the principle of reflection and Schwarz's lemma. He also worked on surfaces of minimal area, the branch of geometry beloved by all who dabble with soap bubbles.



Schwarz's most important work, for the occasion of Weierstrass's seventieth birthday, again dealt with minimal area, specifically whether a minimal surface yields a minimal area. Along the way, Schwarz demonstrated second variation in a multiple integral, constructed a function using successive approximation, and demonstrated the existence of a "least" eigenvalue for certain differential equations. This work also contained the most famous inequality in mathematics, which bears his name.

Schwarz's success obviously stemmed from a matching of his aptitude and training to the mathematical problems of the day. One of his traits, however, could be viewed as either positive or negative—his habit of treating all problems, whether trivial or monumental, with the same level of attention to detail. This might also at least partly explain the decline in productivity in Schwarz's later years.

Schwarz had interests outside mathematics, although his marriage was a mathematical one, since he married Kummer's daughter. Outside mathematics he was the captain of the local voluntary fire brigade, and he assisted the stationmaster at the local railway station by closing the doors of the trains!

## 1.2.4 Length of a Vector

In dealing with objects such as directed line segments in the plane or in space, the intuitive idea of the length of a vector is used to define the dot product. However, sometimes it is more convenient to introduce the inner product first and then define the length, as we shall do now.

norm of a vector  
defined

**1.2.7. Definition.** *The norm, or length, of a vector  $|a\rangle$  in an inner product space is denoted by  $\|a\|$  and defined as  $\|a\| \equiv \sqrt{\langle a|a\rangle}$ . We use the notation  $\|\alpha a + \beta b\|$  for the norm of the vector  $\alpha |a\rangle + \beta |b\rangle$ .*

One can easily show that the norm has the following properties:

1. The norm of the zero vector is zero:  $\|0\| = 0$ .
2.  $\|a\| \geq 0$ , and  $\|a\| = 0$  if and only if  $|a\rangle = |0\rangle$ .
3.  $\|\alpha a\| = |\alpha| \|a\|$  for any<sup>9</sup> complex  $\alpha$ .
4.  $\|a + b\| \leq \|a\| + \|b\|$ . This property is called the **triangle inequality**.

triangle inequality

normed linear space

natural distance in a normed linear space

A vector space on which a norm is defined is called a **normed linear space**. One can introduce the idea of the “distance” between two vectors in a normed linear space. The distance between  $|a\rangle$  and  $|b\rangle$ —denoted by  $d(a, b)$ —is simply the norm of their difference:  $d(a, b) \equiv \|a - b\|$ . It can be readily shown that this has all the properties one expects of the distance (or metric) function introduced in Chapter 0. However, one does not need a normed space to define distance. For example, as explained in Chapter 0, one can define the distance between two points on the surface of a sphere, but the addition of two points on a sphere—a necessary operation for vector space structure—is not defined. Thus the points on a sphere form a metric space, but not a vector space.

parallelogram law

Inner product spaces are automatically normed spaces, but the converse is not, in general, true: There are normed spaces, i.e., spaces satisfying properties 1–4 above that cannot be promoted to inner product spaces. However, if the norm satisfies the **parallelogram law**,

$$\|a + b\|^2 + \|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2, \quad (1.7)$$

then one can define

$$\langle a | b \rangle \equiv \frac{1}{4} \{ \|a + b\|^2 - \|a - b\|^2 - i (\|a + ib\|^2 - \|a - ib\|^2) \} \quad (1.8)$$

and show that it is indeed an inner product. In fact, we have (see [Frie 82, pp. 203–204] for a proof) the following theorem.

**1.2.8. Theorem.** *A normed linear space is an inner product space if and only if the norm satisfies the parallelogram law.*

Now consider any  $N$ -dimensional vector space  $\mathcal{V}$ . Choose a basis  $\{|a_i\rangle\}_{i=1}^N$  in  $\mathcal{V}$ , and for any vector  $|a\rangle$  whose components are  $\{\alpha_i\}_{i=1}^N$  in this basis, define

$$\|a\|^2 \equiv \sum_{i=1}^N |\alpha_i|^2.$$

The reader may check that this defines a norm, and that the norm satisfies the parallelogram law. From Theorem 1.2.8 we have the following:

**1.2.9. Theorem.** *Every finite-dimensional vector space can be turned into an inner product space.*

$\mathbb{C}^n$  has many different distance functions

**1.2.10. Example.** Let the space be  $\mathbb{C}^n$ . The natural inner product of  $\mathbb{C}^n$  gives rise to a norm, which, for the vector  $|a\rangle = (\alpha_1, \alpha_2, \dots, \alpha_n)$  is

$$\|a\| = \sqrt{\langle a|a\rangle} = \sqrt{\sum_{i=1}^n |\alpha_i|^2}.$$

This norm yields the following distance between  $|a\rangle$  and  $|b\rangle = (\beta_1, \beta_2, \dots, \beta_n)$ :

$$d(a, b) = \|a - b\| = \sqrt{\langle a - b|a - b\rangle} = \sqrt{\sum_{i=1}^n |\alpha_i - \beta_i|^2}.$$

One can define other norms, such as  $\|a\|_1 \equiv \sum_{i=1}^n |\alpha_i|$ , which has all the required properties of a norm, and leads to the distance

$$d_1(a, b) = \|a - b\|_1 = \sum_{i=1}^n |\alpha_i - \beta_i|.$$

Another norm defined on  $\mathbb{C}^n$  is given by

$$\|a\|_p \equiv \left( \sum_{i=1}^n |\alpha_i|^p \right)^{1/p},$$

where  $p$  is a positive integer. It is proved in higher mathematical analysis that  $\|\cdot\|_p$  has all the properties of a norm. (The nontrivial part of the proof is to verify the triangle inequality.) The associated distance is

$$d_p(a, b) = \|a - b\|_p = \left( \sum_{i=1}^n |\alpha_i - \beta_i|^p \right)^{1/p}.$$

The other two norms introduced above are special cases, for  $p = 2$  and  $p = 1$ . ■

## 1.3 Linear Transformations

We have made progress in enriching vector spaces with structures such as norms and inner products. However, this enrichment, although important, will be of little value if it is imprisoned in a single vector space. We would like to give vector space properties freedom of movement, so they can go from one space to another. The vehicle that carries these properties is a linear transformation which is the subject of this section. However, first it is instructive to review the concept of a mapping (discussed in Chapter 0) by considering some examples relevant to the present discussion.

**1.3.1. Example.** The following are a few familiar examples of mappings.

<sup>9</sup>The first property follows from this by letting  $\alpha = 0$ .

1. Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be given by  $f(x) = x^2$ .
2. Let  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  be given by  $g(x, y) = x^2 + y^2 - 4$ .
3. Let  $F : \mathbb{R}^2 \rightarrow \mathbb{C}$  be given by  $F(x, y) = U(x, y) + iV(x, y)$ , where  $U : \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $V : \mathbb{R}^2 \rightarrow \mathbb{R}$ .
4. Let  $T : \mathbb{R} \rightarrow \mathbb{R}^2$  be given by  $T(t) = (t + 3, 2t - 5)$ .
5. Motion of a point particle in space can be considered as a mapping  $M : [a, b] \rightarrow \mathbb{R}^3$ , where  $[a, b]$  is an interval of the real line. For each  $t \in [a, b]$ , we define  $M(t) = (x(t), y(t), z(t))$ , where  $x(t)$ ,  $y(t)$ , and  $z(t)$  are real-valued functions of  $t$ . If we identify  $t$  with time, which is assumed to have a value in the interval  $[a, b]$ , then  $M(t)$  describes the path of the particle as a function of time, and  $a$  and  $b$  are the beginning and the end of the motion, respectively. ■

Let us consider an arbitrary mapping  $F : \mathcal{V} \rightarrow \mathcal{W}$  from a vector space  $\mathcal{V}$  to another vector space  $\mathcal{W}$ . It is assumed that the two vector spaces are over the same scalars, say  $\mathbb{C}$ . Consider  $|a\rangle$  and  $|b\rangle$  in  $\mathcal{V}$  and  $|x\rangle$  and  $|y\rangle$  in  $\mathcal{W}$  such that  $F(|a\rangle) = |x\rangle$  and  $F(|b\rangle) = |y\rangle$ . In general,  $F$  does not preserve the vector space structure. That is, the image of a linear combination of vectors is not the same as the linear combination of the images:

$$F(\alpha |a\rangle + \beta |b\rangle) \neq \alpha F(|a\rangle) + \beta F(|b\rangle).$$

This is the case for all the mappings of Example 1.3.1 except the fourth item. There are many applications in which the preservation of the vector space structure (preservation of the linear combination) is desired.

linear transformation,  
linear operator,  
endomorphism

**1.3.2. Definition.** A *linear transformation* from the complex vector space  $\mathcal{V}$  to the complex vector space  $\mathcal{W}$  is a mapping  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  such that

$$\mathbf{T}(\alpha |a\rangle + \beta |b\rangle) = \alpha \mathbf{T}(|a\rangle) + \beta \mathbf{T}(|b\rangle) \quad \forall |a\rangle, |b\rangle \in \mathcal{V} \text{ and } \alpha, \beta \in \mathbb{C}.$$

A linear transformation  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{V}$  is called an *endomorphism* of  $\mathcal{V}$  or a *linear operator* on  $\mathcal{V}$ . The action of a linear transformation on a vector is written without the parentheses:  $\mathbf{T}(|a\rangle) \equiv \mathbf{T}|a\rangle$ .

The same definition applies to real vector spaces. Note that the definition demands that both vector spaces have the same set of scalars: The same scalars multiply vectors in  $\mathcal{V}$  on the LHS and those in  $\mathcal{W}$  on the RHS. An immediate consequence of the definition is the following:

**1.3.3. Box.** Two linear transformations  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  and  $\mathbf{U} : \mathcal{V} \rightarrow \mathcal{W}$  are equal if and only if  $\mathbf{T}|a_i\rangle = \mathbf{U}|a_i\rangle$  for all  $|a_i\rangle$  in some basis of  $\mathcal{V}$ . Thus, a linear transformation is uniquely determined by its action on some basis of its domain space.

The equality in this box is simply the set-theoretic equality of maps discussed in Chapter 0. An important example of a linear transformation occurs when the second vector space,  $\mathcal{W}$ , happens to be the set of scalars,  $\mathbb{C}$  or  $\mathbb{R}$ , in which case the linear transformation is called a **linear functional**.

linear functional

$\mathcal{L}(\mathcal{V}, \mathcal{W})$  is a vector space

The set of linear transformations from  $\mathcal{V}$  to  $\mathcal{W}$  is denoted by  $\mathcal{L}(\mathcal{V}, \mathcal{W})$ , and this set happens to be a vector space. The *zero transformation*,  $\mathbf{0}$ , is defined to take every vector in  $\mathcal{V}$  to the zero vector of  $\mathcal{W}$ . The sum of two linear transformations  $\mathbf{T}$  and  $\mathbf{U}$  is the linear transformation  $\mathbf{T} + \mathbf{U}$ , whose action on a vector  $|a\rangle \in \mathcal{V}$  is defined to be  $(\mathbf{T} + \mathbf{U})|a\rangle \equiv \mathbf{T}|a\rangle + \mathbf{U}|a\rangle$ . Similarly, define  $\alpha\mathbf{T}$  by  $(\alpha\mathbf{T})|a\rangle \equiv \alpha(\mathbf{T}|a\rangle) = \alpha\mathbf{T}|a\rangle$ . The set of endomorphisms of  $\mathcal{V}$  is denoted by  $\mathcal{L}(\mathcal{V})$  rather than  $\mathcal{L}(\mathcal{V}, \mathcal{V})$ .

dual vector space  $\mathcal{V}^*$

The set of linear functionals  $\mathcal{L}(\mathcal{V}, \mathbb{C})$ —or  $\mathcal{L}(\mathcal{V}, \mathbb{R})$  if  $\mathcal{V}$  is a real vector space—is denoted by  $\mathcal{V}^*$  and is called the **dual space** of  $\mathcal{V}$ .

**1.3.4. Example.** The following are some examples of linear operators in various vector spaces. The proofs of linearity are simple in all cases and are left as exercises for the reader.

1. Let  $\{|a_1\rangle, |a_2\rangle, \dots, |a_m\rangle\}$  be an arbitrary finite set of vectors in  $\mathcal{V}$ , and  $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$  an arbitrary set of linear functionals on  $\mathcal{V}$ . Let

$$\mathbf{A} \equiv \sum_{k=1}^m |a_k\rangle \mathbf{f}_k \in \mathcal{L}(\mathcal{V})$$

be defined by  $\mathbf{A}|x\rangle = \sum_{k=1}^m |a_k\rangle \mathbf{f}_k(|x\rangle) = \sum_{k=1}^m \mathbf{f}_k(|x\rangle) |a_k\rangle$ . Then  $\mathbf{A}$  is a linear operator on  $\mathcal{V}$ .

2. Let  $\pi$  be a permutation (shuffling) of the integers  $\{1, 2, \dots, n\}$ . If  $|x\rangle = (\eta_1, \eta_2, \dots, \eta_n)$  is a vector in  $\mathbb{C}^n$ , we can write

$$\mathbf{A}_\pi |x\rangle = (\eta_{\pi(1)}, \eta_{\pi(2)}, \dots, \eta_{\pi(n)}).$$

Then  $\mathbf{A}_\pi$  is a linear operator.

derivative operator

3. For any  $|x\rangle \in \mathcal{P}^c[t]$ , with  $x(t) = \sum_{k=0}^n \alpha_k t^k$ , write  $|y\rangle = \mathbf{D}|x\rangle$ , where  $|y\rangle$  is defined as  $y(t) = \sum_{k=1}^n k\alpha_k t^{k-1}$ . Then  $\mathbf{D}$  is a linear operator, the **derivative operator**.

integration operator

4. For every  $|x\rangle \in \mathcal{P}^c[t]$ , with  $x(t) = \sum_{k=0}^n \alpha_k t^k$ , write  $|y\rangle = \mathbf{S}|x\rangle$ , where  $|y\rangle \in \mathcal{P}^c[t]$  is defined as  $y(t) = \sum_{k=0}^n [\alpha_k / (k+1)] t^{k+1}$ . Then  $\mathbf{S}$  is a linear operator, the **integration operator**.

integration is a linear functional on the space of continuous functions

5. Define the operator  $\text{int} : \mathcal{C}^0(a, b) \rightarrow \mathbb{R}$  by

$$\text{int}(f) = \int_a^b f(t) dt.$$

Then  $\text{int}$  is a linear functional on the vector space  $\mathcal{C}^0(a, b)$ .

6. Let  $\mathcal{C}^n(a, b)$  be the set of real-valued functions defined in the interval  $[a, b]$  whose first  $n$  derivatives exist and are continuous. For any  $|f\rangle \in \mathcal{C}^n(a, b)$  define  $|u\rangle = \mathbf{G}|f\rangle$ , with  $u(t) = g(t)f(t)$  and  $g(t)$  a fixed function in  $\mathcal{C}^n(a, b)$ . Then  $\mathbf{G}$  is linear. In particular, the operation of multiplying by  $t$ , whose operator is denoted by  $\mathbf{T}$ , is linear. ▀

An immediate consequence of Definition 1.3.2 is that the image of the zero vector in  $\mathcal{V}$  is the zero vector in  $\mathcal{W}$ . This is not true for a general mapping, but it is necessarily true for a linear mapping. As the zero vector of  $\mathcal{V}$  is mapped onto the zero vector of  $\mathcal{W}$ , other vectors of  $\mathcal{V}$  may also be dragged along. In fact, we have the following theorem.

**1.3.5. Theorem.** *The set of vectors in  $\mathcal{V}$  that are mapped onto the zero vector of  $\mathcal{W}$  under the linear transformation  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  form a subspace of  $\mathcal{V}$  called the **kernel**, or **null space**, of  $\mathbf{T}$  and denoted by  $\ker \mathbf{T}$ .*

kernel of a linear transformation

*Proof.* The proof is left as an exercise. □

nullity The dimension of  $\ker \mathbf{T}$  is also called the **nullity** of  $\mathbf{T}$ .

The proof of the following is also left as an exercise.

**1.3.6. Theorem.** *The range  $\mathbf{T}(\mathcal{V})$  of a linear transformation  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  is a subspace of  $\mathcal{W}$ . The dimension of  $\mathbf{T}(\mathcal{V})$  is called the **rank** of  $\mathbf{T}$ .*

rank of a linear transformation

**1.3.7. Theorem.** *A linear transformation is 1–1 (injective) iff its kernel is zero.*

*Proof.* The “only if” part is trivial. For the “if” part, suppose  $\mathbf{T}|a_1\rangle = \mathbf{T}|a_2\rangle$ ; then linearity of  $\mathbf{T}$  implies that  $\mathbf{T}(|a_1\rangle - |a_2\rangle) = 0$ . Since  $\ker \mathbf{T} = 0$ , we must have  $|a_1\rangle = |a_2\rangle$ . □

Suppose we start with a basis of  $\ker \mathbf{T}$  and add enough linearly independent vectors to it to get a basis for  $\mathcal{V}$ . Without loss of generality, let us assume that the first  $n$  vectors in this basis form a basis of  $\ker \mathbf{T}$ . So let  $B = \{|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle\}$  be a basis for  $\mathcal{V}$  and  $B' = \{|a_1\rangle, |a_2\rangle, \dots, |a_n\rangle\}$  be a basis for  $\ker \mathbf{T}$ . Here  $N = \dim \mathcal{V}$  and  $n = \dim \ker \mathbf{T}$ . It is straightforward to show that  $\{\mathbf{T}|a_{n+1}\rangle, \dots, \mathbf{T}|a_N\rangle\}$  is a basis for  $\mathbf{T}(\mathcal{V})$ . We therefore have the following result.

dimension theorem **1.3.8. Theorem.** *Let  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  be a linear transformation. Then<sup>10</sup>*

$$\dim \mathcal{V} = \dim \ker \mathbf{T} + \dim \mathbf{T}(\mathcal{V})$$

This theorem is called the **dimension theorem**. One of its consequences is that an injective endomorphism is automatically surjective, and vice versa:

**1.3.9. Proposition.** *An endomorphism of a finite-dimensional vector space is bijective if it is either injective or surjective.*

The dimension theorem is obviously valid only for finite-dimensional vector spaces. In particular, neither surjectivity nor injectivity implies bijectivity for infinite-dimensional vector spaces.

<sup>10</sup>Recall that the dimension of a vector space depends on the scalars used in that space. Although we are dealing with two different vector spaces here, since they are both over the same set of scalars (complex or real), no confusion in the concept of dimension arises.

**1.3.10. Example.** Let us try to find the kernel of  $\mathbf{T} : \mathbb{R}^4 \rightarrow \mathbb{R}^3$  given by

$$\mathbf{T}(x_1, x_2, x_3, x_4) = (2x_1 + x_2 + x_3 - x_4, x_1 + x_2 + 2x_3 + 2x_4, x_1 - x_3 - 3x_4).$$

We must look for  $(x_1, x_2, x_3, x_4)$  such that  $\mathbf{T}(x_1, x_2, x_3, x_4) = (0, 0, 0)$ , or

$$\begin{aligned} 2x_1 + x_2 + x_3 - x_4 &= 0, \\ x_1 + x_2 + 2x_3 + 2x_4 &= 0, \\ x_1 - x_3 - 3x_4 &= 0. \end{aligned}$$

The “solution” to these equations is  $x_1 = x_3 + 3x_4$  and  $x_2 = -3x_3 - 5x_4$ . Thus, to be in  $\ker \mathbf{T}$ , a vector in  $\mathbb{R}^4$  must be of the form

$$(x_3 + 3x_4, -3x_3 - 5x_4, x_3, x_4) = x_3(1, -3, 1, 0) + x_4(3, -5, 0, 1),$$

where  $x_3$  and  $x_4$  are arbitrary real numbers. It follows that  $\ker \mathbf{T}$  consists of vectors that can be written as linear combinations of the two *linearly independent* vectors  $(1, -3, 1, 0)$  and  $(3, -5, 0, 1)$ . Therefore,  $\dim \ker \mathbf{T} = 2$ . Theorem 1.3.8 then says that  $\dim \mathbf{T}(\mathcal{V}) = 2$ ; that is, the range of  $\mathbf{T}$  is two-dimensional. This becomes clear when one notes that

$$\mathbf{T}(x_1, x_2, x_3, x_4) = (2x_1 + x_2 + x_3 - x_4)(1, 0, 1) + (x_1 + x_2 + 2x_3 + 2x_4)(0, 1, -1),$$

and therefore  $\mathbf{T}(x_1, x_2, x_3, x_4)$ , an arbitrary vector in the range of  $\mathbf{T}$ , is a linear combination of *only* two linearly independent vectors,  $(1, 0, 1)$  and  $(0, 1, -1)$ . ■

In many cases, two vector spaces may “look” different, while in reality they are very much the same. For example, the set of complex numbers  $\mathbb{C}$  is a two-dimensional vector space *over the reals*, as is  $\mathbb{R}^2$ . Although we call the vectors of these two spaces by different names, they have very similar properties. This notion of “similarity” is made precise in the following definition.

isomorphism and  
automorphism

**1.3.11. Definition.** A vector space  $\mathcal{V}$  is said to be *isomorphic* to another vector space  $\mathcal{W}$  if there exists a bijective linear mapping  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$ . Then  $\mathbf{T}$  is called an *isomorphism*.<sup>11</sup> A bijective linear map of  $\mathcal{V}$  onto itself is called an *automorphism* of  $\mathcal{V}$ . The set of automorphisms of  $\mathcal{V}$  is denoted by  $GL(\mathcal{V})$ .

For all practical purposes, two isomorphic vector spaces are different manifestations of the “same” vector space. In the example discussed above, the correspondence  $\mathbf{T} : \mathbb{C} \rightarrow \mathbb{R}^2$ , with  $\mathbf{T}(x + iy) = (x, y)$ , establishes an isomorphism between the two vector spaces. It should be emphasized that *only as vector spaces* are  $\mathbb{C}$  and  $\mathbb{R}^2$  isomorphic. If we go beyond the vector space structures, the two sets are quite different. For example,  $\mathbb{C}$  has a natural multiplication for its elements, but  $\mathbb{R}^2$  does not. The following theorem gives a working criterion for isomorphism.

<sup>11</sup>The word “isomorphism,” as we shall see, is used in conjunction with many algebraic structures. To distinguish them, qualifiers need to be used. In the present context, we speak of **linear isomorphism**. We shall use qualifiers when necessary. However, the context usually makes the meaning of isomorphism clear.



**1.3.12. Theorem.** A linear surjective map  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  is an isomorphism if and only if its nullity is zero.

*Proof.* The “only if” part is obvious. To prove the “if” part, assume that the nullity is zero. Then by Theorem 1.3.7,  $\mathbf{T}$  is 1–1. Since it is already surjective,  $\mathbf{T}$  must be bijective.  $\square$

**1.3.13. Theorem.** An isomorphism  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  carries linearly independent sets of vectors onto linearly independent sets of vectors.

*Proof.* Assume that  $\{|a_i\rangle\}_{i=1}^m$  is a set of linearly independent vectors in  $\mathcal{V}$ . To show that  $\{\mathbf{T}|a_i\rangle\}_{i=1}^m$  is linearly independent in  $\mathcal{W}$ , assume that there exist  $\alpha_1, \alpha_2, \dots, \alpha_m$  such that  $\sum_{i=1}^m \alpha_i \mathbf{T}|a_i\rangle = |0\rangle$ . Then the linearity of  $\mathbf{T}$  and Theorem 1.3.12 give  $\mathbf{T}(\sum_{i=1}^m \alpha_i |a_i\rangle) = |0\rangle$ , or  $\sum_{i=1}^m \alpha_i |a_i\rangle = |0\rangle$ , and the linear independence of the  $|a_i\rangle$  implies that  $\alpha_i = 0$  for all  $i$ . Therefore,  $\{\mathbf{T}|a_i\rangle\}_{i=1}^m$  must be linearly independent.  $\square$

The following theorem shows that finite-dimensional vector spaces are severely limited in number:

**1.3.14. Theorem.** Two finite-dimensional vector spaces are isomorphic if and only if they have the same dimension.

*Proof.* Let  $B_{\mathcal{V}} = \{|a_i\rangle\}_{i=1}^N$  be a basis for  $\mathcal{V}$  and  $B_{\mathcal{W}} = \{|b_i\rangle\}_{i=1}^N$  a basis for  $\mathcal{W}$ . Define the linear transformation  $\mathbf{T}|a_i\rangle = |b_i\rangle, i = 1, 2, \dots, N$ . The rest of the proof involves showing that  $\mathbf{T}$  is an isomorphism. We leave this as an exercise for the reader.  $\square$

only two  
 $N$ -dimensional  
vector spaces

A consequence of Theorem 1.3.14 is that all  $N$ -dimensional vector spaces over  $\mathbb{R}$  are isomorphic to  $\mathbb{R}^N$  and all complex  $N$ -dimensional vector spaces are isomorphic to  $\mathbb{C}^N$ . So, for all practical purposes, we have only two  $N$ -dimensional vector spaces,  $\mathbb{R}^N$  and  $\mathbb{C}^N$ .

### 1.3.1 More on Linear Functionals

An example of isomorphism is that between a vector space and its dual space, which we discuss now. Consider an  $N$ -dimensional vector space with a basis  $B = \{|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle\}$ . For any given set of  $N$  scalars,  $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ , define the linear functional  $\mathbf{f}_\alpha$  by  $\mathbf{f}_\alpha |a_i\rangle = \alpha_i$ . When  $\mathbf{f}_\alpha$  acts on any arbitrary vector  $|b\rangle = \sum_{i=1}^N \beta_i |a_i\rangle$  in  $\mathcal{V}$ , the result is

$$\mathbf{f}_\alpha |b\rangle = \mathbf{f}_\alpha \left( \sum_{i=1}^N \beta_i |a_i\rangle \right) = \sum_{i=1}^N \beta_i \mathbf{f}_\alpha |a_i\rangle = \sum_{i=1}^N \beta_i \alpha_i. \tag{1.9}$$

This expression suggests that  $|b\rangle$  can be represented as a column vector with entries  $\beta_1, \beta_2, \dots, \beta_N$  and  $\mathbf{f}_\alpha$  as a row vector with entries  $\alpha_1, \alpha_2, \dots, \alpha_N$ . Then  $\mathbf{f}_\alpha |b\rangle$  is

Every set of  $N$  scalars defines a linear functional.

merely the matrix product<sup>12</sup> of the row vector (on the left) and the column vector (on the right).

$\mathbf{f}_\alpha$  is uniquely determined by the set  $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ . In other words, corresponding to every set of  $N$  scalars there exists a unique linear functional. This leads us to a particular set of functionals,  $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N$  corresponding, respectively, to the sets of scalars  $\{1, 0, 0, \dots, 0\}, \{0, 1, 0, \dots, 0\}, \dots, \{0, 0, 0, \dots, 1\}$ . This means that

$$\begin{aligned} \mathbf{f}_1 |a_1\rangle &= 1 & \text{and} & & \mathbf{f}_1 |a_j\rangle &= 0 & \text{for } j \neq 1, \\ \mathbf{f}_2 |a_2\rangle &= 1 & \text{and} & & \mathbf{f}_2 |a_j\rangle &= 0 & \text{for } j \neq 2, \\ & \vdots & & & & & \\ \mathbf{f}_N |a_N\rangle &= 1 & \text{and} & & \mathbf{f}_N |a_j\rangle &= 0 & \text{for } j \neq N, \end{aligned}$$

or that

$$\mathbf{f}_i |a_j\rangle = \delta_{ij}, \tag{1.10}$$

where  $\delta_{ij}$  is the Kronecker delta.

The functionals of Equation (1.10) form a basis of the dual space  $\mathcal{V}^*$ . To show this, consider an arbitrary  $\mathbf{g} \in \mathcal{V}^*$ , which is uniquely determined by its action on the vectors in a basis  $B = \{|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle\}$ . Let  $\mathbf{g} |a_i\rangle = \gamma_i \in \mathbb{C}$ . Then we claim that  $\mathbf{g} = \sum_{i=1}^N \gamma_i \mathbf{f}_i$ . In fact, consider an arbitrary vector  $|a\rangle$  in  $\mathcal{V}$  with components  $(\alpha_1, \alpha_2, \dots, \alpha_N)$  with respect to  $B$ . Then, on the one hand,  $\mathbf{g} |a\rangle = \mathbf{g}(\sum_{i=1}^N \alpha_i |a_i\rangle) = \sum_{i=1}^N \alpha_i \mathbf{g} |a_i\rangle = \sum_{i=1}^N \alpha_i \gamma_i$ . On the other hand,

$$\begin{aligned} \left(\sum_{i=1}^N \gamma_i \mathbf{f}_i\right) |a\rangle &= \left(\sum_{i=1}^N \gamma_i \mathbf{f}_i\right) \left(\sum_{j=1}^N \alpha_j |a_j\rangle\right) \\ &= \sum_{i=1}^N \gamma_i \sum_{j=1}^N \alpha_j \mathbf{f}_i |a_j\rangle = \sum_{i=1}^N \gamma_i \sum_{j=1}^N \alpha_j \delta_{ij} = \sum_{i=1}^N \gamma_i \alpha_i. \end{aligned}$$

Since the actions of  $\mathbf{g}$  and  $\sum_{i=1}^N \gamma_i \mathbf{f}_i$  yield equal results for arbitrary  $|a\rangle$ , we conclude that  $\mathbf{g} = \sum_{i=1}^N \gamma_i \mathbf{f}_i$ , i.e.,  $\{\mathbf{f}_i\}_{i=1}^N$  span  $\mathcal{V}^*$ . Thus, we have the following result.

**1.3.15. Theorem.** *If  $\mathcal{V}$  is an  $N$ -dimensional vector space with a basis  $B = \{|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle\}$ , then there is a corresponding unique basis  $B^* = \{\mathbf{f}_i\}_{i=1}^N$  in  $\mathcal{V}^*$  with the property that  $\mathbf{f}_i |a_j\rangle = \delta_{ij}$ .*

By this theorem the dual space of an  $N$ -dimensional vector space is also  $N$ -dimensional, and thus isomorphic to it. The basis  $B^*$  is called the **dual basis** of  $B$ . A corollary to Theorem 1.3.15 is that to every vector in  $\mathcal{V}$  there corresponds a

<sup>12</sup>Matrices will be taken up in Chapter 3. Here, we assume only a nodding familiarity with elementary matrix operations.

unique linear functional in  $\mathcal{V}^*$ . This can be seen by noting that every vector  $|a\rangle$  is uniquely determined by its components  $(\alpha_1, \alpha_2, \dots, \alpha_N)$  in a basis  $B$ . The unique linear functional  $\mathbf{f}_a$  corresponding to  $|a\rangle$ , also called the dual of  $|a\rangle$ , is simply  $\sum_{i=1}^N \alpha_i \mathbf{f}_i$ , with  $\mathbf{f}_i \in B^*$ .

annihilator of a vector and a subspace

**1.3.16. Definition.** An annihilator of  $|a\rangle \in \mathcal{V}$  is a linear functional  $\mathbf{f} \in \mathcal{V}^*$  such that  $\mathbf{f}|a\rangle = 0$ . Let  $\mathcal{W}$  be a subspace of  $\mathcal{V}$ . The set of linear functionals in  $\mathcal{V}^*$  that annihilate all vectors in  $\mathcal{W}$  is denoted by  $\mathcal{W}^0$ .

The reader may check that  $\mathcal{W}^0$  is a subspace of  $\mathcal{V}^*$ . Moreover, if we extend a basis  $\{|a_i\rangle\}_{i=1}^k$  of  $\mathcal{W}$  to a basis  $B = \{|a_i\rangle\}_{i=1}^N$  of  $\mathcal{V}$ , then we can show that the functionals  $\{\mathbf{f}_j\}_{j=k+1}^N$ , chosen from the basis  $B^* = \{\mathbf{f}_j\}_{j=1}^N$  dual to  $B$ , span  $\mathcal{W}^0$ . It then follows that

$$\dim \mathcal{V} = \dim \mathcal{W} + \dim \mathcal{W}^0. \tag{1.11}$$

We shall have occasions to use annihilators later on when we discuss symplectic geometry.

We have “dualized” a vector, a basis, and a complete vector space. The only object remaining is a linear transformation.

dual, or pull back, of a linear transformation

**1.3.17. Definition.** Let  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{U}$  be a linear transformation. Define  $\mathbf{T}^* : \mathcal{U}^* \rightarrow \mathcal{V}^*$  by<sup>13</sup>

$$[\mathbf{T}^*(\mathbf{g})] |a\rangle = \mathbf{g}(\mathbf{T}|a\rangle) \quad \forall |a\rangle \in \mathcal{V}, \mathbf{g} \in \mathcal{U}^*,$$

$\mathbf{T}^*$  is called the dual or pull back, of  $\mathbf{T}$ .

One can readily verify that  $\mathbf{T}^* \in \mathcal{L}(\mathcal{U}^*, \mathcal{V}^*)$ , i.e., that  $\mathbf{T}^*$  is a linear operator on  $\mathcal{U}^*$ . Some of the mapping properties of  $\mathbf{T}^*$  are tied to those of  $\mathbf{T}$ . To see this we first consider the kernel of  $\mathbf{T}^*$ . Clearly,  $\mathbf{g}$  is in the kernel of  $\mathbf{T}^*$  if and only if  $\mathbf{g}$  annihilates all vectors of the form  $\mathbf{T}|a\rangle$ , i.e., all vectors in  $\mathbf{T}(\mathcal{V})$ . It follows that  $\mathbf{g}$  is in  $\mathbf{T}(\mathcal{V})^0$ . In particular, if  $\mathbf{T}$  is surjective,  $\mathbf{T}(\mathcal{V}) = \mathcal{U}$ , and  $\mathbf{g}$  annihilates all vectors in  $\mathcal{U}$ , i.e., it is the zero linear functional. We conclude that  $\ker \mathbf{T}^* = 0$ , and therefore,  $\mathbf{T}^*$  is injective. Similarly, one can show that if  $\mathbf{T}$  is injective, then  $\mathbf{T}^*$  is surjective. We summarize the discussion above:

**1.3.18. Proposition.** Let  $\mathbf{T}$  be a linear transformation and  $\mathbf{T}^*$  its pull back. Then  $\ker \mathbf{T}^* = \mathbf{T}(\mathcal{V})^0$ . If  $\mathbf{T}$  is surjective (injective), then  $\mathbf{T}^*$  is injective (surjective). In particular,  $\mathbf{T}^*$  is an isomorphism if  $\mathbf{T}$  is.

It is useful to make a connection between the inner product and linear functionals. To do this, consider a basis  $\{|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle\}$  and let  $\alpha_i = \langle a | a_i \rangle$ . As noted earlier, the set of scalars  $\{\alpha_i\}_{i=1}^N$  defines a unique linear functional  $\mathbf{f}_a$  such that  $\mathbf{f}_a |a_i\rangle = \alpha_i$ . Since  $\langle a | a_i \rangle$  is also equal to  $\alpha_i$ , it is natural to identify  $\mathbf{f}_a$  with

duals and inner products

<sup>13</sup>Do not confuse this “\*” with complex conjugation.

the symbol  $\langle a|$ , and write  $\mathbf{T} : \mathbf{f}_a \mapsto \langle a|$  where  $\mathbf{T}$  is the identification map.

It is also convenient to introduce the notation<sup>14</sup>

$$(|a\rangle)^\dagger \equiv \langle a|, \tag{1.12}$$

dagger of a linear combination of vectors

where the symbol  $\dagger$  means “dual, or dagger of.” Now we ask: How does this dagger operation act on a linear combination of vectors? Let  $|c\rangle = \alpha |a\rangle + \beta |b\rangle$  and take the inner product of  $|c\rangle$  with an arbitrary vector  $|x\rangle$  using linearity in the second factor:  $\langle x|c\rangle = \alpha \langle x|a\rangle + \beta \langle x|b\rangle$ . Now complex conjugate both sides and use the (sesqui)symmetry of the inner product:

$$\begin{aligned} (LHS)^* &= \langle x|c\rangle^* = \langle c|x\rangle, \\ (RHS)^* &= \alpha^* \langle x|a\rangle^* + \beta^* \langle x|b\rangle^* = \alpha^* \langle a|x\rangle + \beta^* \langle b|x\rangle \\ &= (\alpha^* \langle a| + \beta^* \langle b|) |x\rangle. \end{aligned}$$

Since this is true for all  $|x\rangle$ , we must have  $(|c\rangle)^\dagger \equiv \langle c| = \alpha^* \langle a| + \beta^* \langle b|$ . Therefore, in a duality “operation” the complex scalars must be conjugated. So, we have

$$(\alpha |a\rangle + \beta |b\rangle)^\dagger = \alpha^* \langle a| + \beta^* \langle b|. \tag{1.13}$$

Thus, unlike the association  $|a\rangle \leftrightarrow \mathbf{f}_a$  which is linear, the association  $\mathbf{f}_a \leftrightarrow \langle a|$  is not linear, but sesquilinear, i.e., the identification map  $\mathbf{T}$  mentioned above is also sesquilinear:

$$\mathbf{T}(\alpha \mathbf{f}_a + \beta \mathbf{f}_b) = \alpha^* \langle a| + \beta^* \langle b| = \alpha^* \mathbf{T}(\mathbf{f}_a) + \beta^* \mathbf{T}(\mathbf{f}_b);$$

It is convenient to represent  $|a\rangle \in \mathbb{C}^n$  as a column vector

$$|a\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Then the definition of the complex inner product suggests that the dual of  $|a\rangle$  must be represented as a row vector with complex conjugate entries:

$$\langle a| = (\alpha_1^* \quad \alpha_2^* \quad \dots \quad \alpha_n^*), \tag{1.14}$$

and the inner product can be written as the (matrix) product

$$\langle a|b\rangle = (\alpha_1^* \quad \alpha_2^* \quad \dots \quad \alpha_n^*) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \sum_{i=1}^n \alpha_i^* \beta_i.$$

Compare (1.14) with the comments after (1.9). The complex conjugation in (1.14) is the result of the sesquilinearity of the association  $|a\rangle \leftrightarrow \langle a|$ .

<sup>14</sup>The significance of this notation will become clear in Section 2.3.

**1.3.19. Example.** Let  $\mathcal{U}$  and  $\mathcal{V}$  be vector spaces with bases  $B_U = \{|u_i\rangle\}_{i=1}^m$  and  $B_V = \{|v_j\rangle\}_{j=1}^n$ , respectively. Consider an  $mn$ -dimensional vector space  $\mathcal{W}$  whose basis  $B_W$  is in one-to-one correspondence with the pairs  $(|u_i\rangle, |v_j\rangle)$ , and let  $|u_i v_j\rangle$  be the vector corresponding to  $(|u_i\rangle, |v_j\rangle)$ . For  $|u\rangle \in \mathcal{U}$  with components  $\{\alpha_i\}_{i=1}^m$  in  $B_U$  and  $|v\rangle \in \mathcal{V}$  with components  $\{\eta_j\}_{j=1}^n$  in  $B_V$ , define the vector  $|u, v\rangle \in \mathcal{W}$  whose components in  $B_W$  are  $\{\alpha_i \eta_j\}$ . One can easily show that if  $|u\rangle, |u'\rangle$ , and  $|u''\rangle$  are vectors in  $\mathcal{U}$  and  $|u''\rangle = \alpha |u\rangle + \beta |u'\rangle$ , then

$$|u'', v\rangle = \alpha |u, v\rangle + \beta |u', v\rangle$$

The space  $\mathcal{W}$  thus defined is called the **tensor product** of  $\mathcal{U}$  and  $\mathcal{V}$  and denoted by  $\mathcal{U} \otimes \mathcal{V}$ .

One can also define tensor product of three or more vector spaces. Of special interest are tensor products of a vector space and its dual. The **tensor product space**  $\mathcal{V}_{r,s}$  of type  $(r, s)$  is defined as follows:

$$\mathcal{V}_{r,s} = \underbrace{\mathcal{V} \otimes \mathcal{V} \otimes \dots \otimes \mathcal{V}}_{r \text{ times}} \otimes \underbrace{\mathcal{V}^* \otimes \mathcal{V}^* \otimes \dots \otimes \mathcal{V}^*}_{s \text{ times}}$$

We shall come back to this space in Chapter 25. ■

## 1.4 Algebras

In many physical applications, a vector space has a natural “product,” the prime example being the vector space of matrices. It is therefore useful to consider vector spaces for which such a product exists.

algebra defined

**1.4.1. Definition.** An **algebra**  $\mathcal{A}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ) is a vector space over  $\mathbb{C}$  (or  $\mathbb{R}$ ), together with a binary operation  $\mu : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$ , called **multiplication**, that satisfies<sup>15</sup>

$$\mathbf{a}(\beta \mathbf{b} + \gamma \mathbf{c}) = \beta \mathbf{a} \mathbf{b} + \gamma \mathbf{a} \mathbf{c} \quad \forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{A}, \forall \beta, \gamma \in \mathbb{C} \text{ (or } \mathbb{R}\text{)},$$

with a similar relation for multiplication on the right. The dimension of the vector space is called the **dimension of the algebra**. The algebra is called **associative** if the product satisfies  $\mathbf{a}(\mathbf{b} \mathbf{c}) = (\mathbf{a} \mathbf{b}) \mathbf{c}$  and **commutative** if it satisfies  $\mathbf{a} \mathbf{b} = \mathbf{b} \mathbf{a}$ . An algebra with **identity** is an algebra that has an element  $\mathbf{1}$  satisfying  $\mathbf{a} \mathbf{1} = \mathbf{1} \mathbf{a} = \mathbf{a}$ . An element  $\mathbf{b}$  of an algebra with identity is said to be a **left inverse** of  $\mathbf{a}$  if  $\mathbf{b} \mathbf{a} = \mathbf{1}$ . **Right inverse** is defined similarly.

dimension of the algebra; associativity; commutativity; identity; and right and left inverses

**1.4.2. Example.** Define the following product on  $\mathbb{R}^2$ :

$$(x_1, x_2)(y_1, y_2) = (x_1 y_1 - x_2 y_2, x_1 y_2 + x_2 y_1).$$

The reader is urged to verify that this product turns  $\mathbb{R}^2$  into a commutative algebra.

<sup>15</sup>We shall abandon the Dirac bra-and-ket notation in this section due to its clumsiness; instead we use boldface roman letters to denote vectors. It is customary to write  $\mathbf{a} \mathbf{b}$  for  $\mu(\mathbf{a}, \mathbf{b})$ .

Similarly, the vector (cross) product on  $\mathbb{R}^3$  turns it into a nonassociative, noncommutative algebra.

The paradigm of all algebras is the **matrix algebra** whose binary operation is ordinary multiplication of  $n \times n$  matrices. This algebra is associative but not commutative.

All the examples above are finite-dimensional algebras. An example of an infinite-dimensional algebra is  $\mathcal{C}^\infty(a, b)$ , the vector space of infinitely differentiable real-valued functions on a real interval  $(a, b)$ . The multiplication is defined pointwise: If  $f \in \mathcal{C}^\infty(a, b)$  and  $g \in \mathcal{C}^\infty(a, b)$ , then

$$fg(x) \equiv f(x)g(x) \quad \forall x \in (a, b).$$

This algebra is commutative and associative. ■

The last item in the example above has a feature that turns out to be of great significance in all algebras, the product rule for differentiation.

derivation of an algebra defined

**1.4.3. Definition.** A vector space endomorphism  $\mathbf{D} : \mathcal{A} \rightarrow \mathcal{A}$  is called a **derivation** on  $\mathcal{A}$  if it has the additional property

$$\mathbf{D}(\mathbf{ab}) = [\mathbf{D}(\mathbf{a})]\mathbf{b} + \mathbf{a}[\mathbf{D}(\mathbf{b})].$$

**1.4.4. Example.** Let  $\mathcal{A}$  be the set of  $n \times n$  matrices. Define the binary operation, denoted by  $\bullet$ , as

$$\mathbf{A} \bullet \mathbf{B} \equiv \mathbf{AB} - \mathbf{BA},$$

where the RHS is ordinary matrix multiplication. The reader may check that  $\mathcal{A}$  together with this operation becomes an algebra. Now let  $\mathbf{A}$  be a fixed matrix, and define the linear transformation

$$\mathbf{D}_\mathbf{A}(\mathbf{B}) = \mathbf{A} \bullet \mathbf{B}.$$

Then we note that

$$\begin{aligned} \mathbf{D}_\mathbf{A}(\mathbf{B} \bullet \mathbf{C}) &= \mathbf{A} \bullet (\mathbf{B} \bullet \mathbf{C}) = \mathbf{A}(\mathbf{B} \bullet \mathbf{C}) - (\mathbf{B} \bullet \mathbf{C})\mathbf{A} \\ &= \mathbf{A}(\mathbf{BC} - \mathbf{CB}) - (\mathbf{BC} - \mathbf{CB})\mathbf{A} = \mathbf{ABC} - \mathbf{ACB} - \mathbf{BCA} + \mathbf{CBA}. \end{aligned}$$

On the other hand,

$$\begin{aligned} (\mathbf{D}_\mathbf{A}\mathbf{B}) \bullet \mathbf{C} + \mathbf{B} \bullet (\mathbf{D}_\mathbf{A}\mathbf{C}) &= (\mathbf{A} \bullet \mathbf{B}) \bullet \mathbf{C} + \mathbf{B} \bullet (\mathbf{A} \bullet \mathbf{C}) \\ &= (\mathbf{AB} - \mathbf{BA}) \bullet \mathbf{C} + \mathbf{B} \bullet (\mathbf{AC} - \mathbf{CA}) \\ &= (\mathbf{AB} - \mathbf{BA})\mathbf{C} - \mathbf{C}(\mathbf{AB} - \mathbf{BA}) + \mathbf{B}(\mathbf{AC} - \mathbf{CA}) - (\mathbf{AC} - \mathbf{CA})\mathbf{B} \\ &= \mathbf{ABC} + \mathbf{CBA} - \mathbf{BCA} - \mathbf{ACB}. \end{aligned}$$

So,  $\mathbf{D}_\mathbf{A}$  is a derivation on  $\mathcal{A}$ . ■

The linear transformations connecting vector spaces can be modified slightly to accommodate the binary operation of multiplication of the corresponding algebras:

algebra  
homomorphism and  
isomorphism

**1.4.5. Definition.** Let  $\mathcal{A}$  and  $\mathcal{B}$  be algebras. A linear transformation  $\mathbf{T} : \mathcal{A} \rightarrow \mathcal{B}$  is called an **algebra homomorphism** if  $\mathbf{T}(\mathbf{ab}) = \mathbf{T}(\mathbf{a})\mathbf{T}(\mathbf{b})$ . A bijective algebra homomorphism is called an **algebra isomorphism**.

**1.4.6. Example.** Let  $\mathcal{A}$  be  $\mathbb{R}^3$ , and  $\mathcal{B}$  the set of  $3 \times 3$  matrices of the form

$$A = \begin{pmatrix} 0 & a_1 & -a_2 \\ -a_1 & 0 & a_3 \\ a_2 & -a_3 & 0 \end{pmatrix}.$$

Then the map  $\mathbf{T} : \mathcal{A} \rightarrow \mathcal{B}$  defined by

$$\mathbf{T}(\mathbf{a}) = \mathbf{T}(a_1, a_2, a_3) = \begin{pmatrix} 0 & a_1 & -a_2 \\ -a_1 & 0 & a_3 \\ a_2 & -a_3 & 0 \end{pmatrix}$$

can be shown to be a linear isomorphism. Let the cross product be the binary operation on  $\mathcal{A}$ , turning it into an algebra. For  $\mathcal{B}$ , define the binary operation of Example 1.4.4. The reader may check that, with these operations,  $\mathbf{T}$  is extended to an algebra isomorphism. ■

structure constants  
of an algebra

Given an algebra  $\mathcal{A}$  and a basis  $B = \{\mathbf{e}_i\}_{i=1}^N$  for the underlying vector space, one can write

$$\mathbf{e}_i \mathbf{e}_j = \sum_{k=1}^N c_{ij}^k \mathbf{e}_k, \quad c_{ij}^k \in \mathbb{C}. \tag{1.15}$$

The complex numbers  $c_{ij}^k$ , the components of the vector  $\mathbf{e}_i \mathbf{e}_j$  in the basis  $B$ , are called the **structure constants** of  $\mathcal{A}$ . These constants determine the product of any two vectors once they are expressed in terms of the basis vectors of  $B$ . Conversely,

**1.4.7. Box.** Given any  $N$ -dimensional vector space  $\mathcal{V}$ , one can turn it into an algebra by choosing a basis and a set of  $N^3$  numbers  $\{c_{ij}^k\}$  and defining the product of basis vectors by Equation (1.15).

**1.4.8. Example.** Consider the vector space of  $n \times n$  matrices with its standard basis  $\{\mathbf{e}_{ij}\}_{i,j=1}^n$ , where  $\mathbf{e}_{ij}$  has a 1 at the  $ij$ th position and zero everywhere else. This means that  $(\mathbf{e}_{ij})_{lk} = \delta_{il} \delta_{jk}$ , and

$$(\mathbf{e}_{ij} \mathbf{e}_{kl})_{mn} = \sum_{r=1}^n (\mathbf{e}_{ij})_{mr} (\mathbf{e}_{kl})_{rn} = \sum_{r=1}^n \delta_{im} \delta_{jr} \delta_{kr} \delta_{ln} = \delta_{im} \delta_{jk} \delta_{ln} = \delta_{jk} (\mathbf{e}_{il})_{mn},$$

or

$$\mathbf{e}_{ij} \mathbf{e}_{kl} = \delta_{jk} \mathbf{e}_{il}.$$

The structure constants are  $c_{ij,kl}^{mn} = \delta_{im} \delta_{jk} \delta_{ln}$ . Note that one needs a double index to label these constants. ■

**1.4.9. Example.** In the standard basis  $\{e_i\}$  of  $\mathbb{R}^4$ , choose the structure constants as follows:

$$e_1^2 = -e_2^2 = -e_3^2 = -e_4^2 = e_1,$$

$$e_1 e_i = e_i e_1 = e_i \text{ for } i = 2, 3, 4,$$

$$e_i e_j = \sum_k \epsilon_{ijk} e_k \text{ for } i, j = 2, 3, 4,$$

[ $\epsilon_{ijk}$  is defined in Equation (3.19)]. The reader may verify that these relations turn  $\mathbb{R}^4$  into an associative, but noncommutative, algebra, called the **algebra of quaternions** and denoted by  $\mathbb{H}$ . In this context,  $e_1$  is usually denoted by 1, and  $e_2$ ,  $e_3$ , and  $e_4$  by  $i$ ,  $j$ , and  $k$ , respectively, and one writes  $q = x + iy + jz + kw$  for an element of  $\mathbb{H}$ . It then becomes evident that  $\mathbb{H}$  is a generalization of  $\mathbb{C}$ . In analogy with  $\mathbb{C}$ ,  $x$  is called the **real part** of  $q$ , and  $(y, z, w)$  the **pure part** of  $q$ . Similarly, the **conjugate** of  $q$  is  $q^* = x - iy - jz - kw$ . ■

algebra of quaternions

Algebras have a surprisingly rich structure and are used extensively in many branches of mathematics and physics. We shall see their usefulness in our discussion of group theory in Part VII. To close this section, and to complete this introductory discussion of algebras, we cite a few more useful notions.

left, right, and two-sided ideals

**1.4.10. Definition.** Let  $A$  be an algebra. A subspace  $\mathcal{B}$  of  $A$  is called a **subalgebra** of  $A$  if  $\mathcal{B}$  contains the products of all its members. If  $\mathcal{B}$  has the extra property that it contains  $\mathbf{ab}$  for all  $\mathbf{a} \in A$  and  $\mathbf{b} \in \mathcal{B}$ , then  $\mathcal{B}$  is called a **left ideal** of  $A$ . A **right ideal** and a **two-sided ideal** are defined similarly.

It is clear from the definition that an ideal is automatically a subalgebra, and that

**1.4.11. Box.** No proper ideal of an algebra with identity can contain the identity element.

minimal ideal

In fact, no proper left (right) ideal can contain an element that has a left (right) inverse. An ideal can itself contain a proper (sub)ideal. If an ideal does not contain any proper subideal, it is called a **minimal ideal**.

**1.4.12. Example.** The vector space  $\mathcal{C}^0(a, b)$  of all continuous real-valued functions on an interval  $(a, b)$  is turned into a commutative algebra by pointwise multiplication: If  $f, g \in \mathcal{C}^0(a, b)$ , then the product  $fg$  is defined by  $fg(x) \equiv f(x)g(x)$  for all  $x \in (a, b)$ . The set of functions that vanish at a given fixed point  $c \in (a, b)$  constitutes an ideal in  $\mathcal{C}^0(a, b)$ . Since the algebra is commutative, the ideal is two-sided. ■

ideals generated by an element of an algebra

One can easily construct left ideals for an algebra  $\mathcal{A}$ : Take any element  $\mathbf{x} \in \mathcal{A}$  and consider the set

$$\mathcal{A}\mathbf{x} \equiv \{\mathbf{ax} \mid \mathbf{a} \in \mathcal{A}\}.$$



The reader may check that  $\mathcal{A}x$  is a left ideal. A right ideal can be constructed similarly. To construct a two-sided ideal, consider the set

$$\mathcal{A}x\mathcal{A} \equiv \{\mathbf{axb} \mid \mathbf{a}, \mathbf{b} \in \mathcal{A}\}.$$

These are all called **ideals generated by  $x$** .

## 1.5 Problems

**1.1.** Let  $\mathbb{R}^+$  denote the set of positive real numbers. Define the “sum” of two elements of  $\mathbb{R}^+$  to be their usual product, and define scalar multiplication by elements of  $\mathbb{R}$  as being given by  $r \cdot p = p^r$  where  $r \in \mathbb{R}$  and  $p \in \mathbb{R}^+$ . With these operations, show that  $\mathbb{R}^+$  is a vector space over  $\mathbb{R}$ .

**1.2.** Show that the intersection of two subspaces is also a subspace.

**1.3.** For each of the following subsets of  $\mathbb{R}^3$  determine whether it is a subspace of  $\mathbb{R}^3$ :

- (a)  $\{(x, y, z) \in \mathbb{R}^3 \mid x + y - 2z = 0\}$ ;
- (b)  $\{(x, y, z) \in \mathbb{R}^3 \mid x + y - 2z = 3\}$ ;
- (c)  $\{(x, y, z) \in \mathbb{R}^3 \mid xyz = 0\}$ .

**1.4.** Prove that the components of a vector in a given basis are unique.

**1.5.** Show that the following vectors form a basis for  $\mathbb{C}^n$  (or  $\mathbb{R}^n$ ).

$$|a_1\rangle = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}, \quad |a_2\rangle = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad \dots, \quad |a_n\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

**1.6.** Let  $\mathcal{W}$  be a subspace of  $\mathbb{R}^5$  defined by

$$\mathcal{W} = \{(x_1, \dots, x_5) \in \mathbb{R}^5 \mid x_1 = 3x_2 + x_3, x_2 = x_5, \text{ and } x_4 = 2x_3\}.$$

Find a basis for  $\mathcal{W}$ .

**1.7.** Show that the inner product of any vector with  $|0\rangle$  is zero.

**1.8.** Prove Theorem 1.1.5.

**1.9.** Find  $a_0, b_0, b_1, c_0, c_1,$  and  $c_2$  such that the polynomials  $a_0, b_0 + b_1t,$  and  $c_0 + c_1t + c_2t^2$  are mutually orthonormal in the interval  $[0, 1]$ . The inner product is as defined for polynomials in Example 1.2.3 with  $w(t) = 1$ .

**1.10.** Given the linearly independent vectors  $x(t) = t^n$ , for  $n = 0, 1, 2, \dots$  in  $\mathcal{P}^c[t]$ , use the Gram–Schmidt process to find the orthonormal polynomials  $e_0(t)$ ,  $e_1(t)$ , and  $e_2(t)$

(a) when the inner product is defined as  $\langle x | y \rangle = \int_{-1}^1 x^*(t)y(t) dt$ .

(b) when the inner product is defined with a nontrivial weight function:

$$\langle x | y \rangle = \int_{-\infty}^{\infty} e^{-t^2} x^*(t)y(t) dt.$$

Hint: Use the following result:

$$\int_{-\infty}^{\infty} e^{-t^2} t^n dt = \begin{cases} \sqrt{\pi} & \text{if } n = 0, \\ 0 & \text{if } n \text{ is odd,} \\ \sqrt{\pi} \frac{1 \cdot 3 \cdot 5 \cdots (n-1)}{2^{n/2}} & \text{if } n \text{ is even.} \end{cases}$$

**1.11.** (a) Use the Gram–Schmidt process to find an orthonormal set of vectors out of  $(1, -1, 1)$ ,  $(-1, 0, 1)$ , and  $(2, -1, 2)$ .

(b) Are these three vectors linearly independent? If not, find a zero linear combination of them by using part (a).

**1.12.** (a) Use the Gram–Schmidt process to find an orthonormal set of vectors out of  $(1, -1, 2)$ ,  $(-2, 1, -1)$ , and  $(-1, -1, 4)$ .

(b) Are these three vectors linearly independent? If not, find a zero linear combination of them by using part (a).

**1.13.** Show that

$$\begin{aligned} & \int_{-\infty}^{\infty} (t^{10} - t^6 + 5t^4 - 5)e^{-t^4} dt \\ & \leq \sqrt{\int_{-\infty}^{\infty} (t^4 - 1)^2 e^{-t^4} dt} \sqrt{\int_{-\infty}^{\infty} (t^6 + 5)^2 e^{-t^4} dt}. \end{aligned}$$

**1.14.** Show that

$$\begin{aligned} & \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy (x^5 - x^3 + 2x^2 - 2)(y^5 - y^3 + 2y^2 - 2)e^{-(x^4+y^4)} \\ & \leq \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy (x^4 - 2x^2 + 1)(y^6 + 4y^3 + 4)e^{-(x^4+y^4)}. \end{aligned}$$

Hint: Define an appropriate inner product and use the Schwarz inequality.

**1.15.** Show that for any set of  $n$  complex numbers  $\alpha_1, \alpha_2, \dots, \alpha_n$ , we have

$$|\alpha_1 + \alpha_2 + \cdots + \alpha_n|^2 \leq n(|\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_n|^2).$$

Hint: Apply the Schwarz inequality to  $(1, 1, \dots, 1)$  and  $(\alpha_1, \alpha_2, \dots, \alpha_n)$ .

**1.16.** Using the Schwarz inequality show that if  $\{\alpha_i\}_{i=1}^{\infty}$  and  $\{\beta_i\}_{i=1}^{\infty}$  are in  $\mathbb{C}^{\infty}$ , then  $\sum_{i=1}^{\infty} \alpha_i^* \beta_i$  is convergent.

**1.17.** Show that  $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  given by  $\mathbf{T}(x, y) = (x^2 + y^2, x + y, 2x - y)$  is not a linear mapping.

**1.18.** Verify that all the transformations of Example 1.3.4 are linear.

**1.19.** Let  $\pi$  be the permutation that takes  $(1, 2, 3)$  to  $(3, 1, 2)$ . Find

$$\mathbf{A}_{\pi} |e_i\rangle, \quad i = 1, 2, 3,$$

where  $\{|e_i\rangle\}_{i=1}^3$  is the standard basis of  $\mathbb{R}^3$  (or  $\mathbb{C}^3$ ), and  $\mathbf{A}_{\pi}$  is as defined in Example 1.3.4.

**1.20.** Show that if  $\mathbf{T} \in \mathcal{L}(\mathbb{C}, \mathbb{C})$ , then there exists  $\alpha \in \mathbb{C}$  such that  $\mathbf{T}|a\rangle = \alpha|a\rangle$  for all  $|a\rangle \in \mathbb{C}$ .

**1.21.** Show that if  $\{|a_i\rangle\}_{i=1}^n$  spans  $\mathcal{V}$  and  $\mathbf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$  is surjective, then  $\{\mathbf{T}|a_i\rangle\}_{i=1}^n$  spans  $\mathcal{W}$ .

**1.22.** Give an example of a function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that

$$f(\alpha|a\rangle) = \alpha f(|a\rangle) \quad \forall \alpha \in \mathbb{R} \text{ and } |a\rangle \in \mathbb{R}^2$$

but  $f$  is not linear. Hint: Consider a homogeneous function of degree 1.

**1.23.** Show that the following transformations are linear:

(a)  $\mathcal{V}$  is  $\mathbb{C}$  over the reals and  $\mathbf{C}|z\rangle = |z^*\rangle$ . Is  $\mathbf{C}$  linear if instead of real numbers, complex numbers are used as scalars?

(b)  $\mathcal{V}$  is  $\mathcal{P}^c[t]$  and  $\mathbf{T}|x(t)\rangle = |x(t+1)\rangle - |x(t)\rangle$ .

**1.24.** Verify that the kernel of a transformation  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  is a subspace of  $\mathcal{V}$ , and that  $\mathbf{T}(\mathcal{V})$  is a subspace of  $\mathcal{W}$ .

**1.25.** Let  $\mathcal{V}$  and  $\mathcal{W}$  be finite dimensional vector spaces. Show that if  $\mathbf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$  is surjective, then  $\dim \mathcal{W} \leq \dim \mathcal{V}$ .

**1.26.** Suppose that  $\mathcal{V}$  is finite dimensional and  $\mathbf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$  is not zero. Prove that there exists a subspace  $\mathcal{U}$  of  $\mathcal{V}$  such that  $\ker \mathbf{T} \cap \mathcal{U} = \{0\}$  and  $\mathbf{T}(\mathcal{V}) = \mathbf{T}(\mathcal{U})$ .

**1.27.** Show that  $\mathcal{W}^0$  is a subspace of  $\mathcal{V}^*$  and

$$\dim \mathcal{V} = \dim \mathcal{W} + \dim \mathcal{W}^0.$$

**1.28.** Show that  $\mathbf{T}$  and  $\mathbf{T}^*$  have the same rank. In particular, show that if  $\mathbf{T}$  is injective, then  $\mathbf{T}^*$  is surjective. Hint: Use the dimension theorem for  $\mathbf{T}$  and  $\mathbf{T}^*$  and Equation (1.11).

1.29. Show that (a) the product on  $\mathbb{R}^2$  defined by

$$(x_1, x_2)(y_1, y_2) = (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1)$$

turns  $\mathbb{R}^2$  into an associative and commutative algebra, and (b) the cross product on  $\mathbb{R}^3$  turns it into a nonassociative, noncommutative algebra.

1.30. Fix a vector  $\mathbf{a} \in \mathbb{R}^3$  and define the linear transformation  $\mathbf{D}_\mathbf{a} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  by  $\mathbf{D}_\mathbf{a}(\mathbf{b}) = \mathbf{a} \times \mathbf{b}$ . Show that  $\mathbf{D}_\mathbf{a}$  is a derivation of  $\mathbb{R}^3$  with the cross product as multiplication.

1.31. Show that the linear transformation of Example 1.4.6 is an isomorphism of the two algebras  $\mathcal{A}$  and  $\mathcal{B}$ .

1.32. Write down all the structure constants for the algebra of quaternions. Show that this algebra is associative.

1.33. Show that a quaternion is real iff it commutes with every quaternion and that it is pure iff its square is a nonpositive real number.

1.34. Let  $p$  and  $q$  be two quaternions. Show that

(a)  $(pq)^* = q^*p^*$ .

(b)  $q \in \mathbb{R}$  iff  $q^* = q$ , and  $q \in \mathbb{R}^3$  iff  $q^* = -q$ .

(c)  $qq^* = q^*q$  is a nonnegative real number.

1.35. Show that no *proper* left (right) ideal of an algebra with identity can contain an element that has a left (right) inverse.

1.36. Let  $\mathcal{A}$  be an algebra, and  $\mathbf{x} \in \mathcal{A}$ . Show that  $\mathcal{A}\mathbf{x}$  is a left ideal,  $\mathbf{x}\mathcal{A}$  is a right ideal, and  $\mathcal{A}\mathbf{x}\mathcal{A}$  is a two-sided ideal.

## Additional Reading

1. Axler, S. *Linear Algebra Done Right*, Springer-Verlag, 1996. A small text packed with information. Lots of marginal notes and historical remarks.
2. Greub, W. *Linear Algebra*, 4th ed., Springer-Verlag, 1975. Chapter V has a good discussion of algebras and their properties.
3. Halmos, P. *Finite Dimensional Vector Spaces*, 2nd ed., Van Nostrand, 1958. Another great book by the master expositor.

# Operator Algebra

Recall that a vector space in which one can multiply two vectors to obtain a third vector is called an algebra. In this chapter, we want to investigate the algebra of linear transformations. We have already established that the set of linear transformations  $L(\mathcal{V}, \mathcal{W})$  from  $\mathcal{V}$  to  $\mathcal{W}$  is a vector space. Let us attempt to define a multiplication as well. The best candidate is the composition of linear transformations. If  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{U}$  and  $\mathbf{S} : \mathcal{U} \rightarrow \mathcal{W}$  are linear operators, then the composition  $\mathbf{S} \circ \mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  is also a linear operator, as can easily be verified.

This product, however, is not defined on a single vector space, but is such that it takes an element in  $L(\mathcal{V}, \mathcal{U})$  and another element in a second vector space  $L(\mathcal{U}, \mathcal{W})$  to give an element in yet another vector space  $L(\mathcal{V}, \mathcal{W})$ . An algebra requires a single vector space. We can accomplish this by letting  $\mathcal{V} = \mathcal{U} = \mathcal{W}$ . Then the three spaces of linear transformations collapse to the single space  $L(\mathcal{V}, \mathcal{V})$ , the set of endomorphisms of  $\mathcal{V}$ , which we abbreviate as  $\mathcal{L}(\mathcal{V})$  and to which  $\mathbf{T}$ ,  $\mathbf{S}$ ,  $\mathbf{ST} \equiv \mathbf{S} \circ \mathbf{T}$ , and  $\mathbf{TS} \equiv \mathbf{T} \circ \mathbf{S}$  belong. The space  $\mathcal{L}(\mathcal{V})$  is the algebra of the linear operators on  $\mathcal{V}$ .

## 2.1 Algebra of $\mathcal{L}(\mathcal{V})$

Operator algebra encompasses various operations on, and relations among, operators. One of these relations is the equality of operators, which is intuitively obvious; nevertheless, we make it explicit in (see also Box 1.3.3) the following definition.

operator equality **2.1.1. Definition.** *Two linear operators  $\mathbf{T}, \mathbf{U} \in \mathcal{L}(\mathcal{V})$  are equal if  $\mathbf{T}|a\rangle = \mathbf{U}|a\rangle$  for all  $|a\rangle \in \mathcal{V}$ .*

Because of the linearity of  $\mathbf{T}$  and  $\mathbf{U}$ , we have

**2.1.2. Box.** Two endomorphisms  $\mathbf{T}, \mathbf{U} \in \mathcal{L}(\mathcal{V})$  are equal if  $\mathbf{T}|a_i\rangle = \mathbf{U}|a_i\rangle$  for all  $|a_i\rangle \in B$ , where  $B$  is a basis of  $\mathcal{V}$ . Therefore, an endomorphism is uniquely determined by its action on the vectors of a basis.

The equality of operators can also be established by other, more convenient, methods when an inner product is defined on the vector space. The following two theorems contain the essence of these alternatives.

**2.1.3. Theorem.** An endomorphism  $\mathbf{T}$  of an inner product space is  $\mathbf{0}$  if and only if<sup>1</sup>  $\langle b|\mathbf{T}|a\rangle = \langle b|\mathbf{T}a\rangle = 0$  for all  $|a\rangle$  and  $|b\rangle$ .

*Proof.* Clearly, if  $\mathbf{T} = \mathbf{0}$  then  $\langle b|\mathbf{T}|a\rangle = 0$ . Conversely, if  $\langle b|\mathbf{T}|a\rangle = 0$  for all  $|a\rangle$  and  $|b\rangle$ , then, choosing  $|b\rangle = \mathbf{T}|a\rangle = |\mathbf{T}a\rangle$ , we obtain

$$\langle \mathbf{T}a|\mathbf{T}a\rangle = 0 \quad \forall |a\rangle \quad \Leftrightarrow \quad \mathbf{T}|a\rangle = 0 \quad \forall |a\rangle \quad \Leftrightarrow \quad \mathbf{T} = \mathbf{0}$$

by positive definiteness of the inner product. □

**2.1.4. Theorem.** A linear operator  $\mathbf{T}$  on an inner product space is  $\mathbf{0}$  if and only if  $\langle a|\mathbf{T}|a\rangle = 0$  for all  $|a\rangle$ .

*Proof.* Obviously, if  $\mathbf{T} = \mathbf{0}$ , then  $\langle a|\mathbf{T}|a\rangle = 0$ . Conversely, choose a vector  $\alpha|a\rangle + \beta|b\rangle$ , sandwich  $\mathbf{T}$  between this vector and its dual, and rearrange terms to obtain what is known as the **polarization identity**:

polarization identity

$$\begin{aligned} \alpha^*\beta \langle a|\mathbf{T}|b\rangle + \alpha\beta^* \langle b|\mathbf{T}|a\rangle &= \langle \alpha a + \beta b|\mathbf{T}|\alpha a + \beta b\rangle \\ &\quad - |\alpha|^2 \langle a|\mathbf{T}|a\rangle - |\beta|^2 \langle b|\mathbf{T}|b\rangle. \end{aligned}$$

According to the assumption of the theorem, the RHS is zero. Thus, if we let  $\alpha = \beta = 1$  we obtain  $\langle a|\mathbf{T}|b\rangle + \langle b|\mathbf{T}|a\rangle = 0$ . Similarly, with  $\alpha = 1$  and  $\beta = i$  we get  $i \langle a|\mathbf{T}|b\rangle - i \langle b|\mathbf{T}|a\rangle = 0$ . These two equations give  $\langle a|\mathbf{T}|b\rangle = 0$  for all  $|a\rangle, |b\rangle$ . By Theorem 2.1.3,  $\mathbf{T} = \mathbf{0}$ . □

To show that two operators  $\mathbf{U}$  and  $\mathbf{T}$  are equal, one can either have them act on an arbitrary vector and show that they give the same result, or one verifies that  $\mathbf{U} - \mathbf{T}$  is the zero operator by means of one of the theorems above. Equivalently, one shows that  $\langle a|\mathbf{T}|b\rangle = \langle a|\mathbf{U}|b\rangle$  or  $\langle a|\mathbf{T}|a\rangle = \langle a|\mathbf{U}|a\rangle$  for all  $|a\rangle, |b\rangle$ . In addition to the zero element, which is present in all algebras,  $\mathcal{L}(\mathcal{V})$  has an identity element,  $\mathbf{1}$ , which satisfies the relation  $\mathbf{1}|a\rangle = |a\rangle$  for all  $|a\rangle \in \mathcal{V}$ . With  $\mathbf{1}$  in our possession, we can ask whether it is possible to find an operator  $\mathbf{T}^{-1}$  with the property that  $\mathbf{T}^{-1}\mathbf{T} = \mathbf{T}\mathbf{T}^{-1} = \mathbf{1}$ . Generally speaking, only bijective mappings have inverses. Therefore, *only automorphisms of a vector space are invertible*.

<sup>1</sup>It is convenient here to use the notation  $|\mathbf{T}a\rangle$  for  $\mathbf{T}|a\rangle$ . This would then allow us to write the dual of the vector as  $\langle \mathbf{T}a|$ , emphasizing that it is indeed the bra associated with  $\mathbf{T}|a\rangle$ .

**2.1.5. Example.** Let the linear operator  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  be defined by

$$T(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3, x_1 + x_3).$$

We want to see whether  $T$  is invertible and, if so, find its inverse.  $T$  has an inverse if and only if it is bijective. By the comments after Theorem 1.3.8 this is the case if and only if  $T$  is either surjective or injective. The latter is equivalent to  $\ker T = \{0\}$ . But  $\ker T$  is the set of all vectors satisfying  $T(x_1, x_2, x_3) = (0, 0, 0)$ , or

$$x_1 + x_2 = 0, \quad x_2 + x_3 = 0, \quad x_1 + x_3 = 0.$$

The reader may check that the unique solution to these equations is  $x_1 = x_2 = x_3 = 0$ . Thus, the only vector belonging to  $\ker T$  is the zero vector. Therefore,  $T$  has an inverse.

To find  $T^{-1}$  apply  $T^{-1}T = \mathbf{1}$  to  $(x_1, x_2, x_3)$ :

$$(x_1, x_2, x_3) = T^{-1}T(x_1, x_2, x_3) = T^{-1}(x_1 + x_2, x_2 + x_3, x_1 + x_3).$$

This equation demonstrates how  $T^{-1}$  acts on vectors. To make this more apparent, we let  $x_1 + x_2 = x$ ,  $x_2 + x_3 = y$ ,  $x_1 + x_3 = z$ , solve for  $x_1$ ,  $x_2$ , and  $x_3$  in terms of  $x$ ,  $y$ , and  $z$ , and substitute in the preceding equation to obtain

$$T^{-1}(x, y, z) = \frac{1}{2}(x - y + z, x + y - z, -x + y + z).$$

Rewriting this equation in terms of  $x_1$ ,  $x_2$ , and  $x_3$  gives

$$T^{-1}(x_1, x_2, x_3) = \frac{1}{2}(x_1 - x_2 + x_3, x_1 + x_2 - x_3, -x_1 + x_2 + x_3).$$

We can easily verify that  $T^{-1}T = \mathbf{1}$  and that  $TT^{-1} = \mathbf{1}$ . ■

The following theorem, whose proof is left as an exercise, describes some properties of the inverse operator.

**2.1.6. Theorem.** *The inverse of a linear operator is unique. If  $T$  and  $S$  are two invertible linear operators, then  $TS$  is also invertible, and*

$$(TS)^{-1} = S^{-1}T^{-1}.$$

*An endomorphism  $T : \mathcal{V} \rightarrow \mathcal{V}$  is invertible if and only if it carries a basis of  $\mathcal{V}$  onto another basis of  $\mathcal{V}$ .*

### 2.1.1 Polynomials of Operators

With products and sums of operators defined, we can construct polynomials of operators. We define powers of  $T$  inductively as  $T^m = TT^{m-1} = T^{m-1}T$  for all positive integers  $m \geq 1$ . The consistency of this equation (for  $m = 1$ ) demands that  $T^0 = \mathbf{1}$ . It follows that polynomials such as

$$p(T) = \alpha_0\mathbf{1} + \alpha_1T + \alpha_2T^2 + \cdots + \alpha_nT^n$$

can be defined.

**2.1.7. Example.** Let  $\mathbf{T}_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be the linear operator that rotates vectors in the  $xy$ -plane through the angle  $\theta$ , that is,

$$\mathbf{T}_\theta(x, y) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta).$$

We are interested in powers of  $\mathbf{T}_\theta$ :

$$\begin{aligned} \mathbf{T}_\theta^2(x, y) &= \mathbf{T}_\theta \left( \overbrace{x \cos \theta - y \sin \theta}^{x'} \overbrace{x \sin \theta + y \cos \theta}^{y'} \right) \\ &= (x' \cos \theta - y' \sin \theta, x' \sin \theta + y' \cos \theta) \\ &= ((x \cos \theta - y \sin \theta) \cos \theta - (x \sin \theta + y \cos \theta) \sin \theta, \\ &\quad (x \cos \theta - y \sin \theta) \sin \theta + (x \sin \theta + y \cos \theta) \cos \theta) \\ &= (x \cos 2\theta - y \sin 2\theta, x \sin 2\theta + y \cos 2\theta). \end{aligned}$$

Thus,  $\mathbf{T}_\theta^2$  rotates  $(x, y)$  by  $2\theta$ . Similarly, one can show that

$$\mathbf{T}_\theta^3(x, y) = (x \cos 3\theta - y \sin 3\theta, x \sin 3\theta + y \cos 3\theta),$$

and in general,  $\mathbf{T}_\theta^n(x, y) = (x \cos n\theta - y \sin n\theta, x \sin n\theta + y \cos n\theta)$ , which shows that  $\mathbf{T}_\theta^n$  is a rotation of  $(x, y)$  through the angle  $n\theta$ , that is,  $\mathbf{T}_\theta^n = \mathbf{T}_{n\theta}$ . This result could have been guessed because  $\mathbf{T}_\theta^n$  is equivalent to rotating  $(x, y)$   $n$  times, each time by an angle  $\theta$ . ■

Negative powers of an invertible linear operator  $\mathbf{T}$  are defined by  $\mathbf{T}^{-m} = (\mathbf{T}^{-1})^m$ . The exponents of  $\mathbf{T}$  satisfy the usual rules. In particular, for any two integers  $m$  and  $n$  (positive or negative),  $\mathbf{T}^m \mathbf{T}^n = \mathbf{T}^{m+n}$  and  $(\mathbf{T}^m)^n = \mathbf{T}^{mn}$ . The first relation implies that the inverse of  $\mathbf{T}^m$  is  $\mathbf{T}^{-m}$ . One can further generalize the exponent to include fractions and ultimately all real numbers; but we need to wait until Chapter 4, in which we discuss the spectral decomposition theorem.

**2.1.8. Example.** Let us evaluate  $\mathbf{T}_\theta^{-n}$  for the operator of the previous example. First, let us find  $\mathbf{T}_\theta^{-1}$  (see Figure 2.1). We are looking for an operator such that  $\mathbf{T}_\theta^{-1} \mathbf{T}_\theta(x, y) = (x, y)$ , or

$$\mathbf{T}_\theta^{-1}(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta) = (x, y). \quad (2.1)$$

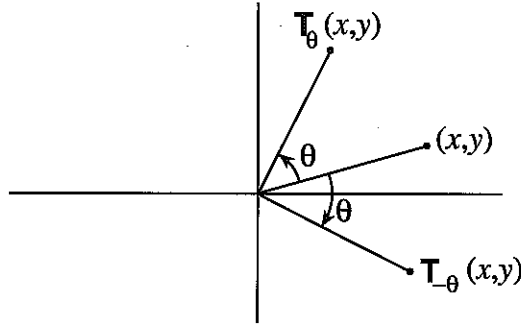
We define  $x' = x \cos \theta - y \sin \theta$  and  $y' = x \sin \theta + y \cos \theta$  and solve  $x$  and  $y$  in terms of  $x'$  and  $y'$  to obtain  $x = x' \cos \theta + y' \sin \theta$  and  $y = -x' \sin \theta + y' \cos \theta$ . Substituting for  $x$  and  $y$  in Equation (2.1) yields

$$\mathbf{T}_\theta^{-1}(x', y') = (x' \cos \theta + y' \sin \theta, -x' \sin \theta + y' \cos \theta).$$

Comparing this with the action of  $\mathbf{T}_\theta$  in the previous example, we discover that the only difference between the two operators is the sign of the  $\sin \theta$  term. We conclude that  $\mathbf{T}_\theta^{-1}$  has the same effect as  $\mathbf{T}_{-\theta}$ . So we have

$$\mathbf{T}_\theta^{-1} = \mathbf{T}_{-\theta} \quad \text{and} \quad \mathbf{T}_\theta^{-n} = (\mathbf{T}_\theta^{-1})^n = (\mathbf{T}_{-\theta})^n = \mathbf{T}_{-n\theta}.$$





**Figure 2.1** The operator  $T_\theta$  and its inverse as they act on a point in the plane.

It is instructive to verify that  $T_\theta^{-n} T_\theta^n = 1$ :

$$\begin{aligned}
 T_\theta^{-n} T_\theta^n(x, y) &= T_\theta^{-n} \left( \overbrace{x \cos n\theta - y \sin n\theta}^{x'} \overbrace{x \sin n\theta + y \cos n\theta}^{y'} \right) \\
 &= (x' \cos n\theta + y' \sin n\theta, -x' \sin n\theta + y' \cos n\theta) \\
 &= ((x \cos n\theta - y \sin n\theta) \cos n\theta + (x \sin n\theta + y \cos n\theta) \sin n\theta, \\
 &\quad - (x \cos n\theta - y \sin n\theta) \sin n\theta + (x \sin n\theta + y \cos n\theta) \cos n\theta) \\
 &= (x(\cos^2 n\theta + \sin^2 n\theta), y(\sin^2 n\theta + \cos^2 n\theta)) = (x, y).
 \end{aligned}$$

Similarly, we can show that  $T_\theta^n T_\theta^{-n}(x, y) = (x, y)$ . ■

One has to keep in mind that  $p(T)$  is not, in general, invertible, even if  $T$  is. In fact, the sum of two invertible operators is not necessarily invertible. For example, although  $T$  and  $-T$  are invertible, their sum, the zero operator, is not.

## 2.1.2 Functions of Operators

We can go one step beyond polynomials of operators and, via Taylor expansion, define functions of them. Consider an ordinary function  $f(x)$ , which has the Taylor expansion

$$f(x) = \sum_{k=0}^{\infty} \frac{(x - x_0)^k}{k!} \left. \frac{d^k f}{dx^k} \right|_{x=x_0}$$

in which  $x_0$  is a point where  $f(x)$  and all its derivatives are defined. To this function, there corresponds a function of the operator  $T$ , defined as

$$f(T) = \sum_{k=0}^{\infty} \left. \frac{d^k f}{dx^k} \right|_{x=x_0} \frac{(T - x_0 \mathbf{1})^k}{k!}. \quad (2.2)$$

Because this series is an infinite sum of operators, difficulties may arise concerning its convergence. However, as will be shown in Chapter 4,  $f(\mathbf{T})$  is always defined for finite-dimensional vector spaces. In fact, it is always a polynomial in  $\mathbf{T}$ . For the time being, we shall think of  $f(\mathbf{T})$  as a formal infinite series. A simplification results when the function can be expanded about  $x = 0$ . In this case we obtain

$$f(\mathbf{T}) = \sum_{k=0}^{\infty} \left. \frac{d^k f}{dx^k} \right|_{x=0} \frac{\mathbf{T}^k}{k!}. \quad (2.3)$$

A widely used function is the exponential, whose expansion is easily found to be

$$e^{\mathbf{T}} \equiv \exp(\mathbf{T}) = \sum_{k=0}^{\infty} \frac{\mathbf{T}^k}{k!}. \quad (2.4)$$

**2.1.9. Example.** Let us evaluate  $\exp(\alpha\mathbf{T})$  when  $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is given by

$$\mathbf{T}(x, y) = (-y, x).$$

We can find a general formula for the action of  $\mathbf{T}^n$  on  $(x, y)$ . Start with  $n = 2$ :

$$\mathbf{T}^2(x, y) = \mathbf{T}(-y, x) = (-x, -y) = -(x, y) = -\mathbf{1}(x, y).$$

Thus,  $\mathbf{T}^2 = -\mathbf{1}$ . From  $\mathbf{T}$  and  $\mathbf{T}^2$  we can easily obtain higher powers of  $\mathbf{T}$ . For example:  $\mathbf{T}^3 = \mathbf{T}(\mathbf{T}^2) = -\mathbf{T}$ ,  $\mathbf{T}^4 = \mathbf{T}^2\mathbf{T}^2 = \mathbf{1}$ , and in general,

$$\begin{aligned} \mathbf{T}^{2n} &= (-1)^n \mathbf{1} & \text{for } n = 0, 1, 2, \dots \\ \mathbf{T}^{2n+1} &= (-1)^n \mathbf{T} & \text{for } n = 0, 1, 2, \dots \end{aligned}$$

Thus,

$$\begin{aligned} \exp(\alpha\mathbf{T}) &= \sum_{n \text{ odd}} \frac{(\alpha\mathbf{T})^n}{n!} + \sum_{n \text{ even}} \frac{(\alpha\mathbf{T})^n}{n!} = \sum_{k=0}^{\infty} \frac{(\alpha\mathbf{T})^{2k+1}}{(2k+1)!} + \sum_{k=0}^{\infty} \frac{(\alpha\mathbf{T})^{2k}}{(2k)!} \\ &= \sum_{k=0}^{\infty} \frac{\alpha^{2k+1} \mathbf{T}^{2k+1}}{(2k+1)!} + \sum_{k=0}^{\infty} \frac{\alpha^{2k} \mathbf{T}^{2k}}{(2k)!} = \sum_{k=0}^{\infty} \frac{(-1)^k \alpha^{2k+1}}{(2k+1)!} \mathbf{T} + \sum_{k=0}^{\infty} \frac{(-1)^k \alpha^{2k}}{(2k)!} \mathbf{1} \\ &= \mathbf{T} \sum_{k=0}^{\infty} \frac{(-1)^k \alpha^{2k+1}}{(2k+1)!} + \mathbf{1} \sum_{k=0}^{\infty} \frac{(-1)^k \alpha^{2k}}{(2k)!}. \end{aligned}$$

The two series are recognized as  $\sin \alpha$  and  $\cos \alpha$ , respectively. Therefore, we get

$$e^{\alpha\mathbf{T}} = \mathbf{T} \sin \alpha + \mathbf{1} \cos \alpha,$$

which shows that  $e^{\alpha\mathbf{T}}$  is a polynomial (of first degree) in  $\mathbf{T}$ .

The action of  $e^{\alpha\mathbf{T}}$  on  $(x, y)$  is given by

$$\begin{aligned} e^{\alpha\mathbf{T}}(x, y) &= (\sin \alpha \mathbf{T} + \cos \alpha \mathbf{1})(x, y) = \sin \alpha \mathbf{T}(x, y) + \cos \alpha \mathbf{1}(x, y) \\ &= (\sin \alpha)(-y, x) + (\cos \alpha)(x, y) \\ &= (-y \sin \alpha, x \sin \alpha) + (x \cos \alpha, y \cos \alpha) \\ &= (x \cos \alpha - y \sin \alpha, x \sin \alpha + y \cos \alpha). \end{aligned}$$

The reader will recognize the final expression as a rotation in the  $xy$ -plane through an angle  $\alpha$ . Thus, we can think of  $e^{\alpha T}$  as a rotation operator of angle  $\alpha$  about the  $z$ -axis. In this context  $T$  is called the *generator* of the rotation. ■

### 2.1.3 Commutators

The result of multiplication of two operators depends on the order in which the operators appear. This means that if  $T, U \in \mathcal{L}(V)$ , then  $TU \in \mathcal{L}(V)$  and  $UT \in \mathcal{L}(V)$ ; however, in general  $UT \neq TU$ . When this is the case, we say that  $U$  and  $T$  *do not commute*. The extent to which two operators fail to commute is given in the following definition.

commutator defined

**2.1.10. Definition.** The *commutator*  $[U, T]$  of the two operators  $U$  and  $T$  in  $\mathcal{L}(V)$  is another operator in  $\mathcal{L}(V)$ , defined as

$$[U, T] \equiv UT - TU.$$

An immediate consequence of this definition is the following:

**2.1.11. Proposition.** For  $S, T, U \in \mathcal{L}(V)$  and  $\alpha, \beta \in \mathbb{C}$  (or  $\mathbb{R}$ ), we have

$[U, T] = -[T, U],$	<i>antisymmetry</i>
$[\alpha U, \beta T] = \alpha\beta[U, T],$	<i>linearity</i>
$[S, T + U] = [S, T] + [S, U],$	<i>linearity in the right entry</i>
$[S + T, U] = [S, U] + [T, U],$	<i>linearity in the left entry</i>
$[ST, U] = S[T, U] + [S, U]T,$	<i>right derivation property</i>
$[S, TU] = [S, T]U + T[S, U],$	<i>left derivation property</i>
$[S, T], U] + [[U, S], T] + [[T, U], S] = 0.$	<i>Jacobi identity</i>

*Proof.* In almost all cases the proof follows immediately from the definition. The only minor exceptions are the derivation properties. We prove the left derivation property:

$$\begin{aligned} [S, TU] &= S(TU) - (TU)S = STU - TUS + \underbrace{TSU - TSU}_{=0} \\ &= (ST - TS)U + T(SU - US) = [S, T]U + T[S, U]. \end{aligned}$$

The right derivation property is proved in exactly the same way. □

A useful consequence of the definition and Proposition 2.1.11 is

$$[A, A^m] = 0 \quad \text{for } m = 0, \pm 1, \pm 2, \dots$$

In particular,  $[A, 1] = 0$  and  $[A, A^{-1}] = 0$ .

## 2.2 Derivatives of Functions of Operators

Up to this point we have been discussing the algebraic properties of operators, static objects that obey certain algebraic rules and fulfill the static needs of some applications. However, physical quantities are dynamic, and if we want operators to represent physical quantities, we must allow them to change with time. This dynamism is best illustrated in quantum mechanics, where physical observables are represented by operators.

a time-dependent operator does not commute with itself at different times

Let us consider a mapping  $\mathbf{H} : \mathbb{R} \rightarrow \mathcal{L}(\mathcal{V})$ , which<sup>2</sup> takes in a real number and gives out a linear operator on the vector space  $\mathcal{V}$ . We denote the image of  $t \in \mathbb{R}$  by  $\mathbf{H}(t)$ , which acts on the underlying vector space  $\mathcal{V}$ . The physical meaning of this is that as  $t$  (usually time) varies, its image  $\mathbf{H}(t)$  also varies. Therefore, for different values of  $t$ , we have *different* operators. In particular,  $[\mathbf{H}(t), \mathbf{H}(t')] \neq \mathbf{0}$  for  $t \neq t'$ . A concrete example is an operator that is a linear combination of the operators  $\mathbf{D}$  and  $\mathbf{T}$  introduced in Example 1.3.4, with time-dependent scalars. To be specific, let  $\mathbf{H}(t) = \mathbf{D} \cos \omega t + \mathbf{T} \sin \omega t$ , where  $\omega$  is a constant. As time passes,  $\mathbf{H}(t)$  changes its identity from  $\mathbf{D}$  to  $\mathbf{T}$  and back to  $\mathbf{D}$ . Most of the time it has a hybrid identity! Since  $\mathbf{D}$  and  $\mathbf{T}$  do not commute, values of  $\mathbf{H}(t)$  for different times do not necessarily commute.

Of particular interest are operators that can be written as  $\exp \mathbf{H}(t)$ , where  $\mathbf{H}(t)$  is a “simple” operator; i.e., the dependence of  $\mathbf{H}(t)$  on  $t$  is simpler than the corresponding dependence of  $\exp \mathbf{H}(t)$ . We have already encountered such a situation in Example 2.1.9, where it was shown that the operation of rotation around the  $z$ -axis could be written as  $\exp \alpha \mathbf{T}$ , and the action of  $\mathbf{T}$  on  $(x, y)$  was a great deal simpler than the corresponding action of  $\exp \alpha \mathbf{T}$ .

Such a state of affairs is very common in physics. In fact, it can be shown that many operators of physical interest can be written as a product of simpler operators, each being of the form  $\exp \alpha \mathbf{T}$ . For example, we know from Euler’s theorem in mechanics that an arbitrary rotation in three dimensions can be written as a product of three simpler rotations, each being a rotation through a so-called *Euler angle* about an axis.

derivative of an operator

**2.2.1. Definition.** For the mapping  $\mathbf{H} : \mathbb{R} \rightarrow \mathcal{L}(\mathcal{V})$ , we define the *derivative* as

$$\frac{d\mathbf{H}}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{H}(t + \Delta t) - \mathbf{H}(t)}{\Delta t}.$$

This derivative also belongs to  $\mathcal{L}(\mathcal{V})$ .

As long as we keep track of the order, practically all the rules of differentiation apply to operators. For example,

$$\frac{d}{dt}(\mathbf{UT}) = \frac{d\mathbf{U}}{dt}\mathbf{T} + \mathbf{U}\frac{d\mathbf{T}}{dt}.$$

<sup>2</sup>Strictly speaking, the domain of  $\mathbf{H}$  must be an interval  $[a, b]$  of the real line, because  $\mathbf{H}$  may not be defined for all  $\mathbb{R}$ . However, for our purposes, such a fine distinction is not necessary.

We are not allowed to change the order of multiplication on the RHS, not even when both operators being multiplied are the same on the LHS. For instance, if we let  $\mathbf{U} = \mathbf{T} = \mathbf{H}$  in the preceding equation, we obtain

$$\frac{d}{dt}(\mathbf{H}^2) = \frac{d\mathbf{H}}{dt}\mathbf{H} + \mathbf{H}\frac{d\mathbf{H}}{dt}$$

This is *not*, in general, equal to  $2\mathbf{H}\frac{d\mathbf{H}}{dt}$ .

**2.2.2. Example.** Let us find the derivative of  $\exp(t\mathbf{H})$ , where  $\mathbf{H}$  is independent of  $t$ . Using Definition 2.2.1, we have

$$\frac{d}{dt}\exp(t\mathbf{H}) = \lim_{\Delta t \rightarrow 0} \frac{\exp[(t + \Delta t)\mathbf{H}] - \exp(t\mathbf{H})}{\Delta t}.$$

However, for infinitesimal  $\Delta t$  we have

$$\begin{aligned} \exp[(t + \Delta t)\mathbf{H}] - \exp(t\mathbf{H}) &= e^{t\mathbf{H}}e^{\Delta t\mathbf{H}} - e^{t\mathbf{H}} \\ &= e^{t\mathbf{H}}(\mathbf{1} + \mathbf{H}\Delta t) - e^{t\mathbf{H}} = e^{t\mathbf{H}}\mathbf{H}\Delta t. \end{aligned}$$

Therefore,

$$\frac{d}{dt}\exp(t\mathbf{H}) = \lim_{\Delta t \rightarrow 0} \frac{e^{t\mathbf{H}}\mathbf{H}\Delta t}{\Delta t} = e^{t\mathbf{H}}\mathbf{H}.$$

Since  $\mathbf{H}$  and  $e^{t\mathbf{H}}$  commute,<sup>3</sup> we also have

$$\frac{d}{dt}\exp(t\mathbf{H}) = \mathbf{H}e^{t\mathbf{H}}.$$

Note that in deriving the equation for the derivative of  $e^{t\mathbf{H}}$ , we have used the relation  $e^{t\mathbf{H}}e^{\Delta t\mathbf{H}} = e^{(t+\Delta t)\mathbf{H}}$ . This may seem trivial, but it will be shown later that in general,  $e^{\mathbf{S}+\mathbf{T}} \neq e^{\mathbf{S}}e^{\mathbf{T}}$ . ■

Now let us evaluate the derivative of a more general time-dependent operator,  $\exp[\mathbf{H}(t)]$ :

$$\frac{d}{dt}\exp[\mathbf{H}(t)] = \lim_{\Delta t \rightarrow 0} \frac{\exp[\mathbf{H}(t + \Delta t)] - \exp[\mathbf{H}(t)]}{\Delta t}.$$

If  $\mathbf{H}(t)$  possesses a derivative, we have, to the first order in  $\Delta t$ ,

$$\mathbf{H}(t + \Delta t) = \mathbf{H}(t) + \Delta t \frac{d}{dt}\mathbf{H},$$

and we can write  $\exp[\mathbf{H}(t + \Delta t)] = \exp[\mathbf{H}(t) + \Delta t d\mathbf{H}/dt]$ . It is very tempting to factor out the  $\exp[\mathbf{H}(t)]$  and expand the remaining part. However, as we will see presently, this is not possible in general. As preparation, consider the following example, which concerns the integration of an operator.

**2.2.3. Example.** The Schrödinger equation  $i \frac{\partial}{\partial t} |\psi(t)\rangle = \mathbf{H} |\psi(t)\rangle$  can be turned into an evolution operator  $|\psi(t)\rangle = \mathbf{U}(t) |\psi(0)\rangle$ , and substitute in the Schrödinger equation to obtain

$$i \frac{\partial}{\partial t} \mathbf{U}(t) |\psi(0)\rangle = \mathbf{H} \mathbf{U}(t) |\psi(0)\rangle .$$

Ignoring the arbitrary vector  $|\psi(0)\rangle$  results in a differential equation in  $\mathbf{U}(t)$ . For the purposes of this example, let us consider an operator differential equation of the form  $d\mathbf{U}/dt = \mathbf{H}\mathbf{U}(t)$ , where  $\mathbf{H}$  is not dependent on  $t$ . We can find a solution to such an equation by repeated differentiation followed by Taylor series expansion. Thus,

$$\begin{aligned} \frac{d^2 \mathbf{U}}{dt^2} &= \mathbf{H} \frac{d}{dt} \mathbf{U} = \mathbf{H} [\mathbf{H} \mathbf{U}(t)] = \mathbf{H}^2 \mathbf{U}(t), \\ \frac{d^3 \mathbf{U}}{dt^3} &= \frac{d}{dt} [\mathbf{H}^2 \mathbf{U}(t)] = \mathbf{H}^2 \frac{d}{dt} \mathbf{U} = \mathbf{H}^3 \mathbf{U}(t). \end{aligned}$$

In general  $d^n \mathbf{U}/dt^n = \mathbf{H}^n \mathbf{U}(t)$ . Assuming that  $\mathbf{U}(t)$  is well-defined at  $t = 0$ , the above relations say that all derivatives of  $\mathbf{U}(t)$  are also well-defined at  $t = 0$ . Therefore, we can expand  $\mathbf{U}(t)$  around  $t = 0$  to obtain

$$\begin{aligned} \mathbf{U}(t) &= \sum_{n=0}^{\infty} \frac{t^n}{n!} \left( \frac{d^n \mathbf{U}}{dt^n} \right)_{t=0} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{H}^n \mathbf{U}(0) \\ &= \left( \sum_{n=0}^{\infty} \frac{(t\mathbf{H})^n}{n!} \right) \mathbf{U}(0) = e^{t\mathbf{H}} \mathbf{U}(0). \end{aligned}$$

Let us see under what conditions we have  $\exp(\mathbf{S} + \mathbf{T}) = \exp(\mathbf{S}) \exp(\mathbf{T})$ . We consider only the case where the commutator of the two operators commutes with both of them:  $[\mathbf{T}, [\mathbf{S}, \mathbf{T}]] = \mathbf{0} = [\mathbf{S}, [\mathbf{S}, \mathbf{T}]]$ . Now consider the operator  $\mathbf{U}(t) = e^{t\mathbf{S}} e^{t\mathbf{T}} e^{-t(\mathbf{S}+\mathbf{T})}$  and differentiate it using the result of Example 2.2.2 and the product rule for differentiation:

$$\begin{aligned} \frac{d}{dt} \mathbf{U} &= \mathbf{S} e^{t\mathbf{S}} e^{t\mathbf{T}} e^{-t(\mathbf{S}+\mathbf{T})} + e^{t\mathbf{S}} \mathbf{T} e^{t\mathbf{T}} e^{-t(\mathbf{S}+\mathbf{T})} - e^{t\mathbf{S}} e^{t\mathbf{T}} (\mathbf{S} + \mathbf{T}) e^{-t(\mathbf{S}+\mathbf{T})} \\ &= \mathbf{S} e^{t\mathbf{S}} e^{t\mathbf{T}} e^{-t(\mathbf{S}+\mathbf{T})} - e^{t\mathbf{S}} e^{t\mathbf{T}} \mathbf{S} e^{-t(\mathbf{S}+\mathbf{T})}. \end{aligned} \tag{2.5}$$

The three factors of  $\mathbf{U}(t)$  are present in all terms; however, they are not always next to one another. We can switch the operators if we introduce a commutator. For instance,  $e^{t\mathbf{T}} \mathbf{S} = \mathbf{S} e^{t\mathbf{T}} + [e^{t\mathbf{T}}, \mathbf{S}]$ .

It is left as a problem for the reader to show that if  $[\mathbf{S}, \mathbf{T}]$  commutes with  $\mathbf{S}$  and  $\mathbf{T}$ , then  $[e^{t\mathbf{T}}, \mathbf{S}] = -t[\mathbf{S}, \mathbf{T}]e^{t\mathbf{T}}$ , and therefore,  $e^{t\mathbf{T}} \mathbf{S} = \mathbf{S} e^{t\mathbf{T}} - t[\mathbf{S}, \mathbf{T}]e^{t\mathbf{T}}$ . Substituting this in Equation (2.5) and noting that  $e^{t\mathbf{S}} \mathbf{S} = \mathbf{S} e^{t\mathbf{S}}$  yields  $d\mathbf{U}/dt = t[\mathbf{S}, \mathbf{T}]\mathbf{U}(t)$ . The solution to this equation is

$$\mathbf{U}(t) = \exp \left( \frac{t^2}{2} [\mathbf{S}, \mathbf{T}] \right) \Rightarrow e^{t\mathbf{S}} e^{t\mathbf{T}} e^{-t(\mathbf{S}+\mathbf{T})} = \exp \left( \frac{t^2}{2} [\mathbf{S}, \mathbf{T}] \right)$$

<sup>3</sup>This is a consequence of a more general result that if two operators commute, any pair of functions of those operators also commute (see Problem 2.14).

because  $\mathbf{U}(0) = \mathbf{1}$ . We thus have the following:

**2.2.4. Proposition.** *Let  $\mathbf{S}, \mathbf{T} \in \mathcal{L}(\mathcal{V})$ . If  $[\mathbf{S}, [\mathbf{S}, \mathbf{T}]] = \mathbf{0} = [\mathbf{T}, [\mathbf{S}, \mathbf{T}]]$ , then the Baker–Campbell–Hausdorff formula*

$$e^{t\mathbf{S}}e^{t\mathbf{T}}e^{-(t^2/2)[\mathbf{S}, \mathbf{T}]} = e^{t(\mathbf{S}+\mathbf{T})}. \tag{2.6}$$

In particular,  $e^{t\mathbf{S}}e^{t\mathbf{T}} = e^{t(\mathbf{S}+\mathbf{T})}$  if and only if  $[\mathbf{S}, \mathbf{T}] = \mathbf{0}$ .

If  $t = 1$ , Equation (2.6) reduces to

$$e^{\mathbf{S}}e^{\mathbf{T}}e^{-(1/2)[\mathbf{S}, \mathbf{T}]} = e^{\mathbf{S}+\mathbf{T}}. \tag{2.7}$$

Now assume that both  $\mathbf{H}(t)$  and its derivative commute with  $[\mathbf{H}, d\mathbf{H}/dt]$ . Letting  $\mathbf{S} = \mathbf{H}(t)$  and  $\mathbf{T} = \Delta t d\mathbf{H}/dt$  in (2.7), we obtain

$$\begin{aligned} e^{\mathbf{H}(t+\Delta t)} &= e^{\mathbf{H}(t)+\Delta t d\mathbf{H}/dt} \\ &= e^{\mathbf{H}(t)}e^{\Delta t(d\mathbf{H}/dt)}e^{-[\mathbf{H}(t), \Delta t d\mathbf{H}/dt]/2}. \end{aligned}$$

For infinitesimal  $\Delta t$ , this yields

$$\begin{aligned} e^{\mathbf{H}(t+\Delta t)} &= e^{\mathbf{H}(t)} \left(1 + \Delta t \frac{d\mathbf{H}}{dt}\right) \left(1 - \frac{1}{2}\Delta t \left[\mathbf{H}(t), \frac{d\mathbf{H}}{dt}\right]\right) \\ &= e^{\mathbf{H}(t)} \left\{1 + \Delta t \frac{d\mathbf{H}}{dt} - \frac{1}{2}\Delta t \left[\mathbf{H}(t), \frac{d\mathbf{H}}{dt}\right]\right\}, \end{aligned}$$

and we have

$$\frac{d}{dt}e^{\mathbf{H}(t)} = e^{\mathbf{H}}\frac{d\mathbf{H}}{dt} - \frac{1}{2}e^{\mathbf{H}}\left[\mathbf{H}, \frac{d\mathbf{H}}{dt}\right].$$

We can also write

$$\begin{aligned} e^{\mathbf{H}(t+\Delta t)} &= e^{[\mathbf{H}(t)+\Delta t d\mathbf{H}/dt]} = e^{[\Delta t d\mathbf{H}/dt+\mathbf{H}(t)]} \\ &= e^{[\Delta t d\mathbf{H}/dt]}e^{\mathbf{H}(t)}e^{-[\Delta t d\mathbf{H}/dt, \mathbf{H}(t)]/2}, \end{aligned}$$

which yields

$$\frac{d}{dt}e^{\mathbf{H}(t)} = \frac{d\mathbf{H}}{dt}e^{\mathbf{H}} + \frac{1}{2}e^{\mathbf{H}}\left[\mathbf{H}, \frac{d\mathbf{H}}{dt}\right].$$

Adding the above two expressions and dividing by 2 yields the following symmetric expression for the derivative:

$$\frac{d}{dt}e^{\mathbf{H}(t)} = \frac{1}{2}\left(\frac{d\mathbf{H}}{dt}e^{\mathbf{H}} + e^{\mathbf{H}}\frac{d\mathbf{H}}{dt}\right) = \frac{1}{2}\left\{\frac{d\mathbf{H}}{dt}, e^{\mathbf{H}}\right\},$$

anticommutator where  $\{\mathbf{S}, \mathbf{T}\} \equiv \mathbf{ST} + \mathbf{TS}$  is called the **anticommutator** of the operators  $\mathbf{S}$  and  $\mathbf{T}$ . We, therefore, have the following proposition.

**2.2.5. Proposition.** Let  $\mathbf{H} : \mathbb{R} \rightarrow \mathcal{L}(\mathcal{V})$  and assume that  $\mathbf{H}$  and its derivative commute with  $[\mathbf{H}, d\mathbf{H}/dt]$ . Then

$$\frac{d}{dt} e^{\mathbf{H}(t)} = \frac{1}{2} \left\{ \frac{d\mathbf{H}}{dt}, e^{\mathbf{H}} \right\}.$$

In particular, if  $[\mathbf{H}, d\mathbf{H}/dt] = 0$ , then

$$\frac{d}{dt} e^{\mathbf{H}(t)} = \frac{d\mathbf{H}}{dt} e^{\mathbf{H}} = e^{\mathbf{H}} \frac{d\mathbf{H}}{dt}.$$

A frequently encountered operator is  $\mathbf{F}(t) = e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}}$ , where  $\mathbf{A}$  and  $\mathbf{B}$  are  $t$ -independent. It is straightforward to show that

$$\frac{d\mathbf{F}}{dt} = [\mathbf{A}, \mathbf{F}(t)] \quad \text{and} \quad \frac{d}{dt} [\mathbf{A}, \mathbf{F}(t)] = \left[ \mathbf{A}, \frac{d\mathbf{F}}{dt} \right].$$

Using these results, we can write

$$\frac{d^2\mathbf{F}}{dt^2} = \frac{d}{dt} [\mathbf{A}, \mathbf{F}(t)] = [\mathbf{A}, [\mathbf{A}, \mathbf{F}(t)]] \equiv \mathbf{A}^2[\mathbf{F}(t)],$$

and in general,  $d^n\mathbf{F}/dt^n = \mathbf{A}^n[\mathbf{F}(t)]$ , where  $\mathbf{A}^n[\mathbf{F}(t)]$  is defined inductively as  $\mathbf{A}^n[\mathbf{F}(t)] = [\mathbf{A}, \mathbf{A}^{n-1}[\mathbf{F}(t)]]$ , with  $\mathbf{A}^0[\mathbf{F}(t)] \equiv \mathbf{F}(t)$ . For example,

$$\mathbf{A}^3[\mathbf{F}(t)] = [\mathbf{A}, \mathbf{A}^2[\mathbf{F}(t)]] = [\mathbf{A}, [\mathbf{A}, \mathbf{A}[\mathbf{F}(t)]]] = [\mathbf{A}, [\mathbf{A}, [\mathbf{A}, \mathbf{F}(t)]]].$$

Evaluating  $\mathbf{F}(t)$  and all its derivatives at  $t = 0$  and substituting in the Taylor expansion about  $t = 0$ , we get

$$\mathbf{F}(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left. \frac{d^n\mathbf{F}}{dt^n} \right|_{t=0} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{A}^n[\mathbf{F}(0)] = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{A}^n[\mathbf{B}].$$

That is,

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{A}^n[\mathbf{B}] \equiv \mathbf{B} + t[\mathbf{A}, \mathbf{B}] + \frac{t^2}{2!} [\mathbf{A}, [\mathbf{A}, \mathbf{B}]] + \dots$$

Sometimes this is written symbolically as

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \left( \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{A}^n \right) [\mathbf{B}] \equiv e^{t\mathbf{A}}[\mathbf{B}],$$

where the RHS is merely an abbreviation of the infinite sum in the middle.

For  $t = 1$  we obtain a widely used formula:

$$e^{\mathbf{A}}\mathbf{B}e^{-\mathbf{A}} = e^{\mathbf{A}}[\mathbf{B}] = \left( \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{A}^n \right) [\mathbf{B}] \equiv \mathbf{B} + [\mathbf{A}, \mathbf{B}] + \frac{1}{2!} [\mathbf{A}, [\mathbf{A}, \mathbf{B}]] + \dots$$



If  $\mathbf{A}$  commutes with  $[\mathbf{A}, \mathbf{B}]$ , then the infinite series truncates at the second term, and we have

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \mathbf{B} + t[\mathbf{A}, \mathbf{B}].$$

For instance, if  $\mathbf{A}$  and  $\mathbf{B}$  are replaced by  $\mathbf{D}$  and  $\mathbf{T}$  of Example 1.3.4, we get (see Problem 2.3)

$$e^{t\mathbf{D}}\mathbf{T}e^{-t\mathbf{D}} = \mathbf{T} + t[\mathbf{D}, \mathbf{T}] = \mathbf{T} + t\mathbf{1}.$$

generator of translation

The RHS shows that the operator  $\mathbf{T}$  has been *translated* by an amount  $t$  (more precisely, by  $t$  times the unit operator). We therefore call  $\exp(t\mathbf{D})$  the translation operator of  $\mathbf{T}$  by  $t$ , and we call  $\mathbf{D}$  the *generator* of translation. With a little modification  $\mathbf{T}$  and  $\mathbf{D}$  become, respectively, the position and momentum operators in quantum mechanics. Thus,

momentum as generator of translation

**2.2.6. Box.** *Momentum is the generator of translation in quantum mechanics.*

But more of this later!

## 2.3 Conjugation of Operators

We have discussed the notion of the dual of a vector in conjunction with inner products. We now incorporate linear operators into this notion. Let  $|b\rangle, |c\rangle \in \mathcal{V}$  and assume that  $|c\rangle = \mathbf{T}|b\rangle$ . We know that there are linear functionals in the dual space  $\mathcal{V}^*$  that are associated with  $(|b\rangle)^\dagger = \langle b|$  and  $(|c\rangle)^\dagger = \langle c|$ . Is there a linear operator belonging to  $\mathcal{L}(\mathcal{V}^*)$  that somehow corresponds to  $\mathbf{T}$ ? In other words, can we find a linear operator that relates  $\langle b|$  and  $\langle c|$  just as  $\mathbf{T}$  relates  $|b\rangle$  and  $|c\rangle$ ? The answer comes in the following definition.

adjoint of an operator

**2.3.1. Definition.** Let  $\mathbf{T} \in \mathcal{L}(\mathcal{V})$  and  $|a\rangle, |b\rangle \in \mathcal{V}$ . The *adjoint*, or *hermitian conjugate*, of  $\mathbf{T}$  is denoted by  $\mathbf{T}^\dagger$  and defined by

$$\langle a|\mathbf{T}|b\rangle^* = \langle b|\mathbf{T}^\dagger|a\rangle. \tag{2.8}$$

The LHS of Equation (2.8) can be written as  $\langle a|c\rangle^*$  or  $\langle c|a\rangle$ , in which case we can identify

$$\langle c| = \langle b|\mathbf{T}^\dagger \Rightarrow (\mathbf{T}|b\rangle)^\dagger = \langle b|\mathbf{T}^\dagger. \tag{2.9}$$

This equation is sometimes used as the definition of the hermitian conjugate. From Equation (2.8), the reader may easily verify that  $\mathbf{1}^\dagger = \mathbf{1}$ . Thus, using the unit operator for  $\mathbf{T}$ , (2.9) justifies Equation (1.12).

Some of the properties of conjugation are listed in the following theorem, whose proof is left as an exercise.

**2.3.2. Theorem.** Let  $\mathbf{U}, \mathbf{T} \in \mathcal{L}(\mathcal{V})$  and  $\alpha \in \mathbb{C}$ . Then

1.  $(\mathbf{U} + \mathbf{T})^\dagger = \mathbf{U}^\dagger + \mathbf{T}^\dagger$ .
2.  $(\mathbf{UT})^\dagger = \mathbf{T}^\dagger \mathbf{U}^\dagger$ .
3.  $(\alpha \mathbf{T})^\dagger = \alpha^* \mathbf{T}^\dagger$ .
4.  $((\mathbf{T}^\dagger)^\dagger) = \mathbf{T}$ .

The last identity holds for finite-dimensional vector spaces; it does not apply to infinite-dimensional vector spaces in general.

In previous examples dealing with linear operators  $\mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , an element of  $\mathbb{R}^n$  was denoted by a row vector, such as  $(x, y)$  for  $\mathbb{R}^2$  and  $(x, y, z)$  for  $\mathbb{R}^3$ . There was no confusion, because we were operating only in  $\mathcal{V}$ . However, since elements of both  $\mathcal{V}$  and  $\mathcal{V}^*$  are required when discussing  $\mathbf{T}$ ,  $\mathbf{T}^*$ , and  $\mathbf{T}^\dagger$ , it is helpful to make a distinction between them. We therefore resort to the convention introduced in Example 1.2.3 by which

**2.3.3. Box.** *Kets are represented as column vectors and bras as row vectors.*

**2.3.4. Example.** Let us find the hermitian conjugate of the operator  $\mathbf{T} : \mathbb{C}^3 \rightarrow \mathbb{C}^3$  given by

$$\mathbf{T} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 - i\alpha_2 + \alpha_3 \\ i\alpha_1 - \alpha_3 \\ \alpha_1 - \alpha_2 + i\alpha_3 \end{pmatrix}.$$

Introduce

$$|a\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \quad \text{and} \quad |b\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

with dual vectors  $\langle a| = (\alpha_1^* \ \alpha_2^* \ \alpha_3^*)$  and  $\langle b| = (\beta_1^* \ \beta_2^* \ \beta_3^*)$ , respectively. We use Equation (2.8) to find  $\mathbf{T}^\dagger$ :

$$\begin{aligned} \langle b| \mathbf{T}^\dagger |a\rangle &= \langle a| \mathbf{T} |b\rangle^* = \left[ (\alpha_1^* \ \alpha_2^* \ \alpha_3^*) \mathbf{T} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \right]^* \\ &= \left[ (\alpha_1^* \ \alpha_2^* \ \alpha_3^*) \begin{pmatrix} \beta_1 - i\beta_2 + \beta_3 \\ i\beta_1 - \beta_3 \\ \beta_1 - \beta_2 + i\beta_3 \end{pmatrix} \right]^* \\ &= [\alpha_1^* \beta_1 - i\alpha_1^* \beta_2 + \alpha_1^* \beta_3 + i\alpha_2^* \beta_1 - \alpha_2^* \beta_3 + \alpha_3^* \beta_1 - \alpha_3^* \beta_2 + i\alpha_3^* \beta_3]^* \\ &= \alpha_1 \beta_1^* + i\alpha_1 \beta_2^* + \alpha_1 \beta_3^* - i\alpha_2 \beta_1^* - \alpha_2 \beta_3^* + \alpha_3 \beta_1^* - \alpha_3 \beta_2^* - i\alpha_3 \beta_3^* \\ &= \beta_1^* (\alpha_1 - i\alpha_2 + \alpha_3) + \beta_2^* (i\alpha_1 - \alpha_3) + \beta_3^* (\alpha_1 - \alpha_2 - i\alpha_3) \\ &= (\beta_1^* \ \beta_2^* \ \beta_3^*) \begin{pmatrix} \alpha_1 - i\alpha_2 + \alpha_3 \\ i\alpha_1 - \alpha_3 \\ \alpha_1 - \alpha_2 - i\alpha_3 \end{pmatrix}. \end{aligned}$$

Therefore, we obtain

$$\mathbf{T}^\dagger \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 - i\alpha_2 + \alpha_3 \\ i\alpha_1 - \alpha_3 \\ \alpha_1 - \alpha_2 - i\alpha_3 \end{pmatrix}. \quad \blacksquare$$

## 2.4 Hermitian and Unitary Operators

The process of conjugation of linear operators looks much like conjugation of complex numbers. Equation (2.8) alludes to this fact, and Theorem 2.3.2 provides further evidence. It is therefore natural to look for operators that are counterparts of real numbers. One can define complex conjugation for operators and thereby construct real operators. However, these real operators will not be interesting because—as it turns out—they completely ignore the complex character of the vector space. The following alternative definition makes use of *hermitian* conjugation, and the result will have much wider application than is allowed by a mere *complex* conjugation.

hermitian and  
anti-hermitian  
operators

**2.4.1. Definition.** A linear operator  $\mathbf{H} \in \mathcal{L}(\mathcal{V})$  is called *hermitian*, or *self-adjoint*, if  $\mathbf{H}^\dagger = \mathbf{H}$ . Similarly,  $\mathbf{A} \in \mathcal{L}(\mathcal{V})$  is called *anti-hermitian* if  $\mathbf{A}^\dagger = -\mathbf{A}$ .

**Charles Hermite** (1822–1901), one of the most eminent French mathematicians of the nineteenth century, was particularly distinguished for the clean elegance and high artistic quality of his work. As a student, he courted disaster by neglecting his routine assigned work to study the classic masters of mathematics; and though he nearly failed his examinations, he became a first-rate creative mathematician while still in his early twenties. In 1870 he was appointed to a professorship at the Sorbonne, where he trained a whole generation of well-known French mathematicians, including Picard, Borel, and Poincaré.



The character of his mind is suggested by a remark of Poincaré: “Talk with M. Hermite. He never evokes a concrete image, yet you soon perceive that the most abstract entities are to him like living creatures.” He disliked geometry, but was strongly attracted to number theory and analysis, and his favorite subject was elliptic functions, where these two fields touch in many remarkable ways. Earlier in the century the Norwegian genius Abel had proved that the general equation of the fifth degree cannot be solved by functions involving only rational operations and root extractions. One of Hermite’s most surprising achievements (in 1858) was to show that this equation can be solved by elliptic functions.

His 1873 proof of the transcendence of  $e$  was another high point of his career.<sup>4</sup> If he had been willing to dig even deeper into this vein, he could probably have disposed of  $\pi$  as

<sup>4</sup>Transcendental numbers are those that are not roots of polynomials with integer coefficients.

well, but apparently he had had enough of a good thing. As he wrote to a friend, "I shall risk nothing on an attempt to prove the transcendence of the number  $\pi$ . If others undertake this enterprise, no one will be happier than I at their success, but believe me, my dear friend, it will not fail to cost them some efforts." As it turned out, Lindemann's proof nine years later rested on extending Hermite's method.

Several of his purely mathematical discoveries had unexpected applications many years later to mathematical physics. For example, the Hermitian forms and matrices that he invented in connection with certain problems of number theory turned out to be crucial for Heisenberg's 1925 formulation of quantum mechanics, and Hermite polynomials (see Chapter 7) are useful in solving Schrödinger's wave equation.

The following observations strengthen the above conjecture that conjugation of complex numbers and hermitian conjugation of operators are somehow related.

expectation value **2.4.2. Definition.** *The expectation value  $\langle T \rangle_a$  of an operator  $T$  in the "state"  $|a\rangle$  is a complex number defined by  $\langle T \rangle_a = \langle a | T | a \rangle$ .*

The complex conjugate of the expectation value is<sup>5</sup>

$$\langle T \rangle^* = \langle a | T | a \rangle^* = \langle a | T^\dagger | a \rangle.$$

In words,  $T^\dagger$ , the *hermitian conjugate* of  $T$ , has an expectation value that is the *complex conjugate* of the latter's expectation value. In particular, if  $T$  is hermitian—its equal to its hermitian conjugate—its expectation value will be real.

What is the analogue of the known fact that a complex number is the sum of a real number and a pure imaginary one? The decomposition

$$T = \frac{1}{2}(T + T^\dagger) + \frac{1}{2}(T - T^\dagger) \equiv H + A$$

shows that *any operator can be written as a sum of a hermitian operator  $H = \frac{1}{2}(T + T^\dagger)$  and an anti-hermitian operator  $A = \frac{1}{2}(T - T^\dagger)$ .*

We can go even further, because any anti-hermitian operator  $A$  can be written as  $A = i(-iA)$  in which  $-iA$  is hermitian:  $(-iA)^\dagger = (-i)^* A^\dagger = i(-A) = -iA$ . Denoting  $-iA$  by  $H'$ , we write  $T = H + iH'$ , where both  $H$  and  $H'$  are *hermitian*. This is the analogue of the decomposition  $z = x + iy$  in which both  $x$  and  $y$  are *real*.

Clearly, we should expect some departures from a perfect correspondence. This is due to a lack of commutativity among operators. For instance, although the product of two real numbers is real, the product of two hermitian operators is not, in general, hermitian:

$$(TU)^\dagger = U^\dagger T^\dagger = UT \neq TU.$$

<sup>5</sup>When no risk of confusion exists, it is common to drop the subscript "a" and write  $\langle T \rangle$  for the expectation value of  $T$ .

We have seen the relation between expectation values and conjugation properties of operators. The following theorem completely characterizes hermitian operators in terms of their expectation values:

**2.4.3. Theorem.** *A linear transformation  $\mathbf{H}$  on a complex inner product space is hermitian if and only if  $\langle a | \mathbf{H} | a \rangle$  is real for all  $|a\rangle$ .*

*Proof.* We have already pointed out that a hermitian operator has real expectation values. Conversely, assume that  $\langle a | \mathbf{H} | a \rangle$  is real for all  $|a\rangle$ . Then

$$\langle a | \mathbf{H} | a \rangle = \langle a | \mathbf{H} | a \rangle^* = \langle a | \mathbf{H}^\dagger | a \rangle \quad \Leftrightarrow \quad \langle a | \mathbf{H} - \mathbf{H}^\dagger | a \rangle = 0 \quad \forall |a\rangle.$$

By Theorem 2.1.4 we must have  $\mathbf{H} - \mathbf{H}^\dagger = \mathbf{0}$ . □

**2.4.4. Example.** In this example, we illustrate the result of the above theorem with  $2 \times 2$  matrices. The matrix  $\mathbf{H} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$  is hermitian<sup>6</sup> and acts on  $\mathbb{C}^2$ . Let us take an arbitrary vector  $|a\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$  and evaluate  $\langle a | \mathbf{H} | a \rangle$ . We have

$$\mathbf{H} |a\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} -i\alpha_2 \\ i\alpha_1 \end{pmatrix}.$$

Therefore,

$$\begin{aligned} \langle a | \mathbf{H} | a \rangle &= (\alpha_1^* \quad \alpha_2^*) \begin{pmatrix} -i\alpha_2 \\ i\alpha_1 \end{pmatrix} = -i\alpha_1^* \alpha_2 + i\alpha_2^* \alpha_1 \\ &= i\alpha_2^* \alpha_1 + (i\alpha_2^* \alpha_1)^* = 2\operatorname{Re}(i\alpha_2^* \alpha_1), \end{aligned}$$

and  $\langle a | \mathbf{H} | a \rangle$  is real.

For the most general  $2 \times 2$  hermitian matrix  $\mathbf{H} = \begin{pmatrix} \alpha & \beta \\ \beta^* & \gamma \end{pmatrix}$ , where  $\alpha$  and  $\gamma$  are real, we have

$$\mathbf{H} |a\rangle = \begin{pmatrix} \alpha & \beta \\ \beta^* & \gamma \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \alpha\alpha_1 + \beta\alpha_2 \\ \beta^*\alpha_1 + \gamma\alpha_2 \end{pmatrix}$$

and

$$\begin{aligned} \langle a | \mathbf{H} | a \rangle &= (\alpha_1^* \quad \alpha_2^*) \begin{pmatrix} \alpha\alpha_1 + \beta\alpha_2 \\ \beta^*\alpha_1 + \gamma\alpha_2 \end{pmatrix} = \alpha_1^* (\alpha\alpha_1 + \beta\alpha_2) + \alpha_2^* (\beta^*\alpha_1 + \gamma\alpha_2) \\ &= \alpha|\alpha_1|^2 + \alpha_1^* \beta\alpha_2 + \alpha_2^* \beta^* \alpha_1 + \gamma|\alpha_2|^2 \\ &= \alpha|\alpha_1|^2 + \gamma|\alpha_2|^2 + 2\operatorname{Re}(\alpha_1^* \beta\alpha_2). \end{aligned}$$

Again  $\langle a | \mathbf{H} | a \rangle$  is real. ■

**2.4.5. Definition.** *An operator  $\mathbf{A}$  on an inner product space is called **positive** (written  $\mathbf{A} \geq 0$ ) if  $\mathbf{A}$  is hermitian and  $\langle a | \mathbf{A} | a \rangle \geq 0$  for all  $|a\rangle$ .*

<sup>6</sup>We assume that the reader has a casual familiarity with hermitian *matrices*. Think of an  $n \times n$  matrix as a linear operator that acts on column vectors whose elements are components of vectors defined in the standard basis of  $\mathbb{C}^n$  or  $\mathbb{R}^n$ . A hermitian matrix then becomes a hermitian operator.

**2.4.6. Example.** An example of a positive operator is the square of a hermitian operator.<sup>7</sup> We note that for any hermitian operator  $H$  and any vector  $|a\rangle$ , we have  $\langle a|H^2|a\rangle = \langle a|H^\dagger H|a\rangle = \langle Ha|Ha\rangle \geq 0$  because of the positive definiteness of the inner product. ■

An operator  $T$  satisfying the extra condition that  $\langle a|T|a\rangle = 0$  implies  $|a\rangle = 0$  is called **positive definite**. From the discussion of the example above, we conclude that the square of an *invertible* hermitian operator is positive definite.

positive definite operators

The reader may be familiar with two- and three-dimensional rigid rotations and the fact that they preserve distances and the scalar product. Can this be generalized to complex inner product spaces? Let  $|a\rangle, |b\rangle \in \mathcal{V}$ , and let  $U$  be an operator on  $\mathcal{V}$  that preserves the scalar product; that is, given  $|b'\rangle = U|b\rangle$  and  $|a'\rangle = U|a\rangle$ , then  $\langle a'|b'\rangle = \langle a|b\rangle$ . This yields

$$\langle a'|b'\rangle = \langle (U^\dagger|a\rangle)(U|b\rangle) = \langle a|U^\dagger U|b\rangle = \langle a|b\rangle = \langle a|\mathbf{1}|b\rangle.$$

Since this is true for arbitrary  $|a\rangle$  and  $|b\rangle$ , we obtain  $U^\dagger U = \mathbf{1}$ . In the next chapter, when we introduce the concept of the determinant of operators, we shall see that this relation implies that  $U$  and  $U^\dagger$  are both invertible,<sup>8</sup> with each one being the inverse of the other.

**2.4.7. Definition.** Let  $\mathcal{V}$  be a finite-dimensional inner product space. An operator  $U$  is called a **unitary operator** if  $U^\dagger = U^{-1}$ . Unitary operators preserve the inner product of  $\mathcal{V}$ .

unitary operators

**2.4.8. Example.** The linear transformation  $T: \mathbb{C}^3 \rightarrow \mathbb{C}^3$  given by

$$T \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} (\alpha_1 - i\alpha_2)/\sqrt{2} \\ (\alpha_1 + i\alpha_2 - 2\alpha_3)/\sqrt{6} \\ (\alpha_1 - \alpha_2 + \alpha_3 + i(\alpha_1 + \alpha_2 + \alpha_3))/\sqrt{6} \end{pmatrix}$$

is unitary. In fact, let

$$|a\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \quad \text{and} \quad |b\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

with dual vectors  $\langle a| = (\alpha_1^* \ \alpha_2^* \ \alpha_3^*)$  and  $\langle b| = (\beta_1^* \ \beta_2^* \ \beta_3^*)$ , respectively. We use Equation (2.8) and the procedure of Example 2.3.4 to find  $T^\dagger$ . The result is

$$T^\dagger \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \frac{\alpha_1}{\sqrt{2}} + \frac{\alpha_2}{\sqrt{6}} + \frac{\alpha_3(1-i)}{\sqrt{6}} \\ \frac{i\alpha_1}{\sqrt{2}} - \frac{i\alpha_2}{\sqrt{6}} - \frac{\alpha_3(1+i)}{\sqrt{6}} \\ -\frac{2\alpha_2}{\sqrt{6}} + \frac{\alpha_3(1-i)}{\sqrt{6}} \end{pmatrix},$$

<sup>7</sup>This is further evidence that hermitian operators are analogues of real numbers: The square of any real number is positive.

<sup>8</sup>This implication holds only for finite-dimensional vector spaces.

and we can verify that

$$\mathbf{T}^\dagger \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}.$$

Thus  $\mathbf{T}\mathbf{T}^\dagger = \mathbf{1}$ . Similarly, we can show that  $\mathbf{T}^\dagger\mathbf{T} = \mathbf{1}$  and therefore that  $\mathbf{T}$  is unitary. ■

## 2.5 Projection Operators

We have already considered subspaces briefly. The significance of subspaces is that physics frequently takes place not inside the whole vector space, but in one of its subspaces. For instance, although motion generally takes place in a three-dimensional space, it may restrict itself to a plane either because of constraints or due to the nature of the force responsible for the motion. An example is planetary motion, which is confined to a plane because the force of gravity is central. Furthermore, the example of projectile motion teaches us that it is very convenient to “project” the motion onto the horizontal and vertical axes and to study these projections separately. It is, therefore, appropriate to ask how we can go from a full space to one of its subspaces in the context of linear operators. Let us first consider a simple example. A point in the plane is designated by the coordinates  $(x, y)$ . A subspace of the plane is the  $x$ -axis. Is there a *linear* operator,<sup>9</sup> say  $\mathbf{P}_x$ , that acts on such a point and somehow sends it into that subspace? Of course, there are many operators from  $\mathbb{R}^2$  to  $\mathbb{R}$ . However, we are looking for a specific one. We want  $\mathbf{P}_x$  to *project* the point onto the  $x$ -axis. Such an operator has to act on  $(x, y)$  and produce  $(x, 0)$ :  $\mathbf{P}_x(x, y) = (x, 0)$ . Note that if the point already lies on the  $x$ -axis,  $\mathbf{P}_x$  does not change it. In particular, if we apply  $\mathbf{P}_x$  twice, we get the same result as if we apply it only once. And this is true for any point in the plane. Therefore, our operator must have the property  $\mathbf{P}_x^2 = \mathbf{P}_x$ . We can generalize the above discussion in the following definition.<sup>10</sup>

projection operators **2.5.1. Definition.** A hermitian operator  $\mathbf{P} \in \mathcal{L}(\mathcal{V})$  is called a *projection operator* if  $\mathbf{P}^2 = \mathbf{P}$ .

From this definition it immediately follows that the only projection operator with an inverse is the identity operator. (Show this!)

Consider two projection operators  $\mathbf{P}_1$  and  $\mathbf{P}_2$ . We want to investigate conditions under which  $\mathbf{P}_1 + \mathbf{P}_2$  becomes a projection operator. By definition,  $\mathbf{P}_1 + \mathbf{P}_2 = (\mathbf{P}_1 + \mathbf{P}_2)^2 = \mathbf{P}_1^2 + \mathbf{P}_1\mathbf{P}_2 + \mathbf{P}_2\mathbf{P}_1 + \mathbf{P}_2^2$ . So  $\mathbf{P}_1 + \mathbf{P}_2$  is a projection operator if and only if

$$\mathbf{P}_1\mathbf{P}_2 + \mathbf{P}_2\mathbf{P}_1 = 0. \quad (2.10)$$

<sup>9</sup>We want this operator to preserve the vector-space structure of the plane and the axis.

<sup>10</sup>It is sometimes useful to relax the condition of hermiticity. However, in this part of the book, we demand that  $\mathbf{P}$  be hermitian.

Multiply this on the left by  $\mathbf{P}_1$  to get

$$\mathbf{P}_1^2 \mathbf{P}_2 + \mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_1 = 0 \Rightarrow \mathbf{P}_1 \mathbf{P}_2 + \mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_1 = 0.$$

Now multiply the same equation on the right by  $\mathbf{P}_1$  to get

$$\mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_1 + \mathbf{P}_2 \mathbf{P}_1^2 = 0 \Rightarrow \mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_1 + \mathbf{P}_2 \mathbf{P}_1 = 0.$$

These last two equations yield

$$\mathbf{P}_1 \mathbf{P}_2 - \mathbf{P}_2 \mathbf{P}_1 = 0. \tag{2.11}$$

The solution to Equations (2.10) and (2.11) is  $\mathbf{P}_1 \mathbf{P}_2 = \mathbf{P}_2 \mathbf{P}_1 = 0$ . We therefore have the following result.

**2.5.2. Proposition.** *Let  $\mathbf{P}_1, \mathbf{P}_2 \in \mathcal{L}(\mathcal{V})$  be projection operators. Then  $\mathbf{P}_1 + \mathbf{P}_2$  is a projection operator if and only if  $\mathbf{P}_1 \mathbf{P}_2 = \mathbf{P}_2 \mathbf{P}_1 = 0$ . Projection operators satisfying this condition are called **orthogonal projection operators**.*

orthogonal projection operators

More generally, if there is a set  $\{\mathbf{P}_i\}_{i=1}^m$  of projection operators satisfying

$$\mathbf{P}_i \mathbf{P}_j = \begin{cases} \mathbf{P}_i & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

then  $\mathbf{P} = \sum_{i=1}^m \mathbf{P}_i$  is also a projection operator. Given a normal vector  $|e\rangle$ , one can show easily that  $\mathbf{P} = |e\rangle \langle e|$  is a projection operator:

- **P is hermitian:**  $\mathbf{P}^\dagger = (|e\rangle \langle e|)^\dagger = (\langle e|)^\dagger (|e\rangle)^\dagger = |e\rangle \langle e|.$
- **P equals its square:**  $\mathbf{P}^2 = (|e\rangle \langle e|)(|e\rangle \langle e|) = |e\rangle \underbrace{\langle e|e\rangle}_{=1} \langle e| = |e\rangle \langle e|.$

In fact, we can take an orthonormal basis  $B = \{|e_i\rangle\}_{i=1}^N$  and construct a set of projection operators  $\{\mathbf{P}_i = |e_i\rangle \langle e_i|\}_{i=1}^N$ . The operators  $\mathbf{P}_i$  are mutually orthogonal. Thus, their sum  $\sum_{i=1}^N \mathbf{P}_i$  is also a projection operator.

**2.5.3. Proposition.** *Let  $B = \{|e_i\rangle\}_{i=1}^N$  be an orthonormal basis for  $\mathcal{V}_N$ . Then the set  $\{\mathbf{P}_i = |e_i\rangle \langle e_i|\}_{i=1}^N$  consists of mutually orthogonal projection operators, and  $\sum_{i=1}^N \mathbf{P}_i = \sum_{i=1}^N |e_i\rangle \langle e_i| = \mathbf{1}$ . This relation is called the **completeness relation**.*

completeness relation

*Proof.* The mutual orthogonality of the  $\mathbf{P}_i$  is an immediate consequence of the orthonormality of the  $|e_i\rangle$ . To show the second part, consider an arbitrary vector  $|a\rangle$ , written in terms of  $|e_i\rangle$ :  $|a\rangle = \sum_{j=1}^N \alpha_j |e_j\rangle$ . Apply  $\mathbf{P}_i$  to both sides to obtain

$$\mathbf{P}_i |a\rangle = \mathbf{P}_i \left( \sum_{j=1}^N \alpha_j |e_j\rangle \right) = \sum_{j=1}^N \alpha_j \mathbf{P}_i |e_j\rangle = \sum_{j=1}^N \alpha_j |e_i\rangle \underbrace{\langle e_i | e_j \rangle}_{\delta_{ij}} = \alpha_i |e_i\rangle.$$



Therefore, we have

$$\mathbf{1}|a\rangle = \sum_{i=1}^N \alpha_i |e_i\rangle = \sum_{i=1}^N \mathbf{P}_i |a\rangle = \left( \sum_{i=1}^N \mathbf{P}_i \right) |a\rangle.$$

Since this holds for an arbitrary  $|a\rangle$ , the two operators must be equal.  $\square$

If we choose only the first  $m < N$  vectors instead of the entire basis, then the projection operator  $\mathbf{P}^{(m)} \equiv \sum_{i=1}^m |e_i\rangle \langle e_i|$  projects arbitrary vectors into the subspace spanned by the first  $m$  basis vectors  $\{|e_i\rangle\}_{i=1}^m$ . In other words, when  $\mathbf{P}^{(m)}$  acts on any vector  $|a\rangle \in \mathcal{V}$ , the result will be a linear combination of only the first  $m$  vectors. The simple proof of this fact is left as an exercise. These points are illustrated in the following example.

**2.5.4. Example.** Consider three orthonormal vectors  $\{|e_i\rangle\}_{i=1}^3 \in \mathbb{R}^3$  given by

$$|e_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad |e_2\rangle = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad |e_3\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

The projection operators associated with each of these can be obtained by noting that  $\langle e_i|$  is a row vector. Therefore,

$$\mathbf{P}_1 = |e_1\rangle \langle e_1| = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Similarly,

$$\mathbf{P}_2 = \frac{1}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{pmatrix}$$

and

$$\mathbf{P}_3 = \frac{1}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Note that  $\mathbf{P}_i$  projects onto the line along  $|e_i\rangle$ . This can be tested by letting  $\mathbf{P}_i$  act on an arbitrary vector and showing that the resulting vector is perpendicular to the other two vectors. For example, let  $\mathbf{P}_2$  act on an arbitrary column vector:

$$|a\rangle \equiv \mathbf{P}_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{6} \begin{pmatrix} x - y + 2z \\ -x + y - 2z \\ 2x - 2y + 4z \end{pmatrix}.$$

We verify that  $|a\rangle$  is perpendicular to both  $|e_1\rangle$  and  $|e_3\rangle$ :

$$\langle e_1|a\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \frac{1}{6} \begin{pmatrix} x - y + 2z \\ -x + y - 2z \\ 2x - 2y + 4z \end{pmatrix} = 0.$$

Similarly,  $\langle e_3 | a \rangle = 0$ . So indeed,  $|a\rangle$  is along  $|e_2\rangle$ .

We can find the operator that projects onto the plane formed by  $|e_1\rangle$  and  $|e_2\rangle$ . This is

$$P_1 + P_2 = \frac{1}{3} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}.$$

When this operator acts on an arbitrary column vector, it produces a vector lying in the plane of  $|e_1\rangle$  and  $|e_2\rangle$ , or perpendicular to  $|e_3\rangle$ :

$$|b\rangle \equiv (P_1 + P_2) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2x + y + z \\ x + 2y - z \\ x - y + 2z \end{pmatrix}.$$

It is easy to show that  $\langle e_3 | b \rangle = 0$ . The operators that project onto the other two planes are obtained similarly. Finally, we verify easily that

$$P_1 + P_2 + P_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1. \quad \blacksquare$$

## 2.6 Operators in Numerical Analysis

In numerical calculations, limiting operations involving infinities and zeros are replaced with finite values. The most natural setting for the discussion of such operations is the ideas developed in this chapter. In this section, we shall assume that all operators are invertible, and (rather sloppily) manipulate them with no mathematical justification.

### 2.6.1 Finite-Difference Operators

In all numerical manipulations, a function is considered as a table with two columns. The first column lists the (discrete) values of the independent variable  $x_i$ , and the second column lists the value of the function  $f$  at  $x_i$ . We often write  $f_i$  for  $f(x_i)$ .

forward, backward,  
and central difference  
operators

Three operators that are in use in numerical analysis are the **forward difference operator**  $\Delta$ , the **backward difference operator**  $\nabla$  (not to be confused with the gradient), and the **central difference operator**  $\delta$ . These are defined as follows:

$$\Delta f_i \equiv f_{i+1} - f_i, \quad \nabla f_i \equiv f_i - f_{i-1}, \quad \delta f_i \equiv f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}. \quad (2.12)$$

The last equation has only theoretical significance, because a half-step is not used in the tabulation of functions or in computer calculations. Typically, the data are equally spaced, so  $x_{i+1} - x_i = h$  is the same for all  $i$ . Then  $f_{i\pm 1} = f(x_i \pm h)$ , and we define  $f_{i\pm 1/2} \equiv f(x_i \pm h/2)$ .

We can define products (composition) of the three operators. In particular,  $\Delta^2$  is given by

$$\Delta^2 f_i = \Delta(f_{i+1} - f_i) = f_{i+2} - 2f_{i+1} + f_i. \quad (2.13)$$

Similarly,

$$\nabla^2 f_i = \nabla(f_i - f_{i-1}) = f_i - 2f_{i-1} + f_{i-2}, \quad \delta^2 f_i = f_{i+1} - 2f_i + f_{i-1}. \quad (2.14)$$

We note that

$$\delta^2 f_i = \underbrace{f_{i+1} - f_i}_{=\Delta f_i} - \underbrace{(f_i - f_{i-1})}_{=\nabla f_i} = (\Delta - \nabla) f_i \Rightarrow \delta^2 = \Delta - \nabla.$$

This shows that the three operators are related.

shifting and  
averaging operators

It is convenient to introduce the **shifting** and **averaging** operators, respectively  $\mathbf{E}$  and  $\mu$ , as

$$\mathbf{E}f(x) = f(x+h), \quad \mu f(x) = \frac{1}{2} \left[ f\left(x + \frac{h}{2}\right) + f\left(x - \frac{h}{2}\right) \right]. \quad (2.15)$$

Note that for any positive integer  $n$ ,  $\mathbf{E}^n f(x) = f(x+nh)$ . We generalize this to any real number  $\alpha$ :

$$\mathbf{E}^\alpha f(x) = f(x+\alpha h). \quad (2.16)$$

All the other finite-difference operators can be written in terms of  $\mathbf{E}$ :

$$\Delta = \mathbf{E} - 1, \quad \nabla = 1 - \mathbf{E}^{-1}, \quad \delta = \mathbf{E}^{1/2} - \mathbf{E}^{-1/2}, \quad \mu = \frac{1}{2}(\mathbf{E}^{1/2} + \mathbf{E}^{-1/2}) \quad (2.17)$$

The first two equations of (2.17) can be rewritten as

$$\mathbf{E} = 1 + \Delta, \quad \mathbf{E} = (1 - \nabla)^{-1}. \quad (2.18)$$

We can obtain a useful formula for the shifting operator when it acts on polynomials of degree  $n$  or less. First note that

$$1 - \nabla^{n+1} = (1 - \nabla)(1 + \nabla + \cdots + \nabla^n).$$

But  $\nabla^{n+1}$  annihilates all polynomials of degree  $n$  or less (see Problem 2.33). Therefore, for such polynomials, we have

$$1 = (1 - \nabla)(1 + \nabla + \cdots + \nabla^n),$$

which shows that  $\mathbf{E} = (\mathbf{1} - \nabla)^{-1} = \mathbf{1} + \nabla + \dots + \nabla^n$ . Now let  $n \rightarrow \infty$  and obtain

$$\mathbf{E} = (\mathbf{1} - \nabla)^{-1} = \sum_{k=0}^{\infty} \nabla^k \quad (2.19)$$

for polynomials of any degree and—by Taylor expansion—for any (well-behaved) function.

**2.6.1. Example.** Numerical interpolation illustrates the use of the formulas derived above. Suppose that we are given a table of the values of a function, and we want the value of the function for an  $x$  located between two entries. We cannot use the table directly, but we may use the following procedure.

Assume that the values of the function  $f$  are given for  $x_1, x_2, \dots, x_i, \dots$ , and we are interested in the value of the function for  $x$  such that  $x_i < x < x_{i+1}$ . This corresponds to  $f_{i+r}$ , where  $0 < r < 1$ . We have

$$f_{i+r} = \mathbf{E}^r f_i = (\mathbf{1} + \Delta)^r f_i = \left( \mathbf{1} + r\Delta + \frac{r(r-1)}{2} \Delta^2 + \dots \right) f_i. \quad (2.20)$$

In practice, the infinite sum is truncated after a finite number of terms.

If only two terms are kept, we have

$$f_{i+r} \approx (\mathbf{1} + r\Delta) f_i = f_i + r(f_{i+1} - f_i) = (1-r)f_i + rf_{i+1}. \quad (2.21)$$

In particular, for  $r = \frac{1}{2}$ , Equation (2.21) yields  $f_{i+1/2} \approx \frac{1}{2}(f_i + f_{i+1})$ , which states the reasonable result that the value at the midpoint is approximately equal to the average of the values at the end points.

If the third term of the series in Equation (2.20) is also retained, then

$$\begin{aligned} f_{i+r} &\approx \left[ \mathbf{1} + r\Delta + \frac{r(r-1)}{2} \Delta^2 \right] f_i = f_i + r\Delta f_i + \frac{r(r-1)}{2} \Delta^2 f_i \\ &= f_i + r(f_{i+1} - f_i) + \frac{r(r-1)}{2} (f_{i+2} - 2f_{i+1} + f_i) \\ &= \frac{(2-r)(1-r)}{2} f_i + r(2-r)f_{i+1} + \frac{r(r-1)}{2} f_{i+2}. \end{aligned} \quad (2.22)$$

For  $r = \frac{1}{2}$ , that is, at the midpoint between  $x_i$  and  $x_{i+1}$ , Equation (2.22) yields

$$f_{i+1/2} \approx \frac{3}{8} f_i + \frac{3}{4} f_{i+1} - \frac{1}{8} f_{i+2},$$

which turns out to be a better approximation than Equation (2.21). However, it involves not only the two points on either side of  $x$  but also a relatively distant point,  $x_{i+2}$ . If we were to retain terms up to  $\Delta^k$  for  $k > 2$ , then  $f_{i+r}$  would be given in terms of  $f_i, f_{i+1}, \dots, f_{i+k}$ , and the result would be more accurate than (2.22). Thus, the more information we have about the behavior of a function at distant points, the better we can approximate it at  $x \in (x_i, x_{i+1})$ .

The foregoing analysis was based on forward interpolation. We may want to use backward interpolation, where  $f_{i-r}$  is sought for  $0 < r < 1$ . In such a case we use the backward difference operator

$$f_{i-r} = (\mathbf{E}^{-1})^r f_i = (\mathbf{1} - \nabla)^r f_i = \left( \mathbf{1} - r\nabla + \frac{r(r-1)}{2} \nabla^2 + \dots \right) f_i. \quad \blacksquare$$

**2.6.2. Example.** Let us check the conclusion made in Example 2.6.1 using a calculator and a specific function, say  $\sin x$ . A calculator gives  $\sin(0.1) = 0.0998334$ ,  $\sin(0.2) = 0.1986693$ ,  $\sin(0.3) = 0.2955202$ . Suppose that we want to find  $\sin(0.15)$  by interpolation. Using Equation (2.21) with  $r = \frac{1}{2}$ , we obtain

$$\sin(0.15) \approx \frac{1}{2}[\sin(0.1) + \sin(0.2)] = 0.1492514.$$

On the other hand, using (2.22) with  $r = \frac{1}{2}$  yields

$$\sin(0.15) \approx \frac{3}{8} \sin(0.1) + \frac{3}{4} \sin(0.2) - \frac{1}{8} \sin(0.3) = 0.1494995.$$

The value of  $\sin(0.15)$  obtained using a calculator is 0.1494381. It is clear that (2.22) gives a better estimate than (2.21). ■

### 2.6.2 Differentiation and Integration Operators

differentiation and  
integration operators

The two most important operations of mathematical physics can be written in terms of the finite-difference operators. Define the **differentiation operator**  $\mathbf{D}$  and the **integration operator**  $\mathbf{J}$  by

$$\mathbf{D}f(x) \equiv f'(x), \quad \mathbf{J}f(x) \equiv \int_x^{x+h} f(t) dt. \tag{2.23}$$

Assuming that  $\mathbf{D}^{-1}$  exists, we note that  $f(x) = \mathbf{D}^{-1}[f'(x)]$ . This shows that  $\mathbf{D}^{-1}$  is the operation of antidifferentiation:  $\mathbf{D}^{-1}f(x) = F(x)$ , where  $F$  is any primitive of  $f$ . On the other hand,  $\Delta F(x) = F(x+h) - F(x) = \mathbf{J}f(x)$ . These two equations and the fact that  $\mathbf{J}$  and  $\mathbf{D}$  commute (reader, verify!) show that

$$\Delta \mathbf{D}^{-1} = \mathbf{J} \Rightarrow \mathbf{J}\mathbf{D} = \mathbf{D}\mathbf{J} = \Delta = \mathbf{E} - \mathbf{1}. \tag{2.24}$$

Using the Taylor expansion, we can write

$$\mathbf{E}f(x) = f(x+h) = \sum_{n=0}^{\infty} \frac{h^n}{n!} f^{(n)}(x) = \left[ \sum_{n=0}^{\infty} \frac{h^n}{n!} \mathbf{D}^n \right] f(x) \equiv e^{h\mathbf{D}} f(x), \tag{2.25}$$

“exact” relation  
between shifting and  
differentiation  
operators

or  $\mathbf{E} = e^{h\mathbf{D}}$ . This yields  $h\mathbf{D} = \ln \mathbf{E}$ , or

$$\mathbf{D} = \frac{1}{h} \ln \mathbf{E} = \frac{1}{h} \ln(1 + \Delta) = \frac{1}{h} \left( \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \dots \right). \tag{2.26}$$

**2.6.3. Example.** Let us calculate  $\cos(0.1)$ , considering  $\cos x$  to be  $(d/dx)(\sin x)$  and using the values given in Example 2.6.2. Using Equation (2.26) to second order, we get

$$\begin{aligned} \mathbf{D}f_i &= \frac{1}{h} \left( \Delta - \frac{\Delta^2}{2} \right) f_i = \frac{1}{h} (\Delta f_i - \frac{1}{2} \Delta^2 f_i) \\ &= \frac{1}{h} [f_{i+1} - f_i - \frac{1}{2}(f_{i+2} - 2f_{i+1} + f_i)] = \frac{1}{2h} (-f_{i+2} + 4f_{i+1} - 3f_i). \end{aligned}$$

This gives

$$\cos(0.1) \approx \frac{1}{2}[-0.295520 + 4(0.198669) - 3(0.099833)] = 0.998284.$$

In comparison, the value obtained directly from a calculator is 0.995004. ■

“exact” relation  
between integration  
and difference  
operators

The operator  $\mathbf{J}$  is a special case of a more general operator defined by

$$\begin{aligned} \mathbf{J}_\alpha f(x) &= \int_x^{x+\alpha h} f(t) dt = F(x + \alpha h) - F(x) \\ &= (\mathbf{E}^\alpha - \mathbf{1})F(x) = (\mathbf{E}^\alpha - \mathbf{1})\mathbf{D}^{-1} f(x), \end{aligned}$$

or

$$\mathbf{J}_\alpha = (\mathbf{E}^\alpha - \mathbf{1})\mathbf{D}^{-1} = h \frac{\mathbf{E}^\alpha - \mathbf{1}}{\ln \mathbf{E}} = h \frac{(1 + \Delta)^\alpha - 1}{\ln(1 + \Delta)}, \quad (2.27)$$

where we used Equation (2.26).

### 2.6.3 Numerical Integration

Suppose that we are interested in the numerical value of  $\int_a^b f(x) dx$ . Let  $x_0 \equiv a$  and  $x_N \equiv b$ , and divide the interval  $[a, b]$  into  $N$  equal parts, each of length  $h = (b - a)/N$ . The method commonly used in calculating integrals is to find the integral  $\int_{x_i}^{x_i + \alpha h} f(x) dx$ , where  $\alpha$  is a suitable number, and then add all such integrals. More specifically,

$$I = \int_{x_0}^{x_0 + \alpha h} f(x) dx + \int_{x_0 + \alpha h}^{x_0 + 2\alpha h} f(x) dx + \cdots + \int_{x_0 + (M-1)\alpha h}^{x_0 + M\alpha h} f(x) dx,$$

where  $M$  is a suitably chosen number. In fact, since  $x_N = x_0 + Nh$ , we have  $M\alpha = N$ . We next employ Equation (2.27) to get

$$I = \mathbf{J}_\alpha f_0 + \mathbf{J}_\alpha f_\alpha + \cdots + \mathbf{J}_\alpha f_{(M-1)\alpha} = \mathbf{J}_\alpha \left( \sum_{k=0}^{M-1} f_{k\alpha} \right), \quad (2.28)$$

where  $f_{k\alpha} \equiv f(x_0 + k\alpha h)$ . We thus need an expression for  $\mathbf{J}_\alpha$  to evaluate the integral. Such an expression can be derived elegantly, by noting that

$$\int_0^\alpha \mathbf{E}^s ds = \left. \frac{\mathbf{E}^s}{\ln \mathbf{E}} \right|_0^\alpha = \frac{\mathbf{E}^\alpha - \mathbf{1}}{\ln \mathbf{E}} = \frac{1}{h} \mathbf{J}_\alpha, \quad [\text{by Equation (2.27)}]$$

so that

$$\begin{aligned} \mathbf{J}_\alpha &= h \int_0^\alpha \mathbf{E}^s ds = h \int_0^\alpha (1 + \Delta)^s ds \equiv h \sum_{k=0}^{\infty} a_k \Delta^k, \\ a_k &= \frac{1}{k!} \int_0^\alpha s(s-1) \cdots (s-k+1) ds, \end{aligned} \quad (2.29)$$

where we expanded  $(1 + \Delta)^s$  using the binomial infinite series. Equations (2.28) and (2.29) give the desired evaluation of the integral.

Let us make a few remarks before developing any commonly used rules of approximation. First, once  $h$  is set, the function can be evaluated only at  $x_0 + nh$ , where  $n$  is a positive integer. This means that  $f_n$  is given only for positive integers  $n$ . Thus, in the sum in (2.28)  $k\alpha$  must be an integer. Since  $k$  is an integer, we conclude that  $\alpha$  must be an integer. Second, since  $N = M\alpha$  for some integer  $M$ , we must choose  $N$  to be a multiple of  $\alpha$ . Third, if we are to be able to evaluate  $J_\alpha f_{(M-1)\alpha}$  [the last term in (2.28)],  $J_\alpha$  cannot have powers of  $\Delta$  higher than  $\alpha$ , because  $\Delta^n f_{(M-1)\alpha}$  contains a term of the form

$$f(x_0 + (M - 1)\alpha h + nh) = f(x_N + (n - \alpha)h),$$

which for  $n > \alpha$  gives  $f$  at a point beyond the upper limit. Thus, in the power-series expansion of  $J_\alpha$ , we must make sure that no power of  $\Delta$  beyond  $\alpha$  is retained.

trapezoidal rule for numerical integration

There are several specific  $J_\alpha$ 's commonly used in numerical integration. We will consider these next. The **trapezoidal rule** sets  $\alpha = 1$ . According to the remarks above, we therefore retain terms up to the first power in the expansion of  $J_\alpha$ . Then (2.29) gives  $J_1 = J = h(1 + \frac{1}{2}\Delta)$ . Substituting this in Equation (2.28), we obtain

$$\begin{aligned} I &= h(1 + \frac{1}{2}\Delta) \left( \sum_{k=0}^{N-1} f_k \right) = h \sum_{k=0}^{N-1} [f_k + \frac{1}{2}(f_{k+1} - f_k)] \\ &= \frac{h}{2} \sum_{k=0}^{N-1} (f_k + f_{k+1}) = \frac{h}{2} (f_0 + 2f_1 + \dots + 2f_{N-1} + f_N). \end{aligned} \tag{2.30}$$

Simpson's one-third rule for numerical integration

**Simpson's one-third rule** sets  $\alpha = 2$ . Thus, we have to retain all terms up to the  $\Delta^2$  term. However, for  $\alpha = 2$ , the third power of  $\Delta$  disappears in Equation (2.29), and we get an extra "power" of accuracy for free! Because of this, Simpson's one-third rule is popular for numerical integrations. Equation (2.29) yields  $J_2 = 2h(1 + \Delta + \frac{1}{6}\Delta^2)$ . Substituting this in (2.28) yields

$$\begin{aligned} I &= \frac{h}{3} \sum_{k=0}^{N/2-1} (61 + 6\Delta + \Delta^2) f_{2k} = \frac{h}{3} \sum_{k=0}^{N/2-1} (f_{2k+2} + 4f_{2k+1} + f_{2k}) \\ &= \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + \dots + 4f_{N-1} + f_N). \end{aligned} \tag{2.31}$$

It is understood, of course, that  $N$  is an even integer. The factor  $\frac{1}{3}$  gives this method its name.

Simpson's three-eighths rule for numerical integration

For **Simpson's three-eighths rule**, we set  $\alpha = 3$ , retain terms up to  $\Delta^3$ , and use Equation (2.29) to obtain

$$J_3 = 3h(1 + \frac{3}{2}\Delta + \frac{3}{4}\Delta^2 + \frac{1}{8}\Delta^3) = \frac{3h}{8} (81 + 12\Delta + 6\Delta^2 + \Delta^3).$$

Substituting in (2.28), we get

$$\begin{aligned} I &= \frac{3h}{8} \sum_{k=0}^{N/3-1} (81 + 12\Delta + 6\Delta^2 + \Delta^3) f_{3k} \\ &= \frac{3h}{8} \sum_{k=0}^{N/3-1} (f_{3k+3} + 3f_{3k+2} + 3f_{3k+1} + f_{3k}). \end{aligned} \quad (2.32)$$

**2.6.4. Example.** Let us use Simpson's one-third rule with four intervals to evaluate the familiar integral  $I = \int_0^1 e^x dx$ . With  $h = 0.25$  and  $N = 4$ , Equation (2.31) yields

$$I \approx \frac{0.25}{3} (1 + 4e^{0.25} + 2e^{0.5} + 4e^{0.75} + e) = 1.71832.$$

This is very close to the "exact" result  $e - 1 \approx 1.71828$ . ■

## 2.7 Problems

**2.1.** Consider a linear operator  $\mathbf{T}$  on a finite-dimensional vector space  $\mathcal{V}$ . Show that there exists a polynomial  $p$  such that  $p(\mathbf{T}) = \mathbf{0}$ . Hint: Take a basis  $B = \{|a_i\rangle\}_{i=1}^N$  and consider the vectors  $\{\mathbf{T}^k |a_1\rangle\}_{k=0}^M$  for large enough  $M$  and conclude that there exists a polynomial  $p_1(\mathbf{T})$  such that  $p_1(\mathbf{T}) |a_1\rangle = \mathbf{0}$ . Do the same for  $|a_2\rangle$ , etc. Now take the product of all such polynomials.

**2.2.** Use mathematical induction to show that  $[\mathbf{A}, \mathbf{A}^m] = \mathbf{0}$ .

**2.3.** For  $\mathbf{D}$  and  $\mathbf{T}$  defined in Example 1.3.4:

(a) Show that  $[\mathbf{D}, \mathbf{T}] = \mathbf{1}$ .

(b) Calculate the linear transformations  $\mathbf{D}^3\mathbf{T}^3$  and  $\mathbf{T}^3\mathbf{D}^3$ .

**2.4.** Consider three linear operators  $\mathbf{L}_1$ ,  $\mathbf{L}_2$ , and  $\mathbf{L}_3$  satisfying the commutation relations  $[\mathbf{L}_1, \mathbf{L}_2] = \mathbf{L}_3$ ,  $[\mathbf{L}_3, \mathbf{L}_1] = \mathbf{L}_2$ ,  $[\mathbf{L}_2, \mathbf{L}_3] = \mathbf{L}_1$ , and define the new operators  $\mathbf{L}_\pm = \mathbf{L}_1 \pm i\mathbf{L}_2$ .

(a) Show that the operator  $\mathbf{L}^2 \equiv \mathbf{L}_1^2 + \mathbf{L}_2^2 + \mathbf{L}_3^2$  commutes with  $\mathbf{L}_k$ ,  $k = 1, 2, 3$ .

(b) Show that the set  $\{\mathbf{L}_+, \mathbf{L}_-, \mathbf{L}_3\}$  is closed under commutation, i.e., the commutator of any two of them can be written as a linear combination of the set. Determine these commutators.

(c) Write  $\mathbf{L}^2$  in terms of  $\mathbf{L}_+$ ,  $\mathbf{L}_-$ , and  $\mathbf{L}_3$ .

**2.5.** Prove the rest of Proposition 2.1.11.

**2.6.** Show that if  $[[\mathbf{A}, \mathbf{B}], \mathbf{A}] = \mathbf{0}$ , then for every positive integer  $k$ ,

$$[\mathbf{A}^k, \mathbf{B}] = k\mathbf{A}^{k-1}[\mathbf{A}, \mathbf{B}].$$

Hint: First prove the relation for low values of  $k$ ; then use mathematical induction.



2.7. Show that for  $\mathbf{D}$  and  $\mathbf{T}$  defined in Example 1.3.4,  $[\mathbf{D}^k, \mathbf{T}] = k\mathbf{D}^{k-1}$  and  $[\mathbf{T}^k, \mathbf{D}] = -k\mathbf{T}^{k-1}$ .

2.8. Evaluate the derivative of  $\mathbf{H}^{-1}(t)$  in terms of the derivative of  $\mathbf{H}(t)$ .

2.9. Show that for any  $\alpha, \beta \in \mathbb{R}$  and any  $\mathbf{H} \in \mathcal{L}(\mathcal{V})$ , we have

$$e^{\alpha\mathbf{H}} e^{\beta\mathbf{H}} = e^{(\alpha+\beta)\mathbf{H}}.$$

2.10. Show that  $(\mathbf{U} + \mathbf{T})(\mathbf{U} - \mathbf{T}) = \mathbf{U}^2 - \mathbf{T}^2$  if and only if  $[\mathbf{U}, \mathbf{T}] = \mathbf{0}$ .

2.11. Prove that if  $\mathbf{A}$  and  $\mathbf{B}$  are hermitian, then  $i[\mathbf{A}, \mathbf{B}]$  is also hermitian.

2.12. Find the solution to the operator differential equation

$$\frac{d\mathbf{U}}{dt} = t\mathbf{H}\mathbf{U}(t).$$

Hint: Make the change of variable  $y = t^2$  and use the result of Example 2.2.3.

2.13. Verify that

$$\frac{d}{dt}\mathbf{H}^3 = \left(\frac{d\mathbf{H}}{dt}\right)\mathbf{H}^2 + \mathbf{H}\left(\frac{d\mathbf{H}}{dt}\right)\mathbf{H} + \mathbf{H}^2\left(\frac{d\mathbf{H}}{dt}\right).$$

2.14. Show that if  $\mathbf{A}$  and  $\mathbf{B}$  commute, and  $f$  and  $g$  are arbitrary functions, then  $f(\mathbf{A})$  and  $g(\mathbf{B})$  also commute.

2.15. Assuming that  $[[\mathbf{S}, \mathbf{T}], \mathbf{T}] = \mathbf{0} = [[\mathbf{S}, \mathbf{T}], \mathbf{S}]$ , show that

$$[\mathbf{S}, \exp(t\mathbf{T})] = t[\mathbf{S}, \mathbf{T}]\exp(t\mathbf{T}).$$

Hint: Expand the exponential and use Problem 2.6.

2.16. Prove that

$$\begin{aligned} \exp(\mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3) &= \exp(\mathbf{H}_1)\exp(\mathbf{H}_2)\exp(\mathbf{H}_3) \\ &\quad \cdot \exp\left\{-\frac{1}{2}([\mathbf{H}_1, \mathbf{H}_2] + [\mathbf{H}_1, \mathbf{H}_3] + [\mathbf{H}_2, \mathbf{H}_3])\right\} \end{aligned}$$

provided that  $\mathbf{H}_1, \mathbf{H}_2$ , and  $\mathbf{H}_3$  commute with all the commutators. What is the generalization to  $\mathbf{H}_1 + \mathbf{H}_2 + \cdots + \mathbf{H}_n$ ?

2.17. Denoting the derivative of  $\mathbf{A}(t)$  by  $\dot{\mathbf{A}}$ , show that

$$\frac{d}{dt}[\mathbf{A}, \mathbf{B}] = [\dot{\mathbf{A}}, \mathbf{B}] + [\mathbf{A}, \dot{\mathbf{B}}].$$

2.18. Prove Theorem 2.3.2. Hint: Use Equation (2.8) and Theorem 2.1.3.

2.19. Let  $\mathbf{A}(t) \equiv \exp(t\mathbf{H})\mathbf{A}_0\exp(-t\mathbf{H})$ , where  $\mathbf{H}$  and  $\mathbf{A}_0$  are constant operators. Show that  $d\mathbf{A}/dt = [\mathbf{H}, \mathbf{A}(t)]$ . What happens when  $\mathbf{H}$  commutes with  $\mathbf{A}(t)$ ?

**2.20.** Let  $|f\rangle, |g\rangle \in \mathbb{C}(a, b)$  with the additional property that  $f(a) = g(a) = f(b) = g(b) = 0$ . Show that for such functions, the derivative operator  $\mathbf{D}$  is anti-hermitian. The inner product is defined as usual:

$$\langle f | g \rangle \equiv \int_a^b f^*(t)g(t) dt.$$

Heisenberg  
uncertainty principle

**2.21.** In this problem, you will go through the steps of proving the rigorous statement of the **Heisenberg uncertainty principle**. Denote the expectation (average) value of an operator  $\mathbf{A}$  in a state  $|\Psi\rangle$  by  $A_{\text{avg}}$ . Thus,  $A_{\text{avg}} = \langle A \rangle = \langle \Psi | \mathbf{A} | \Psi \rangle$ . The *uncertainty* (deviation from the mean) in state  $|\Psi\rangle$  of the operator  $\mathbf{A}$  is given by

$$\Delta A = \sqrt{\langle (A - A_{\text{avg}})^2 \rangle} = \sqrt{\langle \Psi | (\mathbf{A} - A_{\text{avg}}\mathbf{1})^2 | \Psi \rangle}.$$

(a) Show that for any two *hermitian* operators  $\mathbf{A}$  and  $\mathbf{B}$ , we have

$$|\langle \Psi | \mathbf{A}\mathbf{B} | \Psi \rangle|^2 \leq \langle \Psi | \mathbf{A}^2 | \Psi \rangle \langle \Psi | \mathbf{B}^2 | \Psi \rangle.$$

Hint: Apply the Schwarz inequality to an appropriate pair of vectors.

(b) Using the above and the triangle inequality for complex numbers, show that

$$|\langle \Psi | [\mathbf{A}, \mathbf{B}] | \Psi \rangle|^2 \leq 4 \langle \Psi | \mathbf{A}^2 | \Psi \rangle \langle \Psi | \mathbf{B}^2 | \Psi \rangle.$$

(c) Define the operators  $\mathbf{A}' = \mathbf{A} - \alpha\mathbf{1}$ ,  $\mathbf{B}' = \mathbf{B} - \beta\mathbf{1}$ , where  $\alpha$  and  $\beta$  are *real* numbers. Show that  $\mathbf{A}'$  and  $\mathbf{B}'$  are hermitian and  $[\mathbf{A}', \mathbf{B}'] = [\mathbf{A}, \mathbf{B}]$ .

(d) Now use all the results above to show the celebrated uncertainty relation

$$(\Delta A)(\Delta B) \geq \frac{1}{2} |\langle \Psi | [\mathbf{A}, \mathbf{B}] | \Psi \rangle|.$$

What does this reduce to for position operator  $\mathbf{x}$  and momentum operator  $\mathbf{p}$  if  $[\mathbf{x}, \mathbf{p}] = i\hbar$ ?

**2.22.** Show that  $\mathbf{U} = \exp \mathbf{A}$  is unitary if and only if  $\mathbf{A}$  is anti-hermitian.

**2.23.** Find  $\mathbf{T}^\dagger$  for each of the following linear operators.

(a)  $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  given by

$$\mathbf{T} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x - y \end{pmatrix}.$$

(b)  $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  given by

$$\mathbf{T} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y - z \\ 3x - y + 2z \\ -x + 2y + 3z \end{pmatrix}.$$

(c)  $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  given by

$$\mathbf{T} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{pmatrix},$$

where  $\theta$  is a real number. What is  $\mathbf{T}^\dagger \mathbf{T}$ ?

(d)  $\mathbf{T} : \mathbb{C}^2 \rightarrow \mathbb{C}^2$  given by

$$\mathbf{T} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 - i\alpha_2 \\ i\alpha_1 + \alpha_2 \end{pmatrix}.$$

(e)  $\mathbf{T} : \mathbb{C}^3 \rightarrow \mathbb{C}^3$  given by

$$\mathbf{T} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 + i\alpha_2 - 2i\alpha_3 \\ -2i\alpha_1 + \alpha_2 + i\alpha_3 \\ i\alpha_1 - 2i\alpha_2 + \alpha_3 \end{pmatrix}.$$

**2.24.** Show that if  $\mathbf{P}$  is a (hermitian) projection operator, so are (a)  $\mathbf{1} - \mathbf{P}$  and (b)  $\mathbf{U}^\dagger \mathbf{P} \mathbf{U}$  for any unitary operator  $\mathbf{U}$ .

**2.25.** For the vector

$$|a\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

(a) Find the associated projection matrix,  $\mathbf{P}_a$ .

(b) Verify that  $\mathbf{P}_a$  does project an arbitrary vector in  $\mathbb{C}^4$  along  $|a\rangle$ .

(c) Verify directly that the matrix  $\mathbf{1} - \mathbf{P}_a$  is also a projection operator.

**2.26.** Let  $|a_1\rangle \equiv \mathbf{a}_1 = (1, 1, -1)$  and  $|a_2\rangle \equiv \mathbf{a}_2 = (-2, 1, -1)$ .

(a) Construct (in the form of a matrix) the projection operators  $\mathbf{P}_1$  and  $\mathbf{P}_2$  that project onto the directions of  $|a_1\rangle$  and  $|a_2\rangle$ , respectively. Verify that they are indeed projection operators.

(b) Construct (in the form of a matrix) the operator  $\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2$  and verify directly that it is a projection operator.

(c) Let  $\mathbf{P}$  act on an arbitrary vector  $(x, y, z)$ . What is the dot product of the resulting vector with the vector  $\mathbf{a}_1 \times \mathbf{a}_2$ ? What can you say about  $\mathbf{P}$  and your conclusion in (b)?

**2.27.** Let  $\mathbf{P}^{(m)} = \sum_{i=1}^m |e_i\rangle \langle e_i|$  be a projection operator constructed out of the first  $m$  orthonormal vectors of the basis  $B = \{|e_i\rangle\}_{i=1}^N$  of  $\mathcal{V}$ . Show that  $\mathbf{P}^{(m)}$  projects into the subspace spanned by the first  $m$  vectors in  $B$ .

**2.28.** What is the length of the projection of the vector  $(3, 4, -4)$  onto a line whose parametric equation is  $x = 2t + 1, y = -t + 3, z = t - 1$ ? Hint: Find a unit vector in the direction of the line and construct its projection operator.

**2.29.** The parametric equation of a line  $L$  in a coordinate system with origin  $O$  is

$$x = 2t + 1, \quad y = t + 1, \quad z = -2t + 2.$$

A point  $P$  has coordinates  $(3, -2, 1)$ .

- (a) Using the projection operators, find the length of the projection of  $\overline{OP}$  on the line  $L$ .  
 (b) Find the vector whose beginning is  $P$  and ends on the line  $L$  and perpendicular to  $L$ .  
 (c) From this vector calculate the distance from  $P$  to the line  $L$ .

2.30. Let the operator  $\mathbf{U} : \mathbb{C}^2 \rightarrow \mathbb{C}^2$  be given by

$$\mathbf{U} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} i \frac{\alpha_1}{\sqrt{2}} - i \frac{\alpha_2}{\sqrt{2}} \\ \frac{\alpha_1}{\sqrt{2}} + \frac{\alpha_2}{\sqrt{2}} \end{pmatrix}.$$

Is  $\mathbf{U}$  unitary?

2.31. Show that the product of two unitary operators is always unitary, but the product of two hermitian operators is hermitian if and only if they commute.

2.32. Let  $\mathbf{S}$  be an operator that is both unitary and hermitian. Show that

- (a)  $\mathbf{S}$  is involutive (i.e.,  $\mathbf{S}^2 = \mathbf{1}$ ).  
 (b)  $\mathbf{S} = \mathbf{P}^+ - \mathbf{P}^-$ , where  $\mathbf{P}^+$  and  $\mathbf{P}^-$  are hermitian projection operators.

2.33. Show that when the forward difference operator is applied to a polynomial, the degree of the polynomial is reduced by 1. (Hint: Consider  $x^n$  first.) Then show that  $\Delta^{n+1}$  annihilates all polynomials of degree  $n$  or less.

2.34. Show that  $\delta^{n+1}$  and  $\nabla^{n+1}$  annihilate any polynomial of degree  $n$ .

2.35. Show that all of the finite-difference operators commute with one another.

2.36. Verify the identities

$$\begin{aligned} \nabla \mathbf{E} &= \delta \mathbf{E}^{1/2} = \Delta, \\ \nabla + \Delta &= 2\mu\delta = \mathbf{E} - \mathbf{E}^{-1}, \\ \mathbf{E}^{-1/2} &= \mu - (\delta/2). \end{aligned}$$

2.37. By writing everything in terms of  $\mathbf{E}$ , show that  $\delta^2 = \Delta - \nabla = \Delta \nabla$ .

2.38. Write expressions for  $\mathbf{E}^{1/2}$ ,  $\Delta$ ,  $\nabla$ , and  $\mu$  in terms of  $\delta$ .

2.39. Show that

$$\mathbf{D} = \frac{2}{h} \sinh^{-1} \left( \frac{\delta}{2} \right).$$

2.40. Show that

$$\mathbf{D}^2 = \frac{1}{h^2} (\Delta^2 - \Delta^3 + \frac{11}{12} \Delta^4 - \frac{5}{6} \Delta^5 - \dots)$$

and derive Equation (2.27).

2.41. Find an expression for  $J_\alpha$  in powers of  $\Delta$ . Retain all terms up to the fourth power.

2.42. Show that for  $\alpha = 2$ , the third power of  $\Delta$  disappears in Equation (2.29).

2.43. Evaluate the following integrals numerically, using six subintervals with the trapezoidal rule, Simpson's one-third rule, and Simpson's three-eighths rule. Compare with the exact result when possible.

$$(a) \int_0^5 x^3 dx.$$

$$(b) \int_0^2 e^{-x^2} dx.$$

$$(c) \int_0^{1.2} x e^x \cos x dx.$$

$$(d) \int_1^4 \frac{1}{x} dx.$$

$$(e) \int_1^4 \ln x dx.$$

$$(f) \int_1^2 e^{x^2} \sin x dx.$$

$$(g) \int_{-1}^1 \frac{dx}{1+x^2}.$$

$$(h) \int_0^1 x e^{x^2} dx.$$

$$(i) \int_0^1 e^x \tan x dx.$$

### Additional Reading

1. Axler, S. *Linear Algebra Done Right*, Springer-Verlag, 1996.
2. Greub, W. *Linear Algebra*, 4th ed., Springer-Verlag, 1975.
3. Hildebrand, F. *Introduction to Numerical Analysis*, 2nd ed., Dover, 1987. Uses operator techniques in numerical analysis. It has a detailed discussion of error analysis, a topic completely ignored in our text.

# Matrices: Operator Representations

So far, our theoretical investigation has been dealing mostly with abstract vectors and abstract operators. As we have seen in examples and problems, concrete representations of vectors and operators are necessary in most applications. Such representations are obtained by choosing a basis and expressing all operations in terms of components of vectors and matrix representations of operators.

## 3.1 Matrices

Let us choose a basis  $B_V = \{|a_i\rangle\}_{i=1}^N$  of a vector space  $\mathcal{V}_N$ , and express an arbitrary vector  $|x\rangle$  in this basis:  $|x\rangle = \sum_{i=1}^N \xi_i |a_i\rangle$ . We write

$$\mathbf{x} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_N \end{pmatrix} \quad (3.1)$$

representation of  
vectors

and say that the column vector  $\mathbf{x}$  **represents**  $|x\rangle$  in  $B_V$ . We can also have a linear transformation  $\mathbf{A} \in \mathcal{L}(\mathcal{V}_N, \mathcal{W}_M)$  act on the basis vectors in  $B_V$  to give vectors in the  $M$ -dimensional vector space  $\mathcal{W}_M$ :  $|w_k\rangle = \mathbf{A} |a_k\rangle$ . The latter can be written as a linear combination of basis vectors  $B_W = \{|b_j\rangle\}_{j=1}^M$  in  $\mathcal{W}_M$ :

$$|w_1\rangle = \sum_{j=1}^M \alpha_{j1} |b_j\rangle, \quad |w_2\rangle = \sum_{j=1}^M \alpha_{j2} |b_j\rangle, \quad \dots, \quad |w_N\rangle = \sum_{j=1}^M \alpha_{jN} |b_j\rangle.$$

Note that the components have an extra subscript to denote which of the  $N$  vectors  $\{|w_i\rangle\}_{i=1}^N$  they are representing. The components can be arranged in a column as

before to give a representation of the corresponding vectors:

$$\mathbf{w}_1 = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{M1} \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{M2} \end{pmatrix}, \quad \dots, \quad \mathbf{w}_N = \begin{pmatrix} \alpha_{1N} \\ \alpha_{2N} \\ \vdots \\ \alpha_{MN} \end{pmatrix}.$$

The operator itself is determined by the collection of all these vectors, i.e., by a matrix. We write this as

$$\mathbf{A} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2N} \\ \vdots & \vdots & & \vdots \\ \alpha_{M1} & \alpha_{M2} & \dots & \alpha_{MN} \end{pmatrix} \quad (3.2)$$

representation of  
operators

and call  $\mathbf{A}$  the **matrix representing  $\mathbf{A}$  in bases  $B_V$  and  $B_W$** . This statement is also summarized symbolically as

$$\mathbf{A}|a_i\rangle = \sum_{j=1}^M \alpha_{ji} |b_j\rangle, \quad i = 1, 2, \dots, N. \quad (3.3)$$

We thus have the following rule:

**3.1.1. Box.** To find the matrix  $\mathbf{A}$  representing  $\mathbf{A}$  in bases  $B_V = \{|a_i\rangle\}_{i=1}^N$  and  $B_W = \{|b_j\rangle\}_{j=1}^M$ , express  $\mathbf{A}|a_i\rangle$  as a linear combination of the vectors in  $B_W$ . The components form the  $i$ th column of  $\mathbf{A}$ .

Now consider the vector  $|y\rangle = \mathbf{A}|x\rangle$  in  $\mathcal{W}_M$ . This vector can be written in two ways: On the one hand,  $|y\rangle = \sum_{j=1}^M \eta_j |b_j\rangle$ . On the other hand,

$$\begin{aligned} |y\rangle &= \mathbf{A}|x\rangle = \mathbf{A} \sum_{i=1}^N \xi_i |a_i\rangle = \sum_{i=1}^N \xi_i \mathbf{A}|a_i\rangle \\ &= \sum_{i=1}^N \xi_i \left( \sum_{j=1}^M \alpha_{ji} |b_j\rangle \right) = \sum_{j=1}^M \left( \sum_{i=1}^N \xi_i \alpha_{ji} \right) |b_j\rangle. \end{aligned}$$

Since  $|y\rangle$  has a unique set of components in the basis  $B_W$ , we conclude that

$$\eta_j = \sum_{i=1}^N \alpha_{ji} \xi_i, \quad j = 1, 2, \dots, M. \quad (3.4)$$

This is written as

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_M \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2N} \\ \vdots & \vdots & & \vdots \\ \alpha_{M1} & \alpha_{M2} & \dots & \alpha_{MN} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_N \end{pmatrix} \Rightarrow y = Ax, \quad (3.5)$$

in which the matrix multiplication rule is understood. This matrix equation is the representation of the operator equation  $|y\rangle = \mathbf{A}|x\rangle$  in the bases  $B_V$  and  $B_W$ .

The operator  $\mathbf{T}_A$   
associated with a  
matrix  $\mathbf{A}$

The construction above indicates that—once the bases are fixed in the two vector spaces—to every operator there corresponds a *unique* matrix. This uniqueness is the result of the uniqueness of the components of vectors in a basis. On the other hand, given an  $M \times N$  matrix  $\mathbf{A}$  with elements  $\alpha_{ij}$ , one can construct a unique linear operator  $\mathbf{T}_A$  defined by its action on the basis vectors (see Box 1.3.3):  $\mathbf{T}_A|a_i\rangle \equiv \sum_{j=1}^M \alpha_{ji}|b_j\rangle$ . Thus, there is a one-to-one correspondence between operators and matrices. This correspondence is in fact a linear isomorphism:

**3.1.2. Proposition.** *The two vector spaces  $\mathcal{L}(\mathcal{V}_N, \mathcal{W}_M)$  and  $\mathcal{M}^{M \times N}$  are isomorphic. An explicit isomorphism is established only when a basis is chosen for each vector space, in which case, an operator is identified with its matrix representation.*

Given the linear transformations  $\mathbf{A} : \mathcal{V}_N \rightarrow \mathcal{W}_M$  and  $\mathbf{B} : \mathcal{W}_M \rightarrow \mathcal{U}_K$ , we can form the composite linear transformation  $\mathbf{B} \circ \mathbf{A} : \mathcal{V}_N \rightarrow \mathcal{U}_K$ . We can also choose bases  $B_V = \{|a_i\rangle\}_{i=1}^N$ ,  $B_W = \{|b_i\rangle\}_{i=1}^M$ ,  $B_U = \{|c_i\rangle\}_{i=1}^K$  for  $\mathcal{V}$ ,  $\mathcal{W}$ , and  $\mathcal{U}$ , respectively. Then  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{B} \circ \mathbf{A}$  will be represented by an  $M \times N$ , a  $K \times M$ , and a  $K \times N$  matrix, respectively, the latter being the matrix product of the other two matrices. Matrices are determined entirely by their elements. For this reason a matrix  $\mathbf{A}$  whose elements are  $\alpha_{11}, \alpha_{12}, \dots$  is sometimes denoted by  $(\alpha_{ij})$ . Similarly, the elements of this matrix are denoted by  $(A)_{ij}$ . So, on the one hand, we have  $(\alpha_{ij}) = \mathbf{A}$ , and on the other hand  $(A)_{ij} = \alpha_{ij}$ . In the context of this notation, therefore, we can write

$$\begin{aligned} (\mathbf{A} + \mathbf{B})_{ij} &= (\mathbf{A})_{ij} + (\mathbf{B})_{ij} \Rightarrow (\alpha_{ij} + \beta_{ij}) = (\alpha_{ij}) + (\beta_{ij}), \\ (\gamma \mathbf{A})_{ij} &= \gamma (\mathbf{A})_{ij} \Rightarrow \gamma (\alpha_{ij}) = (\gamma \alpha_{ij}), \\ (\mathbf{0})_{ij} &= 0, \\ (\mathbf{1})_{ij} &= \delta_{ij}. \end{aligned}$$

A matrix as a representation of a linear operator is well-defined only in reference to a specific basis. A collection of rows and columns of numbers by themselves have no operational meaning. When we manipulate matrices and attach meaning to them, we make an unannounced assumption regarding the basis: We have the standard basis of  $\mathbb{C}^n$  (or  $\mathbb{R}^n$ ) in mind. The following example should clarify this subtlety.



**3.1.3. Example.** Let us find the matrix representation of the linear operator  $\mathbf{A} \in \mathcal{L}(\mathbb{R}^3)$ , given by

$$\mathbf{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y + 2z \\ 3x - z \\ 2y + z \end{pmatrix} \quad (3.6)$$

in the basis

$$B = \left\{ |a_1\rangle = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, |a_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, |a_3\rangle = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

There is a tendency to associate the matrix

$$\begin{pmatrix} 1 & -1 & 2 \\ 3 & 0 & -1 \\ 0 & 2 & 1 \end{pmatrix}$$

with the operator  $\mathbf{A}$ . The following discussion will show that this is false. To obtain the first column of the matrix representing  $\mathbf{A}$ , we consider

$$\mathbf{A}|a_1\rangle = \mathbf{A} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{5}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

So, by Box 3.1.1, the first column of the matrix is

$$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{5}{2} \end{pmatrix}.$$

The other two columns are obtained from

$$\mathbf{A}|a_2\rangle = \mathbf{A} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

$$\mathbf{A}|a_3\rangle = \mathbf{A} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = -\frac{3}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{5}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

giving the second and the third columns, respectively. The whole matrix is then

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & 2 & -\frac{3}{2} \\ -\frac{1}{2} & 1 & \frac{5}{2} \\ \frac{5}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

As long as all vectors are represented by columns whose entries are expansion coefficients of the vectors in  $B$ ,  $\mathbf{A}$  and  $A$  are indistinguishable. However, the action of  $A$  on the column

vector  $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$  will not yield the RHS of Equation (3.6)! Although this is not usually emphasized, the column vector on the LHS of Equation (3.6) is really the vector

$$x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

which is an expansion in terms of the *standard basis* of  $\mathbb{R}^3$  rather than in terms of  $B$ .

We can expand  $\mathbf{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$  in terms of  $B$ , yielding

$$\begin{aligned} \mathbf{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} x - y + 2z \\ 3x - z \\ 2y + z \end{pmatrix} \\ &= (2x - \frac{3}{2}y) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (-x + \frac{1}{2}y + 2z) \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + (x + \frac{3}{2}y - z) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

This says that in the basis  $B$  this vector has the representation

$$\left( \mathbf{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right)_B = \begin{pmatrix} 2x - \frac{3}{2}y \\ -x + \frac{1}{2}y + 2z \\ x + \frac{3}{2}y - z \end{pmatrix}. \quad (3.7)$$

Similarly,  $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$  is represented by

$$\left( \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right)_B = \begin{pmatrix} \frac{1}{2}x + \frac{1}{2}y - \frac{1}{2}z \\ \frac{1}{2}x - \frac{1}{2}y + \frac{1}{2}z \\ -\frac{1}{2}x + \frac{1}{2}y + \frac{1}{2}z \end{pmatrix}. \quad (3.8)$$

Applying  $\mathbf{A}$  to the RHS of (3.8) yields the RHS of (3.7), as it should.  $\blacksquare$

Given any  $M \times N$  matrix  $\mathbf{A}$ , an operator  $\mathbf{T}_A \in \mathcal{L}(\mathcal{V}_N, \mathcal{W}_M)$  can be associated with  $\mathbf{A}$ , and one can construct the kernel and the range of  $\mathbf{T}_A$ . The rank of  $\mathbf{T}_A$  is called the **rank of  $\mathbf{A}$** . Since the rank of an operator is basis independent, this definition makes sense.

rank of a matrix

Now suppose that we choose a basis for the kernel of  $\mathbf{T}_A$  and extend it to a basis of  $\mathcal{V}$ . Let  $\mathcal{V}_1$  denote the span of the remaining basis vectors. Similarly, we choose a basis for  $\mathbf{T}_A(\mathcal{V})$  and extend it to a basis for  $\mathcal{W}$ . In these two bases, the  $M \times N$  matrix representing  $\mathbf{T}_A$  will have all zeros except for an  $r \times r$  submatrix, where  $r$  is the rank of  $\mathbf{T}_A$ . The reader may verify that this submatrix has a nonzero determinant. In fact, the submatrix represents the isomorphism between  $\mathcal{V}_1$  and  $\mathbf{T}_A(\mathcal{V})$ , and, by its very construction, is the largest such matrix. Since the determinant of an operator is basis-independent, we have the following proposition:

**3.1.4. Proposition.** *The rank of a matrix is the dimension of the largest (square) submatrix whose determinant is not zero.*

### 3.2 Operations on Matrices

transpose of a matrix There are two basic operations that one can perform on a matrix to obtain a new one; these are transposition and complex conjugation. The **transpose** of an  $M \times N$  matrix  $A$  is an  $N \times M$  matrix  $A^t$  obtained by interchanging the rows and columns of  $A$ :

$$(A^t)_{ij} = (A)_{ji}, \quad \text{or} \quad (\alpha_{ij})^t = (\alpha_{ji}). \tag{3.9}$$

The following theorem, whose proof follows immediately from the definition of transpose, summarizes the important properties of the operation of transposition.

**3.2.1. Theorem.** *Let  $A$  and  $B$  be two (square) matrices. Then*

$$(a) (A + B)^t = A^t + B^t, \quad (b) (AB)^t = B^t A^t, \quad (c) (A^t)^t = A.$$

Of special interest is a matrix that is identical to its transpose. Such matrices occur frequently in physics and are called **symmetric** matrices. Similarly, **antisymmetric** matrices are those satisfying  $A^t = -A$ . Any matrix  $A$  can be written as  $A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$ , where the first term is symmetric and the second is antisymmetric.

The elements of a symmetric matrix  $A$  satisfy the relation  $\alpha_{ji} = (A^t)_{ij} = (A)_{ij} = \alpha_{ij}$ ; i.e., the matrix is symmetric under reflection through the main diagonal. On the other hand, for an antisymmetric matrix we have  $\alpha_{ji} = -\alpha_{ij}$ . In particular, the diagonal elements of an antisymmetric matrix are all zero.

A (real) matrix satisfying  $A^t A = A A^t = 1$  is called **orthogonal**.

**Complex conjugation** is an operation under which all elements of a matrix are complex conjugated. Denoting the complex conjugate of  $A$  by  $A^*$ , we have  $(A^*)_{ij} = (A)_{ij}^*$ , or  $(\alpha_{ij})^* = (\alpha_{ij}^*)$ . A matrix is real if and only if  $A^* = A$ . Clearly,  $(A^*)^* = A$ .

Under the combined operation of complex conjugation and transposition, the rows and columns of a matrix are interchanged and all of its elements are complex conjugated. This combined operation is called the **adjoint** operation, or **hermitian conjugation**, and is denoted by  $\dagger$ , as with operators. Thus, we have

$$A^\dagger = (A^t)^* = (A^*)^t, \tag{3.10}$$

$$(A^\dagger)_{ij} = (A)_{ji}^* \quad \text{or} \quad (\alpha_{ij})^\dagger = (\alpha_{ji}^*).$$

Two types of matrices are important enough to warrant a separate definition.

**3.2.2. Definition.** A **hermitian** matrix  $H$  satisfies  $H^\dagger = H$ , or, in terms of elements,  $\eta_{ij}^* = \eta_{ji}$ . A **unitary** matrix  $U$  satisfies  $U^\dagger U = U U^\dagger = 1$ , or, in terms of elements,  $\sum_{k=1}^N \mu_{ik} \mu_{jk}^* = \sum_{k=1}^N \mu_{ki}^* \mu_{kj} = \delta_{ij}$ .

**Remarks:** It follows immediately from this definition that  
 1. The diagonal elements of a hermitian matrix are real.

2. The  $k$ th column of a hermitian matrix is the complex conjugate of its  $k$ th row, and vice versa.
3. A real hermitian matrix is symmetric.
4. The rows of an  $N \times N$  unitary matrix, when considered as vectors in  $\mathbb{C}^N$ , form an orthonormal set, as do the columns.
5. A real unitary matrix is orthogonal.

diagonal matrices

It is sometimes possible (and desirable) to transform a matrix into a form in which all of its off-diagonal elements are zero. Such a matrix is called a **diagonal matrix**. A diagonal matrix whose diagonal elements are  $\{\lambda_k\}_{k=1}^N$  is denoted by  $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ .

**3.2.3. Example.** In this example, we derive a useful identity for functions of a diagonal matrix. Let  $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  be a diagonal matrix, and  $f(x)$  a function that has a Taylor series expansion  $f(x) = \sum_{k=0}^{\infty} a_k x^k$ . The same function of  $D$  can be written as

$$\begin{aligned} f(D) &= \sum_{k=0}^{\infty} a_k D^k = \sum_{k=0}^{\infty} a_k [\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)]^k = \sum_{k=0}^{\infty} a_k \text{diag}(\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k) \\ &= \text{diag} \left( \sum_{k=0}^{\infty} a_k \lambda_1^k, \sum_{k=0}^{\infty} a_k \lambda_2^k, \dots, \sum_{k=0}^{\infty} a_k \lambda_n^k \right) = \text{diag}(f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)). \end{aligned}$$

In words, the function of a diagonal matrix is equal to a diagonal matrix whose entries are the same function of the corresponding entries of the original matrix. In the above derivation, we used the following obvious properties of diagonal matrices:

$$\begin{aligned} a \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) &= \text{diag}(a\lambda_1, a\lambda_2, \dots, a\lambda_n), \\ \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) + \text{diag}(\omega_1, \omega_2, \dots, \omega_n) &= \text{diag}(\lambda_1 + \omega_1, \dots, \lambda_n + \omega_n), \\ \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \cdot \text{diag}(\omega_1, \omega_2, \dots, \omega_n) &= \text{diag}(\lambda_1 \omega_1, \dots, \lambda_n \omega_n). \quad \blacksquare \end{aligned}$$

**3.2.4. Example.** (a) A prototypical symmetric matrix is that of the moment of inertia encountered in mechanics. The  $ij$ th element of this matrix is defined as  $I_{ij} \equiv \iiint \rho(x_1, x_2, x_3) x_i x_j dV$ , where  $x_i$  is the  $i$ th Cartesian coordinate of a point in the distribution of mass described by the volume density  $\rho(x_1, x_2, x_3)$ . It is clear that  $I_{ij} = I_{ji}$ , or  $I = I^t$ . The moment of inertia matrix can be represented as

$$I = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix}.$$

It has six independent elements.

(b) An example of an antisymmetric matrix is the electromagnetic field tensor given by

$$F = \begin{pmatrix} 0 & -B_3 & B_2 & E_1 \\ B_3 & 0 & -B_1 & E_2 \\ -B_2 & B_1 & 0 & E_3 \\ -E_1 & -E_2 & -E_3 & 0 \end{pmatrix}.$$

Pauli spin matrices

(c) Examples of hermitian matrices are the  $2 \times 2$  **Pauli spin matrices**:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(d) The most frequently encountered orthogonal matrices are rotations. One such matrix represents the rotation of a 3-dimensional rigid body in terms of **Euler angles** and is used in mechanics. Attaching a coordinate system to the body, a general rotation can be decomposed into a rotation of angle  $\varphi$  about the  $z$ -axis, followed by a rotation of angle  $\theta$  about the *new*  $x$ -axis, followed by a rotation of angle  $\psi$  about the *new*  $z$ -axis. We simply exhibit this matrix in terms of these angles and leave it to the reader to show that it is indeed orthogonal.

$$\begin{pmatrix} \cos \psi \cos \varphi - \sin \psi \cos \theta \sin \varphi & -\cos \psi \sin \varphi - \sin \psi \cos \theta \cos \varphi & \sin \psi \sin \theta \\ \sin \psi \cos \varphi + \cos \psi \cos \theta \sin \varphi & -\sin \psi \sin \varphi + \cos \psi \cos \theta \cos \varphi & -\cos \psi \sin \theta \\ \sin \theta \sin \varphi & \sin \theta \cos \varphi & \cos \theta \end{pmatrix}.$$

### 3.3 Orthonormal Bases

Matrix representation is facilitated by choosing an *orthonormal* basis  $B = \{|e_i\rangle\}_{i=1}^N$ . The matrix elements of an operator  $\mathbf{A}$  can be found in such a basis by “multiplying” both sides of  $\mathbf{A}|e_i\rangle = \sum_{k=1}^N \alpha_{ki} |e_k\rangle$  on the left by  $\langle e_j|$ :

$$\langle e_j | \mathbf{A} | e_i \rangle = \langle e_j | \left( \sum_{k=1}^N \alpha_{ki} | e_k \rangle \right) = \sum_{k=1}^N \alpha_{ki} \underbrace{\langle e_j | e_k \rangle}_{=\delta_{jk}} = \alpha_{ji},$$

or

$$(\mathbf{A})_{ij} = \alpha_{ij} = \langle e_i | \mathbf{A} | e_j \rangle. \quad (3.11)$$

We can also show that in an orthonormal basis, the  $i$ th component  $\xi_i$  of a vector is found by multiplying it by  $\langle e_i |$ . This expression for  $\xi_i$  allows us to write the expansion of  $|x\rangle$  as

$$|x\rangle = \sum_{j=1}^N \underbrace{\langle e_j | x \rangle}_{\xi_j} |e_j\rangle = \sum_{j=1}^N |e_j\rangle \langle e_j | x \rangle \Rightarrow \mathbf{1} = \sum_{j=1}^N |e_j\rangle \langle e_j|, \quad (3.12)$$

which is the same as in Proposition 2.5.3. Let us now investigate the representation of the special operators discussed in Chapter 2 and find the connection between those operators and the matrices encountered in the last section. We begin by calculating the matrix representing the hermitian conjugate of an operator  $\mathbf{T}$ . In an orthonormal basis, the elements of this matrix are given by Equation (3.11),  $\tau_{ij} = \langle e_i | \mathbf{T} | e_j \rangle$ . Taking the complex conjugate of this equation and using the definition of  $\mathbf{T}^\dagger$  given in Equation (2.8), we obtain  $\tau_{ij}^* = \langle e_i | \mathbf{T} | e_j \rangle^* = \langle e_j | \mathbf{T}^\dagger | e_i \rangle$ , or  $(\mathbf{T}^\dagger)_{ij} = \tau_{ji}^*$ . This is precisely how the adjoint of a matrix was defined.

Note how crucially this conclusion depends on the orthonormality of the basis vectors. If the basis were not orthonormal, we could not use Equation (3.11) on which the conclusion is based. Therefore,

**3.3.1. Box.** *Only in an orthonormal basis is the adjoint of an operator represented by the adjoint of the matrix representing that operator.*

In particular, a hermitian operator is represented by a hermitian matrix only if an orthonormal basis is used. The following example illustrates this point.

**3.3.2. Example.** Consider the matrix representation of the hermitian operator  $\mathbf{H}$  in a general—not orthonormal—basis  $B = \{|a_i\rangle\}_{i=1}^N$ . The elements of the matrix corresponding to  $\mathbf{H}$  are given by

$$\mathbf{H}|a_k\rangle = \sum_{j=1}^N \eta_{jk} |a_j\rangle, \quad \text{or} \quad \mathbf{H}|a_i\rangle = \sum_{j=1}^N \eta_{ji} |a_j\rangle. \quad (3.13)$$

Taking the product of the first equation with  $\langle a_i|$  and complex-conjugating the result gives  $\langle a_i|\mathbf{H}|a_k\rangle^* = (\sum_{j=1}^N \eta_{jk} \langle a_i|a_j\rangle)^* = \sum_{j=1}^N \eta_{jk}^* \langle a_j|a_i\rangle$ . But by the definition of a hermitian operator,

$$\langle a_i|\mathbf{H}|a_k\rangle^* = \langle a_k|\mathbf{H}^\dagger|a_i\rangle = \langle a_k|\mathbf{H}|a_i\rangle.$$

So we have  $\langle a_k|\mathbf{H}|a_i\rangle = \sum_{j=1}^N \eta_{jk}^* \langle a_j|a_i\rangle$ .

On the other hand, multiplying the second equation in (3.13) by  $\langle a_k|$  gives  $\langle a_k|\mathbf{H}|a_i\rangle = \sum_{j=1}^N \eta_{ji} \langle a_k|a_j\rangle$ . The only conclusion we can draw from this discussion is  $\sum_{j=1}^N \eta_{jk}^* \langle a_j|a_i\rangle = \sum_{j=1}^N \eta_{ji} \langle a_k|a_j\rangle$ . Because this equation does not say anything about each individual  $\eta_{ij}$ , we cannot conclude, in general, that  $\eta_{ij}^* = \eta_{ji}$ . However, if the  $|a_i\rangle$ 's are orthonormal, then  $\langle a_j|a_i\rangle = \delta_{ji}$  and  $\langle a_k|a_j\rangle = \delta_{kj}$ , and we obtain  $\sum_{j=1}^N \eta_{jk}^* \delta_{ji} = \sum_{j=1}^N \eta_{ji} \delta_{kj}$ , or  $\eta_{ik}^* = \eta_{ki}$ , as expected of a hermitian matrix. ■

Similarly, we expect the matrices representing unitary operators to be unitary only if the basis is orthonormal. This is an immediate consequence of Equation (3.10), but we shall prove it in order to provide yet another example of how the completeness relation, Equation (3.12), is used. Since  $\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$ , we have  $\langle e_i|\mathbf{U}\mathbf{U}^\dagger|e_j\rangle = \langle e_i|\mathbf{1}|e_j\rangle = \delta_{ij}$ . We insert the completeness relation  $\mathbf{1} = \sum_{k=1}^N |e_k\rangle\langle e_k|$  between  $\mathbf{U}$  and  $\mathbf{U}^\dagger$  on the LHS:

$$\langle e_i|\mathbf{U}\left(\sum_{k=1}^N |e_k\rangle\langle e_k|\right)\mathbf{U}^\dagger|e_j\rangle = \sum_{k=1}^N \underbrace{\langle e_i|\mathbf{U}|e_k\rangle}_{\equiv \mu_{ik}} \underbrace{\langle e_k|\mathbf{U}^\dagger|e_j\rangle}_{\equiv \mu_{jk}^*} = \delta_{ij}.$$

This equation gives the first half of the requirement for a unitary matrix given in Definition 3.2.2. By redoing the calculation for  $\mathbf{U}^\dagger\mathbf{U}$ , we could obtain the second half of that requirement.

### 3.4 Change of Basis and Similarity Transformation

It is often advantageous to describe a physical problem in a particular basis because it takes a simpler form there, but the general form of the result may still be of importance. In such cases the problem is solved in one basis, and the result is transformed to other bases. Let us investigate this point in some detail.

Given a basis  $B = \{|a_i\rangle\}_{i=1}^N$ , we can write an arbitrary vector  $|a\rangle$  with components  $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$  in  $B$  as  $|a\rangle = \sum_{i=1}^N \alpha_i |a_i\rangle$ . Now suppose that we change the basis to  $B' = \{|a'_j\rangle\}_{j=1}^N$ . How are the components of  $|a\rangle$  in  $B'$  related to those in  $B$ ? To answer this question, we write  $|a_i\rangle$  in terms of  $B'$  vectors,  $|a_i\rangle = \sum_{j=1}^N \rho_{ji} |a'_j\rangle$ , and substitute for  $|a_i\rangle$  in this expansion of  $|a\rangle$ , obtaining  $|a\rangle = \sum_{i=1}^N \alpha_i \sum_{j=1}^N \rho_{ji} |a'_j\rangle = \sum_{i,j} \alpha_i \rho_{ji} |a'_j\rangle$ . If we denote the  $j$ th component of  $|a\rangle$  in  $B'$  by  $\alpha'_j$ , then this equation tells us that

$$\alpha'_j = \sum_{i=1}^N \rho_{ji} \alpha_i \quad \text{for } j = 1, 2, \dots, N. \quad (3.14)$$

If we use  $a'$ ,  $R$ , and  $a$ , respectively, to designate a column vector with elements  $\alpha'_j$ , an  $N \times N$  matrix with elements  $\rho_{ji}$ , and a column vector with elements  $\alpha_i$ , then Equation (3.14) can be written in matrix form as

$$\begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \vdots \\ \alpha'_N \end{pmatrix} = \begin{pmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1N} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2N} \\ \vdots & \vdots & & \vdots \\ \rho_{N1} & \rho_{N2} & \dots & \rho_{NN} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix}, \quad \text{or } a' = Ra. \quad (3.15)$$

basis transformation  
matrix

The matrix  $R$  is called the **basis transformation matrix**. It is invertible because it is a linear transformation that maps one basis onto another (see Theorem 2.1.6).

What happens to a matrix when we transform the basis? Consider the equation  $|b\rangle = A|a\rangle$ , where  $|a\rangle$  and  $|b\rangle$  have components  $\{\alpha_i\}_{i=1}^N$  and  $\{\beta_i\}_{i=1}^N$ , respectively, in  $B$ . This equation has a corresponding matrix equation  $b = Aa$ . Now, if we change the basis, the components of  $|a\rangle$  and  $|b\rangle$  will change to those of  $a'$  and  $b'$ , respectively. We seek a matrix  $A'$  such that  $b' = A'a'$ . This matrix will clearly be the transform of  $A$ . Using Equation (3.15), we write  $Rb = A'Ra$ , or  $b = R^{-1}A'Ra$ . Comparing this with  $b = Aa$  and applying the fact that both equations hold for arbitrary  $a$  and  $b$ , we conclude that

$$R^{-1}A'R = A, \quad \text{or } A' = RAR^{-1}. \quad (3.16)$$

similarity  
transformation

This is called a **similarity transformation** on  $A$ , and  $A'$  is said to be **similar** to  $A$ .

The transformation matrix  $R$  can easily be found for orthonormal bases  $B = \{|e_i\rangle\}_{i=1}^N$  and  $B' = \{|e'_i\rangle\}_{i=1}^N$ . We have  $|e_i\rangle = \sum_{k=1}^N \rho_{ki} |e'_k\rangle$ . Multiplying this

equation by  $\langle e'_j |$ , we obtain

$$\langle e'_j | e_i \rangle = \sum_{k=1}^N \rho_{ki} \langle e'_j | e'_k \rangle = \sum_{k=1}^N \rho_{ki} \delta_{jk} = \rho_{ji}. \tag{3.17}$$

That is,

**3.4.1. Box.** To find the  $ij$ th element of the matrix that changes the components of a vector in the orthonormal basis  $B$  to those of the same vector in the orthonormal basis  $B'$ , take the  $j$ th ket in  $B$  and multiply it by the  $i$ th bra in  $B'$ .

To find the  $ij$ th element of the matrix that changes  $B'$  into  $B$ , we take the  $j$ th ket in  $B'$  and multiply it by the  $i$ th bra in  $B$ :  $\rho'_{ij} = \langle e_i | e'_j \rangle$ . However, the matrix  $R'$  must be  $R^{-1}$ , as can be seen from Equation (3.15). On the other hand,  $(\rho'_{ij})^* = \langle e_i | e'_j \rangle^* = \langle e'_j | e_i \rangle = \rho_{ji}$ , or

$$(R^{-1})_{ij}^* = \rho_{ji}, \quad \text{or} \quad (R^{-1})_{ij} = \rho_{ji}^* = (R^\dagger)_{ij}. \tag{3.18}$$

This shows that  $R$  is a unitary matrix and yields an important result.

**3.4.2. Theorem.** The matrix that transforms one orthonormal basis into another is necessarily unitary.

From Equations (3.17) and (3.18) we have  $(R^\dagger)_{ij} = \langle e_i | e'_j \rangle$ . Thus,

**3.4.3. Box.** To obtain the  $j$ th column of  $R^\dagger$ , we take the  $j$ th vector in the new basis and successively “multiply” it by  $\langle e_i |$  for  $i = 1, 2, \dots, N$ .

In particular, if the original basis is the standard basis of  $\mathbb{C}^N$  and  $|e'_j\rangle$  is represented by a column vector in that basis, then the  $j$ th column of  $R^\dagger$  is simply the vector  $|e'_j\rangle$ .

**3.4.4. Example.** In this example, we show that the similarity transform of a function of a matrix is the same function of the similarity transform of the matrix:  $Rf(A)R^{-1} = f(RAR^{-1})$ . The proof involves inserting  $1 = R^{-1}R$  between factors of  $A$  in the Taylor series expansion of  $f(A)$ :

$$\begin{aligned} Rf(A)R^{-1} &= R \left( \sum_{k=0}^{\infty} a_k A^k \right) R^{-1} = \sum_{k=0}^{\infty} a_k R A^k R^{-1} = \sum_{k=0}^{\infty} a_k \overbrace{R A A \cdots A}^{k \text{ times}} R^{-1} \\ &= \sum_{k=0}^{\infty} a_k \overbrace{R A R^{-1} R A R^{-1} \cdots R A R^{-1}}^{k \text{ times}} = \sum_{k=0}^{\infty} a_k (R A R^{-1})^k = f(R A R^{-1}). \quad \blacksquare \end{aligned}$$



## 3.5 The Determinant

An important concept associated with linear operators is the determinant. Determinants are usually defined in terms of matrix representations of operators in a particular basis. This may give the impression that determinants are basis dependent. However, we shall show that the value of the determinant of an operator is the same in all bases. In fact, it is possible to define determinants of operators without resort to a specific representation of the operator in terms of matrices (see Section 25.3.1).

Let us first introduce a permutation symbol  $\varepsilon_{i_1 i_2 \dots i_N}$ , which will be used extensively in this chapter. It is defined by

$$\varepsilon_{12\dots N} = 1, \quad \text{and} \quad \varepsilon_{i_1 i_2 \dots i_j \dots i_k \dots i_N} = -\varepsilon_{i_1 i_2 \dots i_k \dots i_j \dots i_N}. \quad (3.19)$$

In other words,  $\varepsilon_{i_1 i_2 \dots i_N}$  is completely antisymmetric (or skew-symmetric) under interchange of any pair of its indices. We will use this permutation symbol to define determinants. An immediate consequence of the definition above is that  $\varepsilon_{i_1 i_2 \dots i_N}$  will be zero if any two of its indices are equal. Also note that  $\varepsilon_{i_1 i_2 \dots i_N}$  is  $+1$  if  $(i_1, i_2, \dots, i_N)$  is an even permutation<sup>1</sup> (shuffling) of  $(1, 2, \dots, N)$ , and  $-1$  if it is an odd permutation of  $(1, 2, \dots, N)$ .

### 3.5.1 Determinant of a Matrix

determinant defined

**3.5.1. Definition.** *The determinant is a mapping,  $\det : \mathcal{M}^{N \times N} \rightarrow \mathbb{C}$ , given in term of the elements  $\alpha_{ij}$  of a matrix  $A$  by*

$$\det A = \sum_{i_1, \dots, i_N=1}^N \varepsilon_{i_1 i_2 \dots i_N} \alpha_{i_1 1} \dots \alpha_{i_N N}.$$

Definition 3.5.1 gives  $\det A$  in terms of an expansion in rows, so the first entry is from the first row, the second from the second row, and so on. It is also possible to expand in terms of columns, as the following theorem shows.

**3.5.2. Theorem.** *The determinant of a matrix  $A$  can be written as*

$$\det A = \sum_{i_1, \dots, i_N}^N \varepsilon_{i_1 i_2 \dots i_N} \alpha_{i_1 1} \dots \alpha_{i_N N}.$$

Therefore,  $\det A = \det A^t$ .

<sup>1</sup>An even permutation means an even number of exchanges of pairs of elements. Thus,  $(2, 3, 1)$  is an even permutation of  $(1, 2, 3)$ , while  $(2, 1, 3)$  is an odd permutation. It can be shown (see Chapter 23) that the parity (evenness or oddness) of a permutation is well-defined; i.e., that although there may be many routes of reaching a permutation from a given (fiducial) permutation via exchanges of pairs of elements, all such routes require either even numbers of exchanges or odd numbers of exchanges.

*Proof.* We shall go into some detail in the proof of only this theorem on determinants to illustrate the manipulations involved in working with the  $\varepsilon$  symbol. We shall not reproduce such details for the other theorems.

In the equation of Definition 3.5.1,  $i_1, i_2, \dots, i_N$  are all different and form a permutation of  $(1, 2, \dots, N)$ . So one of the  $\alpha$ 's must have 1 as its second index. We assume that it is the  $j_1$ th term; that is,  $i_{j_1} = 1$ , and

$$\alpha_{1i_1} \dots \alpha_{j_1 i_{j_1}} \dots \alpha_{Ni_N} = \alpha_{1i_1} \dots \alpha_{j_1 1} \dots \alpha_{Ni_N}$$

We move this term all the way to the left to get  $\alpha_{j_1 1} \alpha_{1i_1} \dots \alpha_{Ni_N}$ . Now we look for the entry with 2 as the second index and assume that it occurs at the  $j_2$ th position; that is,  $i_{j_2} = 2$ . We move this to the left, next to  $\alpha_{j_1 1}$ , and write  $\alpha_{j_1 1} \alpha_{j_2 2} \alpha_{1i_1} \dots \alpha_{Ni_N}$ . We continue in this fashion until we get  $\alpha_{j_1 1} \alpha_{j_2 2} \dots \alpha_{j_N N}$ . Since  $j_1, j_2, \dots, j_N$  is really a reshuffling of  $i_1, i_2, \dots, i_N$ , the summation indices can be changed to  $j_1, j_2, \dots, j_N$ , and we can write<sup>2</sup>

$$\det A = \sum_{j_1, \dots, j_N=1}^N \varepsilon_{i_1 i_2 \dots i_N} \alpha_{j_1 1} \dots \alpha_{j_N N}.$$

If we can show that  $\varepsilon_{i_1 i_2 \dots i_N} = \varepsilon_{j_1 j_2 \dots j_N}$ , we are done. In the equation of the theorem, the sequence of integers  $(i_1, i_2, \dots, i_N)$  is obtained by some shuffling of  $(1, 2, \dots, N)$ . What we have done just now is to reshuffle  $(i_1, i_2, \dots, i_N)$  in reverse order to get back to  $(1, 2, \dots, N)$ . Thus, if the shuffling in the equation of the theorem is even (odd), the reshuffling will also be even (odd). Thus,  $\varepsilon_{i_1 i_2 \dots i_N} = \varepsilon_{j_1 j_2 \dots j_N}$ , and we obtain the first part of the theorem:

$$\det A = \sum_{j_1, \dots, j_N=1}^N \varepsilon_{j_1 j_2 \dots j_N} \alpha_{j_1 1} \dots \alpha_{j_N N}.$$

For the second part, we simply note that the rows of  $A^t$  are columns of  $A$  and vice versa.  $\square$

**3.5.3. Theorem.** *Interchanging two rows (or two columns) of a matrix changes the sign of its determinant.*

*Proof.* The proof is a simple exercise in permutation left for the reader.  $\square$

An immediate consequence of this theorem is the following corollary.

**3.5.4. Corollary.** *The determinant of a matrix with two equal rows (or two equal columns) is zero. Therefore, one can add a multiple of a row (column) to another row (column) of a matrix without changing its determinant.*

<sup>2</sup>The  $\varepsilon$  symbol in the sum is not independent of the  $j$ 's, although it appears without such indices. In reality, the  $i$  indices are "functions" of the  $j$  indices.

Since every term of the determinant in Definition 3.5.1 contains one and only one element from each row, we can write

$$\det A = \alpha_{i1}A_{i1} + \alpha_{i2}A_{i2} + \cdots + \alpha_{iN}A_{iN} = \sum_{j=1}^N \alpha_{ij}A_{ij},$$

cofactor of an  
element of a matrix

where  $A_{ij}$  contains products of elements of the matrix  $A$  other than the element  $\alpha_{ij}$ . Since each element of a row or column occurs at most once in each term of the expansion,  $A_{ij}$  cannot contain any element from the  $i$ th row or the  $j$ th column. The quantity  $A_{ij}$  is called the **cofactor of  $\alpha_{ij}$** , and the above expression is known as the (Laplace) *expansion of  $\det A$  by its  $i$ th row*. Clearly, there is a similar expansion by the  $i$ th column of the determinant, which is obtained by a similar argument using the equation of Theorem 3.5.2. We collect both results in the following equation.

$$\det A = \sum_{j=1}^N \alpha_{ij}A_{ij} = \sum_{j=1}^N \alpha_{ji}A_{ji}. \quad (3.20)$$

**Vandermonde, Alexandre-Thiéophile**, also known as Alexis, Abnit, and Charles-Auguste Vandermonde (1735–1796) had a father, a physician who directed his sickly son toward a musical career. An acquaintanceship with Fontaine, however, so stimulated Vandermonde that in 1771 he was elected to the Académie des Sciences, to which he presented four mathematical papers (his total mathematical production) in 1771–1772. Later, Vandermonde wrote several papers on harmony, and it was said at that time that musicians considered Vandermonde to be a mathematician and that mathematicians viewed him as a musician.

Vandermonde's membership in the Academy led to a paper on experiments with cold, made with Bezout and Lavoisier in 1776, and a paper on the manufacture of steel with Berthollet and Monge in 1786. Vandermonde became an ardent and active revolutionary, being such a close friend of Monge that he was termed "femme de Monge." He was a member of the Commune of Paris and the club of the Jacobins. In 1782 he was director of the Conservatoire des Arts et Métiers and in 1792, chief of the Bureau de l'Habillement des Armées. He joined in the design of a course in political economy for the École Normale and in 1795 was named a member of the Institut National.

Vandermonde is best known for the theory of determinants. Lebesgue believed that the attribution of determinant to Vandermonde was due to a misreading of his notation. Nevertheless, Vandermonde's fourth paper was the first to give a connected exposition of determinants, because he (1) defined a contemporary symbolism that was more complete, simple, and appropriate than that of Leibniz; (2) defined determinants as functions apart from the solution of linear equations presented by Cramer but also treated by Vandermonde; and (3) gave a number of properties of these functions, such as the number and signs of the terms and the effect of interchanging two consecutive indices (rows or columns), which he used to show that a determinant is zero if two rows or columns are identical.

Vandermonde's real and unrecognized claim to fame was lodged in his first paper, in which he approached the general problem of the solvability of algebraic equations through a study of functions invariant under permutations of the roots of the equations. Cauchy

assigned priority in this to Lagrange and Vandermonde. Vandermonde read his paper in November 1770, but he did not become a member of the Academy until 1771, and the paper was not published until 1774. Although Vandermonde's methods were close to those later developed by Abel and Galois for testing the solvability of equations, and although his treatment of the binomial equation  $x^n - 1 = 0$  could easily have led to the anticipation of Gauss's results on constructible polygons, Vandermonde himself did not rigorously or completely establish his results, nor did he see the implications for geometry. Nevertheless, Kronecker dates the modern movement in algebra to Vandermonde's 1770 paper.

Unfortunately, Vandermonde's spurt of enthusiasm and creativity, which in two years produced four insightful mathematical papers at least two of which were of substantial importance, was quickly diverted by the exciting politics of the time and perhaps by poor health.

**3.5.5. Proposition.** *If  $i \neq k$ , then  $\sum_{j=1}^N \alpha_{ij} A_{kj} = 0 = \sum_{j=1}^N \alpha_{ji} A_{jk}$ .*

*Proof.* Consider the matrix  $B$  obtained from  $A$  by replacing row  $k$  by row  $i$  (row  $i$  remains unchanged, of course). The matrix  $B$  has two equal rows, and its determinant is therefore zero. Now, if we expand  $\det B$  by its  $k$ th row according to Equation (3.20), we obtain  $0 = \det B = \sum_{j=1}^N \beta_{kj} B_{kj}$ . But the elements of the  $k$ th row of  $B$  are the elements of the  $i$ th row of  $A$ ; that is,  $\beta_{kj} = \alpha_{ij}$ , and the cofactors of the  $k$ th row of  $B$  are the same as those of  $A$ , that is,  $B_{kj} = A_{kj}$ . Thus, the first equation of the proposition is established. The second equation can be established using expansion by columns. □

minor of a matrix

A **minor of order  $N - 1$**  of an  $N \times N$  matrix  $A$  is the determinant of a matrix obtained by striking out one row and one column of  $A$ . If we strike out the  $i$ th row and  $j$ th column of  $A$ , then the minor is denoted by  $M_{ij}$ .

**3.5.6. Theorem.**  $A_{ij} = (-1)^{i+j} M_{ij}$ .

*Proof.* The proof involves separating  $\alpha_{11}$  from the rest of the terms in the expansion of the determinant. The unique coefficient of  $\alpha_{11}$  is  $A_{11}$  by Equation (3.20). We can show that it is also  $M_{11}$  by examining the  $\varepsilon$  expansion of the determinant and performing the first sum. This will establish the equality  $A_{11} = M_{11}$ . The general equality is obtained by performing enough interchanges of rows and columns of the matrix to bring  $\alpha_{ij}$  into the first-row first-column position, each exchange introducing a negative sign, thus the  $(-1)^{i+j}$  factor. The details are left as an exercise. □

The combination of Equation (3.20) and Theorem 3.5.6 gives the familiar routine of evaluating the determinant of a matrix.

### 3.5.2 Determinants of Products of Matrices

One extremely useful property of determinants is expressed in the following theorem.

**3.5.7. Theorem.**  $\det(AB) = (\det A)(\det B)$ 

*Proof.* The proof consists in keeping track of index shuffling while rearranging the order of products of matrix elements. We shall leave the details as an exercise.  $\square$

**3.5.8. Example.** Let  $O$  and  $U$  denote, respectively, an orthogonal and a unitary  $n \times n$  matrix; that is,  $OO^t = O^tO = 1$ , and  $UU^\dagger = U^\dagger U = 1$ . Taking the determinant of the first equation and using Theorems 3.5.2 and 3.5.7, we obtain  $(\det O)(\det O^t) = (\det O)^2 = \det 1 = 1$ . Therefore, for an orthogonal matrix, we get  $\det O = \pm 1$ .

Orthogonal transformations preserve a real inner product. Among such transformations are the so-called inversions, which, in their simplest form, multiply a vector by  $-1$ . In three dimensions this corresponds to a reflection through the origin. The matrix associated with this operation is  $-1$ :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

which has a determinant of  $-1$ . This is a prototype of other, more complicated, orthogonal transformations whose determinants are  $-1$ .

The other orthogonal transformations, whose determinants are  $+1$ , are of special interest because they correspond to rotations in three dimensions. The set of orthogonal transformations in  $n$  dimensions having the determinant  $+1$  is denoted by  $SO(n)$ . These transformations are special because they have the mathematical structure of a (continuous) group, which finds application in many areas of advanced physics. We shall come back to the topic of group theory later in the book.

We can obtain a similar result for unitary transformations. We take the determinant of both sides of  $U^\dagger U = 1$ :

$$\det(U^*)^t \det U = \det U^* \det U = (\det U)^*(\det U) = |\det U|^2 = 1.$$

Thus, we can generally write  $\det U = e^{i\alpha}$ , with  $\alpha \in \mathbb{R}$ . The set of those transformations with  $\alpha = 0$  forms a group to which  $1$  belongs and that is denoted by  $SU(n)$ . This group has found applications in the attempts at unifying the fundamental interactions.  $\blacksquare$

**3.5.3 Inverse of a Matrix**

One of the most useful properties of the determinant is the simple criterion it gives for a matrix to be invertible. We are ready to investigate this criterion now. We combine Equation (3.20) and the content of Proposition 3.5.5 into a single equation,

$$\sum_{j=1}^N \alpha_{ij} A_{kj} = (\det A) \delta_{ik} = \sum_{j=1}^N \alpha_{ji} A_{jk}, \quad (3.21)$$

and construct a matrix  $C_A$ , whose elements are the cofactors of the elements of the matrix  $A$ :

$$C_A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ \vdots & \vdots & & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NN} \end{pmatrix}. \quad (3.22)$$

Then Equation (3.21) can be written as

$$\sum_{j=1}^N \alpha_{ij} (C_A^t)_{jk} = (\det A) \delta_{ik} = \sum_{j=1}^N (C_A^t)_{kj} \alpha_{ji},$$

or, in matrix form, as

$$AC_A^t = (\det A)1 = C_A^t A. \quad (3.23)$$

**3.5.9. Theorem.** *The inverse of a matrix (if it exists) is unique. The matrix  $A$  has an inverse if and only if  $\det A \neq 0$ . Furthermore,*

$$A^{-1} = \frac{C_A^t}{\det A}, \quad (3.24)$$

where  $C_A$  is the matrix of the cofactors of  $A$ .

*Proof.* Let  $B$  and  $C$  be inverses of  $A$ . Then

$$B = \underbrace{(CA)}_{=1} B = C \underbrace{(AB)}_{=1} = C.$$

For the second part, we note that if  $A$  has an inverse  $B$ , then

$$AB = 1 \Rightarrow \det A \det B = \det 1 = 1,$$

whence  $\det A \neq 0$ . Conversely, if  $\det A \neq 0$ , then dividing both sides of Equation (3.23) by  $\det A$ , we obtain the unique inverse (3.24) of  $A$ .  $\square$

The inverse of a  $2 \times 2$  matrix is easily found:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad (3.25)$$

if  $ad - bc \neq 0$ . There is a more practical way of calculating the inverse of matrices. In the following discussion of this method, we shall confine ourselves simply to stating a couple of definitions and the main theorem, with no attempt at providing any proofs. The practical utility of the method will be illustrated by a detailed analysis of examples.

elementary row  
operation

**3.5.10. Definition.** An *elementary row operation* on a matrix is one of the following: (a) interchange of two rows of the matrix, (b) multiplication of a row by a nonzero number, and (c) addition of a multiple of one row to another.

Elementary column operations are defined analogously.

triangular, or  
row-echelon form of  
a matrix

**3.5.11. Definition.** A matrix is in *triangular*, or *row-echelon*, form if it satisfies the following three conditions:

1. Any row consisting of only zeros is below any row that contains at least one nonzero element.
2. Going from left to right, the first nonzero entry of any row is to the left of the first nonzero entry of any lower row.
3. The first nonzero entry of each row is 1.

**3.5.12. Theorem.** For any invertible  $n \times n$  matrix  $A$ ,

- The  $n \times 2n$  matrix  $(A|1)$  can be transformed into the  $n \times 2n$  matrix  $(1|A^{-1})$  by means of a finite number of elementary row operations.<sup>3</sup>
- If  $(A|1)$  is transformed into  $(1|B)$  by means of elementary row operations, then  $B = A^{-1}$ .

A systematic way of transforming  $(A|1)$  into  $(1|A^{-1})$  is first to bring  $A$  into triangular form and then eliminate all nonzero elements of each column by elementary row operations.

**3.5.13. Example.** Let us evaluate the inverse of

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 2 & 1 & -1 \end{pmatrix}.$$

We start with

$$\left( \begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 2 & 1 & -1 & 0 & 0 & 1 \end{array} \right) \equiv M$$

and apply elementary row operations to  $M$  to bring the left half of it into triangular form. If we denote the  $k$ th row by  $(k)$  and the three operations of Definition 3.5.10, respectively, by  $(k) \leftrightarrow (j)$ ,  $\alpha(k)$ , and  $\alpha(k) + (j)$ , we get

$$\begin{aligned} M &\xrightarrow{-2(1)+(3)} \left( \begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & -3 & 1 & -2 & 0 & 1 \end{array} \right) \xrightarrow{3(2)+(3)} \left( \begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & -5 & -2 & 3 & 1 \end{array} \right) \\ &\xrightarrow{-\frac{1}{5}(3)} \left( \begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2/5 & -3/5 & -1/5 \end{array} \right) \equiv M' \end{aligned}$$

<sup>3</sup>The matrix  $(A|1)$  denotes the  $n \times 2n$  matrix obtained by juxtaposing the  $n \times n$  unit matrix to the right of  $A$ . It can easily be shown that if  $A$ ,  $B$ , and  $C$  are  $n \times n$  matrices, then  $A(B|C) = (AB|AC)$ .

The left half of  $M'$  is in triangular form. However, we want all entries above any 1 in a column to be zero as well, i.e., we want the left-hand matrix to be 1. We can do this by appropriate use of type 3 elementary row operations:

$$\begin{aligned} M' &\xrightarrow{-2(2)+(1)} \left( \begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & -2 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2/5 & -3/5 & -1/5 \end{array} \right) \\ &\xrightarrow{-3(3)+(1)} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -1/5 & -1/5 & 3/5 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2/5 & -3/5 & -1/5 \end{array} \right) \\ &\xrightarrow{2(3)+(2)} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -1/5 & -1/5 & 3/5 \\ 0 & 1 & 0 & 4/5 & -1/5 & -2/5 \\ 0 & 0 & 1 & 2/5 & -3/5 & -1/5 \end{array} \right) \end{aligned}$$

The right half of the resulting matrix is  $A^{-1}$ . ■

**3.5.14. Example.** It is instructive to start with a matrix that is not invertible and show that it is impossible to turn it into 1 by elementary row operations. Consider the matrix

$$B = \begin{pmatrix} 2 & -1 & 3 \\ 1 & -2 & 1 \\ -1 & 5 & 0 \end{pmatrix}.$$

Let us systematically bring it into triangular form:

$$\begin{aligned} M &= \left( \begin{array}{ccc|ccc} 2 & -1 & 3 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ -1 & 5 & 0 & 0 & 0 & 1 \end{array} \right) \xrightarrow{(1) \leftrightarrow (2)} \left( \begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 2 & -1 & 3 & 1 & 0 & 0 \\ -1 & 5 & 0 & 0 & 0 & 1 \end{array} \right) \\ &\xrightarrow{-2(1)+(2)} \left( \begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 3 & 1 & 1 & -2 & 0 \\ -1 & 5 & 0 & 0 & 0 & 1 \end{array} \right) \xrightarrow{(1)+(3)} \left( \begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 3 & 1 & 1 & -2 & 0 \\ 0 & 3 & 1 & 0 & 1 & 1 \end{array} \right) \\ &\xrightarrow{-(2)+(3)} \left( \begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 3 & 1 & 1 & -2 & 0 \\ 0 & 0 & 0 & -1 & 3 & 1 \end{array} \right) \xrightarrow{\frac{1}{3}(2)} \left( \begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1/3 & 1/3 & -2/3 & 0 \\ 0 & 0 & 0 & -1 & 3 & 1 \end{array} \right) \end{aligned}$$

The matrix  $B$  is now in triangular form, but its third row contains all zeros. There is no way we can bring this into the form of a unit matrix. We therefore conclude that  $B$  is not invertible. This is, of course, obvious, since it can easily be verified that  $B$  has a vanishing determinant. ■

We mentioned earlier that the determinant is a property of linear transformations, although they are defined in terms of matrices that represent them. We can now show this. First, note that taking the determinant of both sides of  $AA^{-1} = 1$ ,



one obtains  $\det(A^{-1}) = 1/\det A$ . Now recall that the representations of an operator in two different bases are related via a similarity transformation. Thus, if  $\mathbf{A}$  is represented by  $\mathbf{A}$  in one basis and by  $\mathbf{A}'$  in another, then there exists an invertible matrix  $\mathbf{R}$  such that  $\mathbf{A}' = \mathbf{R}\mathbf{A}\mathbf{R}^{-1}$ . Taking the determinant of both sides, we get

$$\det \mathbf{A}' = \det \mathbf{R} \det \mathbf{A} \frac{1}{\det \mathbf{R}} = \det \mathbf{A}.$$

Thus, the determinant is an intrinsic property of the operator, independent of the basis chosen in which to represent the operator.

### 3.6 The Trace

Another intrinsic quantity associated with an operator that is usually defined in terms of matrices is given in the following definition.

**3.6.1. Definition.** Let  $\mathbf{A}$  be an  $N \times N$  matrix. The mapping  $\text{tr} : \mathcal{M}^{N \times N} \rightarrow \mathbb{C}$  (or  $\mathbb{R}$ ) given by  $\text{tr} \mathbf{A} = \sum_{i=1}^N \alpha_{ii}$  is called the **trace** of  $\mathbf{A}$ .

trace of a square matrix

**3.6.2. Theorem.** The trace is a linear mapping. Furthermore,

$$\text{tr} \mathbf{A}' = \text{tr} \mathbf{A} \quad \text{and} \quad \text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA}).$$

*Proof.* The linearity of the trace and the first identity follow directly from the definition. To prove the second identity of the theorem, we use the definitions of the trace and the matrix product:

$$\begin{aligned} \text{tr}(\mathbf{AB}) &= \sum_{i=1}^N (\mathbf{AB})_{ii} = \sum_{i=1}^N \sum_{j=1}^N (\mathbf{A})_{ij} (\mathbf{B})_{ji} = \sum_{i=1}^N \sum_{j=1}^N (\mathbf{B})_{ji} (\mathbf{A})_{ij} \\ &= \sum_{j=1}^N \left( \sum_{i=1}^N (\mathbf{B})_{ji} (\mathbf{A})_{ij} \right) = \sum_{j=1}^N (\mathbf{BA})_{jj} = \text{tr}(\mathbf{BA}). \end{aligned}$$

□

connection between trace and determinant

**3.6.3. Example.** In this example, we show a very useful connection between the trace and the determinant that holds when a matrix is only infinitesimally different from the unit matrix. Let us calculate the determinant of  $\mathbf{1} + \epsilon \mathbf{A}$  to first order in  $\epsilon$ . Using the definition of determinant, we write

$$\begin{aligned} \det(\mathbf{1} + \epsilon \mathbf{A}) &= \sum_{i_1, \dots, i_n=1}^n \epsilon_{i_1 \dots i_n} (\delta_{1i_1} + \epsilon \alpha_{1i_1}) \dots (\delta_{ni_n} + \epsilon \alpha_{ni_n}) \\ &= \sum_{i_1, \dots, i_n=1}^n \epsilon_{i_1 \dots i_n} \delta_{1i_1} \dots \delta_{ni_n} + \epsilon \sum_{k=1}^n \sum_{i_1, \dots, i_n=1}^n \epsilon_{i_1 \dots i_n} \delta_{1i_1} \dots \delta_{ki_k} \dots \delta_{ni_n} \alpha_{ki_k}. \end{aligned}$$

The first sum is just the product of all the Kronecker deltas. In the second sum,  $\hat{\delta}_{ki_k}$  means that in the product of the deltas,  $\delta_{ki_k}$  is absent. This term is obtained by multiplying the second term of the  $k$ th parentheses by the first term of all the rest. Since we are interested only in the first power of  $\epsilon$ , we stop at this term. Now, the first sum is reduced to  $\epsilon_{12\dots n} = 1$  after all the Kronecker deltas are summed over. For the second sum, we get

$$\begin{aligned} \epsilon \sum_{k=1}^n \sum_{i_1, \dots, i_n=1}^n \epsilon_{i_1 \dots i_n} \delta_{1i_1} \dots \hat{\delta}_{ki_k} \dots \delta_{ni_n} \alpha_{ki_k} &= \epsilon \sum_{k=1}^n \sum_{i_k=1}^n \epsilon_{12\dots i_k \dots n} \alpha_{ki_k} \\ &= \epsilon \sum_{k=1}^n \epsilon_{12\dots k \dots n} \alpha_{kk} = \epsilon \sum_{k=1}^n \alpha_{kk} = \epsilon \operatorname{tr} A, \end{aligned} \tag{3.26}$$

where the last line follows from the fact that the only nonzero value for  $\epsilon_{12\dots i_k \dots n}$  is obtained when  $i_k$  is equal to the missing index, i.e.,  $k$ , in which case it will be 1. Thus  $\det(1 + \epsilon A) = 1 + \epsilon \operatorname{tr} A$ . ■

Similar matrices have the same trace: If  $A' = RAR^{-1}$ , then

$$\begin{aligned} \operatorname{tr} A' &= \operatorname{tr}(RAR^{-1}) = \operatorname{tr}[R(AR^{-1})] = \operatorname{tr}[(AR^{-1})R] \\ &= \operatorname{tr}[A(R^{-1}R)] = \operatorname{tr}(A1) = \operatorname{tr} A. \end{aligned}$$

The preceding discussion is summarized in the following proposition.

**3.6.4. Proposition.** *To every operator  $A \in \mathcal{L}(\mathcal{V})$  are associated two intrinsic numbers,  $\det A$  and  $\operatorname{tr} A$ , which are the determinant and trace of the matrix representation of the operator in any basis of  $\mathcal{V}$ .*

It follows from this proposition that the result of Example 3.6.3 can be written in terms of operators:

$$\det(1 + \epsilon A) = 1 + \epsilon \operatorname{tr} A. \tag{3.27}$$

A particularly useful formula that can be derived from this equation is the derivative at  $t = 0$  of an operator  $A(t)$  depending on a single variable with the property that  $A(0) = 1$ . To first order in  $t$ , we can write  $A(t) = 1 + t\dot{A}(0)$  where a dot represents differentiating with respect to  $t$ . Substituting this in Equation (3.27) and differentiating with respect to  $t$ , we obtain the important result

$$\left. \frac{d}{dt} \det(A(t)) \right|_{t=0} = \operatorname{tr} \dot{A}(0). \tag{3.28}$$

**3.6.5. Example.** We have seen that the determinant of a *product* of matrices is the product of the determinants. On the other hand, the trace of a *sum* of matrices is the sum of traces. When dealing with numbers, products and sums are related via the logarithm and exponential:  $\alpha\beta = \exp\{\ln \alpha + \ln \beta\}$ . A generalization of this relation exists for diagonalizable matrices. Let  $A$  be such a matrix, i.e., let  $D = RAR^{-1}$  for some similarity transformation  $R$  and some diagonal matrix  $D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ . The determinant of a diagonal matrix is simply the product of its elements:

$$\det D = \lambda_1 \lambda_2 \dots \lambda_n.$$

Taking the natural log of both sides and using the result of Example 3.2.3, we have

$$\ln(\det D) = \ln \lambda_1 + \ln \lambda_2 + \cdots + \ln \lambda_n = \operatorname{tr}(\ln D),$$

which can also be written as  $\det D = \exp\{\operatorname{tr}(\ln D)\}$ .

In terms of  $A$ , this reads  $\det(RAR^{-1}) = \exp\{\operatorname{tr}(\ln(RAR^{-1}))\}$ . Now invoke the invariance of determinant and trace under similarity transformation and the result of Example 3.4.4 to obtain

$$\det A = \exp\{\operatorname{tr}(R(\ln A)R^{-1})\} = \exp\{\operatorname{tr}(\ln A)\}. \quad (3.29)$$

This is an important equation, which is sometimes used to define the determinant of operators in infinite-dimensional vector spaces. ■

Both the determinant and the trace are mappings from  $\mathcal{M}^{N \times N}$  to  $\mathbb{C}$ . The determinant is not a linear mapping, but the trace is; and this opens up the possibility of defining an inner product in the vector space of  $N \times N$  matrices in terms of the trace:

**3.6.6. Proposition.** *For any two matrices  $A, B \in \mathcal{M}^{N \times N}$ , the mapping  $g : \mathcal{M}^{N \times N} \times \mathcal{M}^{N \times N} \rightarrow \mathbb{C}$  defined by  $g(A, B) = \operatorname{tr}(A^\dagger B)$  is a sesquilinear inner product.*

*Proof.* The proof follows directly from the linearity of trace and the definition of hermitian conjugate. □

## 3.7 Problems

**3.1.** Show that if  $|c\rangle = |a\rangle + |b\rangle$ , then in any basis the components of  $|c\rangle$  are equal to the sums of the corresponding components of  $|a\rangle$  and  $|b\rangle$ . Also show that the elements of the matrix representing the sum of two operators are the sums of the elements of the matrices representing those two operators.

**3.2.** Show that the unit operator  $\mathbf{1}$  is represented by the unit matrix in any basis.

**3.3.** The linear operator  $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  is given by

$$\mathbf{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y - 3z \\ x + y - z \end{pmatrix}.$$

Construct the matrix representing  $\mathbf{A}$  in the standard bases of  $\mathbb{R}^3$  and  $\mathbb{R}^2$ .

**3.4.** The linear transformation  $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  is defined as

$$\mathbf{T}(x_1, x_2, x_3) = (x_1 + x_2 - x_3, 2x_1 - x_3, x_1 + 2x_2).$$

Find the matrix representation of  $\mathbf{T}$  in

- (a) the standard basis of  $\mathbb{R}^3$ ,
- (b) the basis consisting of  $|a_1\rangle = (1, 1, 0)$ ,  $|a_2\rangle = (1, 0, -1)$ , and  $|a_3\rangle = (0, 2, 3)$ .

**3.5.** Show that the diagonal elements of an antisymmetric matrix are all zero.

**3.6.** Show that the number of independent *real* parameters for an  $N \times N$

1. (real) symmetric matrix is  $N(N + 1)/2$ ,
2. (real) antisymmetric matrix is  $N(N - 1)/2$ ,
3. (real) orthogonal matrix is  $N(N - 1)/2$ ,
4. (complex) unitary matrix is  $N^2$ ,
5. (complex) hermitian matrix is  $N^2$ .

**3.7.** Show that an arbitrary orthogonal  $2 \times 2$  matrix can be written in one of the following two forms:

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

The first is a pure rotation (its determinant is  $+1$ ), and the second has determinant  $-1$ . The form of the choices is dictated by the assumption that the first entry of the matrix reduces to 1 when  $\theta = 0$ .

**3.8.** Derive the formulas

$$\begin{aligned} \cos(\theta_1 + \theta_2) &= \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2, \\ \sin(\theta_1 + \theta_2) &= \sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2 \end{aligned}$$

by noting that the rotation of the angle  $\theta_1 + \theta_2$  in the  $xy$ -plane is the product of two rotations. (See Problem 3.7.)

**3.9.** Prove that if a matrix  $M$  satisfies  $MM^\dagger = 0$ , then  $M = 0$ . Note that in general,  $M^2 = 0$  does not imply that  $M$  is zero. Find a nonzero  $2 \times 2$  matrix whose square is zero.

**3.10.** Construct the matrix representations of

$$\mathbf{D} : \mathcal{P}_4^c[t] \rightarrow \mathcal{P}_4^c[t] \quad \text{and} \quad \mathbf{T} : \mathcal{P}_3^c[t] \rightarrow \mathcal{P}_4^c[t],$$

the derivative and multiplication-by- $t$  operators. Choose  $\{1, t, t^2, t^3\}$  as your basis of  $\mathcal{P}_3^c[t]$  and  $\{1, t, t^2, t^3, t^4\}$  as your basis of  $\mathcal{P}_4^c[t]$ . Use the matrix of  $\mathbf{D}$  so obtained to find the first, second, third, fourth, and fifth derivatives of a general polynomial of degree 4.

**3.11.** Find the transformation matrix  $R$  that relates the (orthonormal) standard basis of  $\mathbb{C}^3$  to the orthonormal basis obtained from the following vectors via the Gram-Schmidt process:

$$|a_1\rangle = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad |a_2\rangle = \begin{pmatrix} 0 \\ 1 \\ -i \end{pmatrix}, \quad |a_3\rangle = \begin{pmatrix} i \\ 0 \\ -1 \end{pmatrix}.$$

Verify that  $R$  is unitary, as expected from Theorem 3.4.2.

**3.12.** If the matrix representation of an endomorphism  $T$  of  $\mathbb{C}^2$  with respect to the standard basis is  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ , what is its matrix representation with respect to the basis  $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ ?

**3.13.** If the matrix representation of an endomorphism  $T$  of  $\mathbb{C}^3$  with respect to the standard basis is

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$$

what is the representation of  $T$  with respect to the basis

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}?$$

**3.14.** Using Definition 3.5.1, calculate the determinant of a general  $3 \times 3$  matrix and obtain the familiar expansion of such a determinant in terms of the first row of the matrix.

**3.15.** Prove Corollary 3.5.4.

**3.16.** Show that  $\det(\alpha A) = \alpha^N \det A$  for an  $N \times N$  matrix  $A$  and a complex number  $\alpha$ .

**3.17.** Show that  $\det 1 = 1$  for any unit matrix.

**3.18.** Find a specific pair of matrices  $A$  and  $B$  such that  $\det(A+B) \neq \det A + \det B$ . Therefore, the determinant is *not* a linear mapping. Hint: *Any* pair of matrices will most likely work. In fact, the challenge is to find a pair such that  $\det(A+B) = \det A + \det B$ .

**3.19.** Demonstrate Proposition 3.5.5 using an arbitrary  $3 \times 3$  matrix and evaluating the sum explicitly.

**3.20.** Find the inverse of the matrix

$$A = \begin{pmatrix} 3 & -1 & 2 \\ 1 & 0 & -3 \\ -2 & 1 & -1 \end{pmatrix}.$$

**3.21.** Show explicitly that  $\det(AB) = \det A \det B$  for  $2 \times 2$  matrices.

**3.22.** Given three  $N \times N$  matrices  $A$ ,  $B$ , and  $C$  such that  $AB = C$  with  $C$  invertible, show that both  $A$  and  $B$  must be invertible. Thus, any two operators  $A$  and  $B$  on a finite-dimensional vector space satisfying  $AB = \mathbf{1}$  are invertible and each is the inverse of the other. Note: This is not true for infinite-dimensional vector spaces.

**3.23.** Show directly that the similarity transformation induced by  $R$  does not change the determinant or the trace of  $A$  where

$$R = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 2 & 1 & -1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & -2 \\ 1 & -3 & -1 \end{pmatrix}.$$

**3.24.** Find the matrix that transforms the standard basis of  $\mathbb{C}^3$  to the vectors

$$|a_1\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} \\ \frac{1+i}{\sqrt{6}} \end{pmatrix}, \quad |a_2\rangle = \begin{pmatrix} \frac{-i}{\sqrt{2}} \\ \frac{i}{\sqrt{6}} \\ \frac{-1+i}{\sqrt{6}} \end{pmatrix}, \quad |a_3\rangle = \begin{pmatrix} 0 \\ \frac{-2}{\sqrt{6}} \\ \frac{1+i}{\sqrt{6}} \end{pmatrix}.$$

Show that this matrix is unitary.

**3.25.** Consider the three operators  $L_1$ ,  $L_2$ , and  $L_3$  satisfying  $[L_1, L_2] = iL_3$ ,  $[L_3, L_1] = iL_2$ ,  $[L_2, L_3] = iL_1$ . Show that the trace of each of these operators is necessarily zero.

**3.26.** Show that in the expansion of the determinant given in Definition 3.5.1, no two elements of the same row or the same column can appear in each term of the sum.

**3.27.** Find inverses for the following matrices using both methods discussed in this chapter.

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & 2 \\ -1 & -2 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 2 & 1 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & -1 & -2 \end{pmatrix},$$

$$D = \begin{pmatrix} 1/\sqrt{2} & 0 & (1-i)/(2\sqrt{2}) & (1+i)/(2\sqrt{2}) \\ 0 & 1/\sqrt{2} & (1-i)/(2\sqrt{2}) & -(1+i)/(2\sqrt{2}) \\ 1/\sqrt{2} & 0 & -(1-i)/(2\sqrt{2}) & -(1+i)/(2\sqrt{2}) \\ 0 & 1/\sqrt{2} & -(1-i)/(2\sqrt{2}) & (1+i)/(2\sqrt{2}) \end{pmatrix}.$$

**3.28.** Let  $A$  be an operator on  $\mathcal{V}$ . Show that if  $\det A = 0$ , then there exists a nonzero vector  $|x\rangle \in \mathcal{V}$  such that  $A|x\rangle = 0$ .

**3.29.** For which values of  $\alpha$  are the following matrices invertible? Find the inverses whenever possible.

$$A = \begin{pmatrix} 1 & \alpha & 0 \\ \alpha & 1 & \alpha \\ 0 & \alpha & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & 1 & 0 \\ 1 & \alpha & 1 \\ 0 & 1 & \alpha \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & 1 & \alpha \\ 1 & \alpha & 0 \\ \alpha & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & \alpha \\ 1 & \alpha & 1 \end{pmatrix}.$$

**3.30.** Let  $\{\mathbf{a}_i\}_{i=1}^N$  be the set consisting of the  $N$  rows of an  $N \times N$  matrix  $A$  and assume that the  $\mathbf{a}_i$  are orthogonal to each other. Show that

$$|\det A| = \|\mathbf{a}_1\| \|\mathbf{a}_2\| \cdots \|\mathbf{a}_N\|.$$

Hint: Consider  $AA^\dagger$ . What would the result be if  $A$  were a unitary matrix?

**3.31.** Prove that a set of  $n$  homogeneous linear equations in  $n$  unknowns has a nontrivial solution if and only if the determinant of the matrix of coefficients is zero.

**3.32.** Use determinants to show that an antisymmetric matrix whose dimension is odd cannot have an inverse.

**3.33.** Show that  $\text{tr}(|a\rangle\langle b|) = \langle b|a\rangle$ . Hint: Evaluate the trace in an orthonormal basis.

**3.34.** Show that if two invertible  $N \times N$  matrices  $A$  and  $B$  anticommute (that is,  $AB + BA = 0$ ), then (a)  $N$  must be even, and (b)  $\text{tr} A = \text{tr} B = 0$ .

**3.35.** Show that for a spatial rotation  $R_{\hat{n}}(\theta)$  of an angle  $\theta$  about an *arbitrary* axis  $\hat{n}$ ,  $\text{tr} R_{\hat{n}}(\theta) = 1 + 2 \cos \theta$ .

**3.36.** Express the sum of the squares of elements of a matrix as a trace. Show that this sum is invariant under an orthogonal transformation of the matrix.

**3.37.** Let  $S$  and  $A$  be a symmetric and an antisymmetric matrix, respectively, and let  $M$  be a general matrix. Show that

- (a)  $\text{tr} M = \text{tr} M^t$ ,
- (b)  $\text{tr}(SA) = 0$ ; in particular,  $\text{tr} A = 0$ ,
- (c)  $SA$  is antisymmetric if and only if  $[S, A] = 0$ ,
- (d)  $MSM^t$  is symmetric and  $MAM^t$  is antisymmetric.
- (e)  $MHM^\dagger$  is hermitian if  $H$  is.

**3.38.** Find the trace of each of the following linear operators:

(a)  $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  given by

$$\mathbf{T}(x, y, z) = (x + y - z, 2x + 3y - 2z, x - y).$$

(b)  $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  given by

$$\mathbf{T}(x, y, z) = (y - z, x + 2y + z, z - y).$$

(c)  $\mathbf{T} : \mathbb{C}^4 \rightarrow \mathbb{C}^4$  given by

$$\mathbf{T}(x, y, z, w) = (x + iy - z + iw, 2ix + 3y - 2iz - w, x - iy, z + iw).$$

**3.39.** Use Equation (3.29) to derive Equation (3.27).

**3.40.** Suppose that there are two operators  $\mathbf{A}$  and  $\mathbf{B}$  such that  $[\mathbf{A}, \mathbf{B}] = c\mathbf{1}$ , where  $c$  is a constant. Show that the vector space in which such operators are defined cannot be finite-dimensional. Conclude that the position and momentum operators of quantum mechanics can be defined only in infinite dimensions.

## Additional Reading

1. Axler, S. *Linear Algebra Done Right*, Springer-Verlag, 1996.
2. Birkhoff, G. and MacLane, S. *Modern Algebra*, 4th ed., Macmillan, 1977. Discusses matrices from the standpoint of row and column operations.
3. Greub, W. *Linear Algebra*, 4th ed., Springer-Verlag, 1975.



# Spectral Decomposition

The last chapter discussed matrix representation of operators. It was pointed out there that such a representation is basis-dependent. In some bases, the operator may “look” quite complicated, while in others it may take a simple form. In a “special” basis, the operator may look the simplest: It may be a diagonal matrix. This chapter investigates conditions under which a basis exists in which the operator is represented by a diagonal matrix.

## 4.1 Direct Sums

Sometimes it is possible, and convenient, to break up a vector space into special (disjoint) subspaces. For instance, it is convenient to decompose the motion of a projectile into its horizontal and vertical components. Similarly, the study of the motion of a particle in  $\mathbb{R}^3$  under the influence of a central force field is facilitated by decomposing the motion into its projections onto the direction of angular momentum and onto a plane perpendicular to the angular momentum. This corresponds to decomposing a vector in space into a vector, say, in the  $xy$ -plane and a vector along the  $z$ -axis. We can generalize this to any vector space, but first some notation: Let  $\mathcal{U}$  and  $\mathcal{W}$  be subspaces of a vector space  $\mathcal{V}$ . Denote by  $\mathcal{U} + \mathcal{W}$  the collection of all vectors in  $\mathcal{V}$  that can be written as a sum of two vectors, one in  $\mathcal{U}$  and one in  $\mathcal{W}$ . It is easy to show that  $\mathcal{U} + \mathcal{W}$  is a subspace of  $\mathcal{V}$ .

Sum of two  
subspaces defined

**4.1.1. Example.** Let  $\mathcal{U}$  be the  $xy$ -plane and  $\mathcal{W}$  the  $yz$ -plane. These are both subspaces of  $\mathbb{R}^3$ , and so is  $\mathcal{U} + \mathcal{W}$ . In fact,  $\mathcal{U} + \mathcal{W} = \mathbb{R}^3$ , because given any vector  $(x, y, z)$  in  $\mathbb{R}^3$ , we can write it as

$$(x, y, z) = \underbrace{\left(x, \frac{1}{2}y, 0\right)}_{\in \mathcal{U}} + \underbrace{\left(0, \frac{1}{2}y, z\right)}_{\in \mathcal{W}}.$$

This decomposition is not unique: We could also write  $(x, y, z) = (x, \frac{1}{3}y, 0) + (0, \frac{2}{3}y, z)$ , and a host of other relations. ■

**4.1.2. Definition.** Let  $\mathcal{U}$  and  $\mathcal{W}$  be subspaces of a vector space  $\mathcal{V}$  such that  $\mathcal{V} = \mathcal{U} + \mathcal{W}$  and the only vector common to both  $\mathcal{U}$  and  $\mathcal{W}$  is the zero vector. Then we say that  $\mathcal{V}$  is the **direct sum** of  $\mathcal{U}$  and  $\mathcal{W}$  and write

$$\mathcal{V} = \mathcal{U} \oplus \mathcal{W}.$$

**4.1.3. Proposition.** Let  $\mathcal{U}$  and  $\mathcal{W}$  be subspaces of  $\mathcal{V}$ . Then  $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$  if and only if any vector in  $\mathcal{V}$  can be written uniquely as a vector in  $\mathcal{U}$  plus a vector in  $\mathcal{W}$ .

*Proof.* Assume  $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$ , and let  $|v\rangle \in \mathcal{V}$  be written as a sum of a vector in  $\mathcal{U}$  and a vector in  $\mathcal{W}$  in two different ways:

$$|v\rangle = |u\rangle + |w\rangle = |u'\rangle + |w'\rangle \quad \Leftrightarrow \quad |u\rangle - |u'\rangle = |w'\rangle - |w\rangle.$$

The LHS is in  $\mathcal{U}$ . Since it is equal to the RHS—which is in  $\mathcal{W}$ —it must be in  $\mathcal{W}$  as well. Therefore, the LHS must equal zero, as must the RHS. Thus,  $|u\rangle = |u'\rangle$ ,  $|w'\rangle = |w\rangle$ , and there is only one way that  $|v\rangle$  can be written as a sum of a vector in  $\mathcal{U}$  and a vector in  $\mathcal{W}$ .

Conversely, if  $|a\rangle \in \mathcal{U}$  and also  $|a\rangle \in \mathcal{W}$ , then one can write

$$|a\rangle = \underbrace{|a\rangle}_{\text{in } \mathcal{U}} + \underbrace{|0\rangle}_{\text{in } \mathcal{W}} \quad \text{and} \quad |a\rangle = \underbrace{|0\rangle}_{\text{in } \mathcal{U}} + \underbrace{|a\rangle}_{\text{in } \mathcal{W}}.$$

Uniqueness of the decomposition of  $|a\rangle$  implies that  $|a\rangle = |0\rangle$ . Therefore, the only vector common to both  $\mathcal{U}$  and  $\mathcal{W}$  is the zero vector. This implies that  $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$ . □

**4.1.4. Proposition.** If  $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$ , then  $\dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{W}$ .

*Proof.* Let  $\{|u_i\rangle\}_{i=1}^m$  be a basis for  $\mathcal{U}$  and  $\{|w_i\rangle\}_{i=1}^k$  a basis for  $\mathcal{W}$ . Then it is easily verified that  $\{|u_1\rangle, |u_2\rangle, \dots, |u_m\rangle, |w_1\rangle, |w_2\rangle, \dots, |w_k\rangle\}$  is a basis for  $\mathcal{V}$ . The details are left as an exercise. □

We can generalize the notion of the direct sum to more than two subspaces. For example, we can write  $\mathbb{R}^3 = \mathcal{X} \oplus \mathcal{Y} \oplus \mathcal{Z}$ , where  $\mathcal{X}$ ,  $\mathcal{Y}$ , and  $\mathcal{Z}$  are the one-dimensional subspaces corresponding to the three axes. Now assume that

$$\mathcal{V} = \mathcal{U}_1 \oplus \mathcal{U}_2 \oplus \dots \oplus \mathcal{U}_r, \tag{4.1}$$

i.e.,  $\mathcal{V}$  is the direct sum of  $r$  of its subspaces that have no common vectors among themselves except the zero vector and have the property that any vector in  $\mathcal{V}$  can be written (uniquely) as a sum of vectors one from each subspace. Define the linear operator  $\mathbf{P}_j$  by  $\mathbf{P}_j |u\rangle = |u_j\rangle$  where  $|u\rangle = \sum_{j=1}^r |u_j\rangle$ ,  $|u_j\rangle \in \mathcal{U}_j$ . Then it is readily

verified that  $\mathbf{P}_j^2 = \mathbf{P}_j$  and  $\mathbf{P}_j \mathbf{P}_k = 0$  for  $j \neq k$ . Thus, the  $\mathbf{P}_j$ 's are (not necessarily hermitian) projection operators. Furthermore, for an arbitrary vector  $|u\rangle$ , we have

$$|u\rangle = \sum_{j=1}^r |u_j\rangle = \sum_{j=1}^r \mathbf{P}_j |u\rangle = \left( \sum_{j=1}^r \mathbf{P}_j \right) |u\rangle \quad \forall |u\rangle \in \mathcal{V}.$$

Since this is true for all vectors, we have the identity

$$\sum_{j=1}^r \mathbf{P}_j = \mathbf{1}. \tag{4.2}$$

orthogonal  
complement of a  
subspace

**4.1.5. Definition.** Let  $\mathcal{V}$  be an inner product space. Let  $\mathcal{M}$  be any subspace of  $\mathcal{V}$ . Denote by  $\mathcal{M}^\perp$  the set of all vectors in  $\mathcal{V}$  orthogonal to all the vectors in  $\mathcal{M}$ .  $\mathcal{M}^\perp$  (pronounced "em perp") is called the **orthogonal complement** of  $\mathcal{M}$ .

**4.1.6. Proposition.**  $\mathcal{M}^\perp$  is a subspace of  $\mathcal{V}$ .

*Proof.* In fact, if  $|a\rangle, |b\rangle \in \mathcal{M}^\perp$ , then for any vector  $|c\rangle \in \mathcal{M}$ , we have

$$\langle c | (\alpha |a\rangle + \beta |b\rangle) \rangle = \alpha \overbrace{\langle c | a \rangle}^{=0} + \beta \overbrace{\langle c | b \rangle}^{=0} = 0.$$

So,  $\alpha |a\rangle + \beta |b\rangle \in \mathcal{M}^\perp$  for arbitrary  $\alpha, \beta \in \mathbb{C}$  and  $|a\rangle, |b\rangle \in \mathcal{M}^\perp$ . □

If  $\mathcal{V}$  of Equation (4.1) is an inner product space, and the subspaces are mutually orthogonal, then for arbitrary  $|u\rangle, |v\rangle \in \mathcal{V}$ ,

$$\langle u | \mathbf{P}_j |v\rangle = \langle u | v_j \rangle = \langle u_j | v_j \rangle = \langle v_j | u_j \rangle^* = \langle v | u_j \rangle^* = \langle v | \mathbf{P}_j |u\rangle^*$$

which shows that  $\mathbf{P}_j$  is hermitian. In Chapter 2, we assumed the projection operators to be hermitian. Now we see that only in an inner product space (and only if the subspaces of a direct sum are orthogonal) do we recover the hermiticity of projection operators.

**4.1.7. Example.** Consider an orthonormal basis  $B_M = \{|e_i\rangle\}_{i=1}^m$  for  $\mathcal{M}$ , and extend it to a basis  $B = \{|e_i\rangle\}_{i=1}^N$  for  $\mathcal{V}$ . Now construct a (hermitian) projection operator  $\mathbf{P} = \sum_{i=1}^m |e_i\rangle \langle e_i|$ . This is the operator that projects an arbitrary vector in  $\mathcal{V}$  onto the subspace  $\mathcal{M}$ . It is straightforward to show that  $\mathbf{1} - \mathbf{P}$  is the projection operator that projects onto  $\mathcal{M}^\perp$  (see Problem 4.2).

An arbitrary vector  $|a\rangle \in \mathcal{V}$  can be written as

$$|a\rangle = (\mathbf{P} + \mathbf{1} - \mathbf{P}) |a\rangle = \underbrace{\mathbf{P} |a\rangle}_{\text{in } \mathcal{M}} + \underbrace{(\mathbf{1} - \mathbf{P}) |a\rangle}_{\text{in } \mathcal{M}^\perp}.$$

Furthermore, the only vector that can be in both  $\mathcal{M}$  and  $\mathcal{M}^\perp$  is the zero vector, because it is the only vector orthogonal to itself. ■

From this example and the remarks immediately preceding it we may conclude the following:

**4.1.8. Proposition.** *If  $\mathcal{V}$  is an inner product space, then  $\mathcal{V} = \mathcal{M} \oplus \mathcal{M}^\perp$  for any subspace  $\mathcal{M}$ . Furthermore, the projection operators corresponding to  $\mathcal{M}$  and  $\mathcal{M}^\perp$  are hermitian.*

## 4.2 Invariant Subspaces

This section explores the possibility of obtaining subspaces by means of the action of a linear operator on vectors of an  $N$ -dimensional vector space  $\mathcal{V}$ . Let  $|a\rangle$  be any vector in  $\mathcal{V}$ , and  $\mathbf{A}$  a linear operator on  $\mathcal{V}$ . The vectors

$$|a\rangle, \mathbf{A}|a\rangle, \mathbf{A}^2|a\rangle, \dots, \mathbf{A}^N|a\rangle$$

are linearly dependent (there are  $N + 1$  of them!). Let  $\mathcal{M} \equiv \text{Span}\{\mathbf{A}^k|a\rangle\}_{k=0}^N$ . It follows that,  $m \equiv \dim \mathcal{M} \leq \dim \mathcal{V}$ , and  $\mathcal{M}$  has the property that for any vector  $|x\rangle \in \mathcal{M}$  the vector  $\mathbf{A}|x\rangle$  also belongs to  $\mathcal{M}$  (show this!). In other words, no vector in  $\mathcal{M}$  "leaves" the subspace when acted on by  $\mathbf{A}$ .

invariant subspace;  
reduction of an  
operator

**4.2.1. Definition.** *A subspace  $\mathcal{M}$  is an invariant subspace of the operator  $\mathbf{A}$  if  $\mathbf{A}$  transforms vectors of  $\mathcal{M}$  into vectors of  $\mathcal{M}$ . This is written succinctly as  $\mathbf{A}(\mathcal{M}) \subset \mathcal{M}$ . We say that  $\mathcal{M}$  reduces  $\mathbf{A}$  if both  $\mathcal{M}$  and  $\mathcal{M}^\perp$  are invariant subspaces of  $\mathbf{A}$ .*

Starting with a basis of  $\mathcal{M}$ , we can extend it to a basis  $B = \{|a_i\rangle\}_{i=1}^N$  of  $\mathcal{V}$  whose first  $m$  vectors span  $\mathcal{M}$ . The matrix representation of  $\mathbf{A}$  in such a basis is given by the relation  $\mathbf{A}|a_i\rangle = \sum_{j=1}^N \alpha_{ji}|a_j\rangle, i = 1, 2, \dots, N$ . If  $i \leq m$ , then  $\alpha_{ji} = 0$  for  $j > m$ , because  $\mathbf{A}|a_i\rangle$  belongs to  $\mathcal{M}$  when  $i \leq m$  and therefore can be written as a linear combination of only  $\{|a_1\rangle, |a_2\rangle, \dots, |a_m\rangle\}$ . Thus, the matrix representation of  $\mathbf{A}$  in  $B$  will have the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

matrix representation  
of an operator in a  
subspace

where  $\mathbf{A}_{11}$  is an  $m \times m$  matrix,  $\mathbf{A}_{12}$  an  $m \times (N - m)$  matrix,  $\mathbf{0}_{21}$  the  $(N - m) \times m$  zero matrix, and  $\mathbf{A}_{22}$  an  $(N - m) \times (N - m)$  matrix. We say that  $\mathbf{A}_{11}$  represents the operator  $\mathbf{A}$  in the  $m$ -dimensional subspace  $\mathcal{M}$ .

block diagonal matrix  
defined

It may also happen that the subspace spanned by the remaining basis vectors in  $B$ , namely  $|a_{m+1}\rangle, |a_{m+2}\rangle, \dots, |a_N\rangle$ , is also an invariant subspace of  $\mathbf{A}$ . Then  $\mathbf{A}_{12}$  will be zero, and  $\mathbf{A}$  will take a **block diagonal** form:<sup>1</sup>

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix}$$

reducible and  
irreducible matrices

A matrix representation of an operator that can be brought into this form by a suitable choice of basis is said to be **reducible**; otherwise, it is called **irreducible**.

<sup>1</sup>From now on, we shall denote all zero matrices by the same symbol regardless of their dimensionality.

A reducible matrix  $\mathbf{A}$  is denoted in two different ways:<sup>2</sup>

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix} \Leftrightarrow \mathbf{A} = \mathbf{A}_1 \oplus \mathbf{A}_2. \tag{4.3}$$

For example, when  $\mathcal{M}$  reduces  $\mathbf{A}$  and one chooses a basis the first  $m$  vectors of which are in  $\mathcal{M}$  and the remaining ones in  $\mathcal{M}^\perp$ , then  $\mathbf{A}$  is reducible. We have seen on a number of occasions the significance of the hermitian conjugate of an operator (e.g., in relation to hermitian and unitary operators). The importance of this operator will be borne out further when we study the spectral theorem later in this chapter. Let us now investigate some properties of the adjoint of an operator in the context of invariant subspaces.

condition for invariance

**4.2.2. Lemma.** *A subspace  $\mathcal{M}$  of an inner product space  $\mathcal{V}$  is invariant under the linear operator  $\mathbf{A}$  if and only if  $\mathcal{M}^\perp$  is invariant under  $\mathbf{A}^\dagger$ .*

*Proof.* The proof is left as a problem. □

An immediate consequence of the above lemma and the two identities  $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$  and  $(\mathcal{M}^\perp)^\perp = \mathcal{M}$  is contained in the following theorem.

**4.2.3. Theorem.** *A subspace of  $\mathcal{V}$  reduces  $\mathbf{A}$  if and only if it is invariant under both  $\mathbf{A}$  and  $\mathbf{A}^\dagger$ .*

**4.2.4. Lemma.** *Let  $\mathcal{M}$  be a subspace of  $\mathcal{V}$  and  $\mathbf{P}$  the hermitian projection operator onto  $\mathcal{M}$ . Then  $\mathcal{M}$  is invariant under the linear operator  $\mathbf{A}$  if and only if  $\mathbf{AP} = \mathbf{PAP}$ .*

*Proof.* Suppose  $\mathcal{M}$  is invariant. Then for any  $|x\rangle$  in  $\mathcal{V}$ , we have

$$\mathbf{P}|x\rangle \in \mathcal{M} \Rightarrow \mathbf{AP}|x\rangle \in \mathcal{M} \Rightarrow \mathbf{PAP}|x\rangle = \mathbf{AP}|x\rangle.$$

Since the last equality holds for arbitrary  $|x\rangle$ , we have  $\mathbf{AP} = \mathbf{PAP}$ .

Conversely, suppose  $\mathbf{AP} = \mathbf{PAP}$ . For any  $|y\rangle \in \mathcal{M}$ , we have

$$\mathbf{P}|y\rangle = |y\rangle \Rightarrow \underbrace{\mathbf{AP}}_{=\mathbf{PAP}}|y\rangle = \mathbf{A}|y\rangle = \mathbf{P}(\mathbf{AP}|y\rangle) \in \mathcal{M}.$$

Therefore,  $\mathcal{M}$  is invariant under  $\mathbf{A}$ . □

**4.2.5. Theorem.** *Let  $\mathcal{M}$  be a subspace of  $\mathcal{V}$ ,  $\mathbf{P}$  the hermitian projection operator of  $\mathcal{V}$  onto  $\mathcal{M}$ , and  $\mathbf{A}$  a linear operator on  $\mathcal{V}$ . Then  $\mathcal{M}$  reduces  $\mathbf{A}$  if and only if  $\mathbf{A}$  and  $\mathbf{P}$  commute.*

---

<sup>2</sup>It is common to use a single subscript for submatrices of a block diagonal matrix, just as it is common to use a single subscript for entries of a diagonal matrix.

*Proof.* Suppose  $\mathcal{M}$  reduces  $\mathbf{A}$ . Then by Theorem 4.2.3,  $\mathcal{M}$  is invariant under both  $\mathbf{A}$  and  $\mathbf{A}^\dagger$ . Lemma 4.2.4 then implies

$$\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{A}\mathbf{P} \quad \text{and} \quad \mathbf{A}^\dagger\mathbf{P} = \mathbf{P}\mathbf{A}^\dagger\mathbf{P}. \quad (4.4)$$

Taking the adjoint of the second equation yields  $(\mathbf{A}^\dagger\mathbf{P})^\dagger = (\mathbf{P}\mathbf{A}^\dagger\mathbf{P})^\dagger$ , or  $\mathbf{P}\mathbf{A} = \mathbf{P}\mathbf{A}\mathbf{P}$ . This equation together with the first equation of (4.4) yields  $\mathbf{P}\mathbf{A} = \mathbf{A}\mathbf{P}$ .

Conversely, suppose that  $\mathbf{P}\mathbf{A} = \mathbf{A}\mathbf{P}$ . Then  $\mathbf{P}^2\mathbf{A} = \mathbf{P}\mathbf{A}\mathbf{P}$ , whence  $\mathbf{P}\mathbf{A} = \mathbf{P}\mathbf{A}\mathbf{P}$ . Taking adjoints gives  $\mathbf{A}^\dagger\mathbf{P} = \mathbf{P}\mathbf{A}^\dagger\mathbf{P}$ , because  $\mathbf{P}$  is hermitian. By Lemma 4.2.4,  $\mathcal{M}$  is invariant under  $\mathbf{A}^\dagger$ . Similarly, from  $\mathbf{P}\mathbf{A} = \mathbf{A}\mathbf{P}$ , we get  $\mathbf{P}\mathbf{A}\mathbf{P} = \mathbf{A}\mathbf{P}^2$ , whence  $\mathbf{P}\mathbf{A}\mathbf{P} = \mathbf{A}\mathbf{P}$ . Once again by Lemma 4.2.4,  $\mathcal{M}$  is invariant under  $\mathbf{A}$ . By Theorem 4.2.3,  $\mathcal{M}$  reduces  $\mathbf{A}$ .  $\square$

The main goal of the remaining part of this chapter is to prove that certain operators, e.g. hermitian operators, are diagonalizable, that is, that we can always find an (orthonormal) basis in which they are represented by a diagonal matrix.

### 4.3 Eigenvalues and Eigenvectors

Let us begin by considering eigenvalues and eigenvectors, which are generalizations of familiar concepts in two and three dimensions. Consider the operation of rotation about the  $z$ -axis by an angle  $\theta$  denoted by  $\mathbf{R}_z(\theta)$ . Such a rotation takes any vector  $(x, y)$  in the  $xy$ -plane to a new vector  $(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)$ . Thus, unless  $(x, y) = (0, 0)$  or  $\theta$  is an integer multiple of  $2\pi$ , the vector will change. Is there a nonzero vector that is so special (*eigen*, in German) that it does not change when acted on by  $\mathbf{R}_z(\theta)$ ? As long as we confine ourselves to two dimensions, the answer is no. But if we lift ourselves up from the two-dimensional  $xy$ -plane, we encounter many such vectors, all of which lie along the  $z$ -axis.

The foregoing example can be generalized to any rotation (normally specified by Euler angles). In fact, the methods developed in this section can be used to show that a general rotation, given by Euler angles, always has an unchanged vector lying along the axis around which the rotation takes place. This concept is further generalized in the following definition.

eigenvalue and  
eigenvector

**4.3.1. Definition.** A scalar  $\lambda$  is an *eigenvalue* and a nonzero vector  $|a\rangle$  is an *eigenvector* of the linear transformation  $\mathbf{A} \in \mathcal{L}(\mathcal{V})$  if

$$\mathbf{A}|a\rangle = \lambda|a\rangle. \quad (4.5)$$

**4.3.2. Proposition.** Add the zero vector to the set of all eigenvectors of  $\mathbf{A}$  belonging to the same eigenvalue  $\lambda$ , and denote the span of the resulting set by  $\mathcal{M}_\lambda$ . Then  $\mathcal{M}_\lambda$  is a subspace of  $\mathcal{V}$ , and every (nonzero) vector in  $\mathcal{M}_\lambda$  is an eigenvector of  $\mathbf{A}$  with eigenvalue  $\lambda$ .

*Proof.* The proof follows immediately from the above definition and the definition of a subspace.  $\square$

eigenspace **4.3.3. Definition.** The subspace  $\mathcal{M}_\lambda$  is referred to as the **eigenspace** of  $\mathbf{A}$  corresponding to the eigenvalue  $\lambda$ . Its dimension is called the **geometric multiplicity** of  $\lambda$ . An eigenvalue is called **simple** if its geometric multiplicity is 1. The set of spectrum eigenvalues of  $\mathbf{A}$  is called the **spectrum** of  $\mathbf{A}$ .

By their very construction, eigenspaces corresponding to different eigenvalues have no vectors in common except the zero vector. This can be demonstrated by noting that if  $|v\rangle \in \mathcal{M}_\lambda \cap \mathcal{M}_\mu$  for  $\lambda \neq \mu$ , then

$$0 = (\mathbf{A} - \lambda \mathbf{1}) |v\rangle = \mathbf{A} |v\rangle - \lambda |v\rangle = \mu |v\rangle - \lambda |v\rangle = \underbrace{(\mu - \lambda)}_{\neq 0} |v\rangle \Rightarrow |v\rangle = 0.$$

Let us rewrite Equation (4.5) as  $(\mathbf{A} - \lambda \mathbf{1}) |a\rangle = 0$ . This equation says that  $|a\rangle$  is an eigenvector of  $\mathbf{A}$  if and only if  $|a\rangle$  belongs to the kernel of  $\mathbf{A} - \lambda \mathbf{1}$ . If the latter is invertible, then its kernel will consist of only the zero vector, which is not acceptable as a solution of Equation (4.5). Thus, if we are to obtain nontrivial solutions,  $\mathbf{A} - \lambda \mathbf{1}$  must have no inverse. This is true if and only if

$$\det(\mathbf{A} - \lambda \mathbf{1}) = 0. \tag{4.6}$$

characteristic polynomial and characteristic roots of an operator

The determinant in Equation (4.6) is a polynomial in  $\lambda$ , called the **characteristic polynomial** of  $\mathbf{A}$ . The roots of this polynomial are called **characteristic roots** and are simply the eigenvalues of  $\mathbf{A}$ . Now, any polynomial of degree greater than or equal to 1 has at least one (complex) root, which yields the following theorem.

**4.3.4. Theorem.** Every operator on a finite-dimensional vector space over  $\mathbb{C}$  has at least one eigenvalue and therefore at least one eigenvector.

Let  $\lambda_1, \lambda_2, \dots, \lambda_p$  be the distinct roots of the characteristic polynomial of  $\mathbf{A}$ , and let  $\lambda_j$  occur  $m_j$  times. Then <sup>3</sup>

$$\det(\mathbf{A} - \lambda \mathbf{1}) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_p - \lambda)^{m_p} = \prod_{j=1}^p (\lambda_j - \lambda)^{m_j}. \tag{4.7}$$

For  $\lambda = 0$ , this gives

$$\det \mathbf{A} = \lambda_1^{m_1} \lambda_2^{m_2} \dots \lambda_p^{m_p} = \prod_{j=1}^p \lambda_j^{m_j}. \tag{4.8}$$

Equation (4.8) states that the determinant of an operator is the product of all its eigenvalues. In particular, if one of the eigenvalues is zero, then the operator is not invertible.

<sup>3</sup> $m_j$  is called the **algebraic multiplicity** of  $\lambda_j$ .

**4.3.5. Example.** Let us find the eigenvalues of a projection operator  $\mathbf{P}$ . If  $|a\rangle$  is an eigenvector, then  $\mathbf{P}|a\rangle = \lambda|a\rangle$ . Applying  $\mathbf{P}$  on both sides again, we obtain

$$\mathbf{P}^2|a\rangle = \lambda\mathbf{P}|a\rangle = \lambda(\lambda|a\rangle) = \lambda^2|a\rangle.$$

But  $\mathbf{P}^2 = \mathbf{P}$ ; thus,  $\mathbf{P}|a\rangle = \lambda^2|a\rangle$ . It follows that  $\lambda^2|a\rangle = \lambda|a\rangle$ , or  $(\lambda^2 - \lambda)|a\rangle = 0$ . Since  $|a\rangle \neq 0$ , we must have  $\lambda(\lambda - 1) = 0$ , or  $\lambda = 0, 1$ . Thus, the only eigenvalues of a projection operator are 0 and 1. The presence of zero as an eigenvalue of  $\mathbf{P}$  is an indication that  $\mathbf{P}$  is not invertible. ■

**4.3.6. Example.** To be able to see the difference between algebraic and geometric multiplicities, consider the matrix  $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ , whose characteristic polynomial is  $(1 - \lambda)^2$ . Thus, the matrix has only one eigenvalue,  $\lambda = 1$ , with algebraic multiplicity  $m_1 = 2$ . However, the most general vector  $|a\rangle$  satisfying  $(\mathbf{A} - 1)|a\rangle = 0$  is easily shown to be of the form  $\begin{pmatrix} \alpha \\ 0 \end{pmatrix}$ . This shows that  $\mathcal{M}_{\lambda=1}$  is one-dimensional, i.e., the geometric multiplicity of  $\lambda$  is 1. ■

As mentioned at the beginning of this chapter, it is useful to represent an operator by as simple a matrix as possible. The simplest matrix is a diagonal matrix. This motivates the following definition:

diagonalizable  
operators

**4.3.7. Definition.** A linear operator  $\mathbf{A}$  on a vector space  $\mathcal{V}$  is said to be **diagonalizable** if there is a basis for  $\mathcal{V}$  all of whose vectors are eigenvectors of  $\mathbf{A}$ .

**4.3.8. Theorem.** Let  $\mathbf{A}$  be a diagonalizable operator on a vector space  $\mathcal{V}$  with distinct eigenvalues  $\{\lambda_j\}_{j=1}^r$ . There are (not necessarily hermitian) projection operators  $\mathbf{P}_j$  on  $\mathcal{V}$  such that

$$(1) \mathbf{1} = \sum_{j=1}^r \mathbf{P}_j, \quad (2) \mathbf{P}_i \mathbf{P}_j = 0 \text{ for } i \neq j, \quad (3) \mathbf{A} = \sum_{j=1}^r \lambda_j \mathbf{P}_j.$$

*Proof.* Let  $\mathcal{M}_j$  denote the eigenspace corresponding to the eigenvalue  $\lambda_j$ . Since the eigenvectors span  $\mathcal{V}$  and the only common vector of two different eigenspaces is the zero vector (see comments after Definition 4.3.3), we have

$$\mathcal{V} = \mathcal{M}_1 \oplus \mathcal{M}_2 \oplus \cdots \oplus \mathcal{M}_r.$$

This immediately gives (1) and (2) if we use Equations (4.1) and (4.2) where  $\mathbf{P}_j$  is the projection operator onto  $\mathcal{M}_j$ .

To prove (3), let  $|v\rangle$  be an arbitrary vector in  $\mathcal{V}$ . Then  $|v\rangle$  can be written uniquely as a sum of vectors each coming from one eigenspace. Therefore,

$$\mathbf{A}|v\rangle = \sum_{j=1}^r \mathbf{A}|v_j\rangle = \sum_{j=1}^r \lambda_j |v_j\rangle = \left( \sum_{j=1}^r \lambda_j \mathbf{P}_j \right) |v\rangle.$$

Since this equality holds for all vectors  $|v\rangle$ , (3) follows. □



## 4.4 Spectral Decomposition

This section derives one of the most powerful theorems in the theory of linear operators, the spectral decomposition theorem. We shall derive the theorem for operators that generalize hermitian and unitary operators.

normal operator  
defined

**4.4.1. Definition.** A *normal operator* is an operator on an inner product space that commutes with its adjoint.

An important consequence of this definition is that

$$\|\mathbf{A}x\| = \|\mathbf{A}^\dagger x\| \quad \text{if and only if} \quad \mathbf{A} \text{ is normal.} \quad (4.9)$$

**4.4.2. Proposition.** Let  $\mathbf{A}$  be a normal operator on  $\mathcal{V}$ . Then  $|x\rangle$  is an eigenvector of  $\mathbf{A}$  with eigenvalue  $\lambda$  if and only if  $|x\rangle$  is an eigenvector of  $\mathbf{A}^\dagger$  with eigenvalue  $\lambda^*$ .

*Proof.* By Equation (4.9), the fact that  $(\mathbf{A} - \lambda\mathbf{1})^\dagger = \mathbf{A}^\dagger - \lambda^*\mathbf{1}$ , and the fact that  $\mathbf{A} - \lambda\mathbf{1}$  is normal (reader, verify), we have  $\|(\mathbf{A} - \lambda\mathbf{1})x\| = 0$  if and only if  $\|(\mathbf{A}^\dagger - \lambda^*\mathbf{1})x\| = 0$ . Since it is only the zero vector that has the zero norm, we get

$$(\mathbf{A} - \lambda\mathbf{1})|x\rangle = 0 \quad \text{if and only if} \quad (\mathbf{A}^\dagger - \lambda^*\mathbf{1})|x\rangle = 0.$$

This proves the proposition.  $\square$

We obtain a useful consequence of this proposition by applying it to a hermitian operator  $\mathbf{H}$  and a unitary operator<sup>4</sup>  $\mathbf{U}$ . In the first case, we get

$$\mathbf{H}|x\rangle = \lambda|x\rangle = \mathbf{H}^\dagger|x\rangle = \lambda^*|x\rangle \Rightarrow (\lambda - \lambda^*)|x\rangle = 0 \Rightarrow \lambda = \lambda^*.$$

Therefore,  $\lambda$  is real. In the second case, we write

$$|x\rangle = \mathbf{1}|x\rangle = \mathbf{U}\mathbf{U}^\dagger|x\rangle = \mathbf{U}(\lambda^*|x\rangle) = \lambda^*\mathbf{U}|x\rangle = \lambda\lambda^*|x\rangle \Rightarrow \lambda\lambda^* = 1.$$

Therefore,  $\lambda$  is unimodular (has absolute value equal to 1). We summarize the foregoing discussion:

**4.4.3. Corollary.** The eigenvalues of a hermitian operator are real. The eigenvalues of a unitary operator have unit absolute value.

**4.4.4. Example.** Let us find the eigenvalues and eigenvectors of the hermitian matrix  $\mathbf{H} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ . We have

$$\det(\mathbf{H} - \lambda\mathbf{1}) = \det \begin{pmatrix} -\lambda & -i \\ i & -\lambda \end{pmatrix} = \lambda^2 - 1.$$

Thus, the eigenvalues,  $\lambda_1 = 1$  and  $\lambda_2 = -1$ , are real, as expected.

<sup>4</sup>Obviously, both are normal operators.

To find the eigenvectors, we write

$$0 = (\mathbf{H} - \lambda_1 \mathbf{1}) |a_1\rangle = (\mathbf{H} - \mathbf{1}) |a_1\rangle = \begin{pmatrix} -1 & -i \\ i & -1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} -\alpha_1 - i\alpha_2 \\ i\alpha_1 - \alpha_2 \end{pmatrix}$$

or  $\alpha_2 = i\alpha_1$ , which gives  $|a_1\rangle = \begin{pmatrix} \alpha_1 \\ i\alpha_1 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ i \end{pmatrix}$ , where  $\alpha_1$  is an arbitrary complex number. Also,

$$0 = (\mathbf{H} - \lambda_2 \mathbf{1}) |a_2\rangle = (\mathbf{H} + \mathbf{1}) |a_2\rangle = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \beta_1 - i\beta_2 \\ i\beta_1 + \beta_2 \end{pmatrix}$$

or  $\beta_2 = -i\beta_1$ , which gives  $|a_2\rangle = \begin{pmatrix} \beta_1 \\ -i\beta_1 \end{pmatrix} = \beta_1 \begin{pmatrix} 1 \\ -i \end{pmatrix}$ , where  $\beta_1$  is an arbitrary complex number.

Always normalize the eigenvectors!

It is desirable, in most situations, to orthonormalize the eigenvectors. In the present case, they are already orthogonal. This is a property shared by all eigenvectors of a hermitian (in fact, normal) operator stated in the next theorem. We therefore need to merely normalize the eigenvectors:

$$1 = \langle a_1 | a_1 \rangle = \alpha_1^* (1 \quad -i) \alpha_1 \begin{pmatrix} 1 \\ i \end{pmatrix} = 2|\alpha_1|^2$$

or  $|\alpha_1| = 1/\sqrt{2}$  and  $\alpha_1 = e^{i\varphi}/\sqrt{2}$  for some  $\varphi \in \mathbb{R}$ . A similar result is obtained for  $\beta_1$ . The choice  $\varphi = 0$  yields

$$|e_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{and} \quad |e_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad \blacksquare$$

The following theorem proves for all normal operators the orthogonality property of their eigenvectors illustrated in the example above for a simple hermitian operator.

**4.4.5. Theorem.** *An eigenspace of a normal operator reduces that operator. Moreover, eigenspaces of a normal operator are mutually orthogonal.*

*Proof.* The first part of the theorem is a trivial consequence of Proposition 4.4.2 and Theorem 4.2.3. To prove the second part, let  $|u\rangle \in \mathcal{M}_\lambda$  and  $|v\rangle \in \mathcal{M}_\mu$  with  $\lambda \neq \mu$ . Then, using Theorem 4.2.3 once more, we obtain

$$\lambda \langle v | u \rangle = \langle v | \lambda u \rangle = \langle v | \mathbf{A}u \rangle = \langle \mathbf{A}^\dagger v | u \rangle = \langle \mu^* v | u \rangle = \mu \langle v | u \rangle$$

It follows that  $(\lambda - \mu) \langle v | u \rangle = 0$  and since  $\lambda \neq \mu$ ,  $\langle v | u \rangle = 0$ . □

spectral decomposition theorem

**4.4.6. Theorem.** (Spectral Decomposition Theorem) *Let  $\mathbf{A}$  be a normal operator on a finite-dimensional complex inner product space  $\mathcal{V}$ . Let  $\lambda_1, \lambda_2, \dots, \lambda_r$  be its distinct eigenvalues. Then there exist nonzero (hermitian) projection operators  $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_r$  such that*

$$1. \quad \mathbf{P}_i \mathbf{P}_j = \mathbf{0} \quad \forall i \neq j;$$

$$2. \quad \sum_{i=1}^r \mathbf{P}_i = \mathbf{1};$$

$$3. \sum_{i=1}^r \lambda_i \mathbf{P}_i = \mathbf{A}.$$

*Proof.* Let  $\mathbf{P}_i$  be the operator that projects onto the eigenspace  $\mathcal{M}_i$  corresponding to eigenvalue  $\lambda_i$ . By comments after Proposition 4.1.6, these projection operators are hermitian. Because of Theorem 4.4.5, the only vector common to any two distinct eigenspaces is the zero vector. So, it makes sense to talk about the direct sum of these eigenspaces. Let  $\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2 \oplus \cdots \oplus \mathcal{M}_r$  and  $\mathbf{P} = \sum_{i=1}^r \mathbf{P}_i$ , where  $\mathbf{P}$  is the orthogonal projection operator onto  $\mathcal{M}$ . Since  $\mathbf{A}$  commutes with every  $\mathbf{P}_i$  (Theorem 4.2.5), it commutes with  $\mathbf{P}$ . Hence, by Theorem 4.2.5,  $\mathcal{M}$  reduces  $\mathbf{A}$ , i.e.,  $\mathcal{M}^\perp$  is also invariant under  $\mathbf{A}$ . Now regard the restriction of  $\mathbf{A}$  to  $\mathcal{M}^\perp$  as an operator in its own right on the finite-dimensional vector space  $\mathcal{M}^\perp$ . Theorem 4.3.4 now forces  $\mathbf{A}$  to have at least one eigenvector in  $\mathcal{M}^\perp$ . But this is impossible because all eigenvectors of  $\mathbf{A}$  have been accounted for in its eigenspaces. The only resolution is for  $\mathcal{M}^\perp$  to be zero. This gives

$$\mathcal{V} = \mathcal{M}_1 \oplus \mathcal{M}_2 \oplus \cdots \oplus \mathcal{M}_r \quad \text{and} \quad \mathbf{1} = \sum_{i=1}^r \mathbf{P}_i.$$

The second equation follows from the first and Equations (4.1) and (4.2). The remaining part of the theorem follows from arguments similar to those used in the proof of Theorem 4.3.8.  $\square$

We can now establish the connection between the diagonalizability of a normal operator and the spectral theorem. In each subspace  $\mathcal{M}_i$ , we choose an orthonormal basis. The union of all these bases is clearly a basis for the whole space  $\mathcal{V}$ . Let us label these basis vectors  $|e_j^s\rangle$ , where the subscript indicates the subspace and the superscript indicates the particular vector in that subspace. Clearly,  $\langle e_j^s | e_{j'}^{s'} \rangle = \delta_{ss'} \delta_{jj'}$  and  $\mathbf{P}_j = \sum_{s=1}^{m_j} |e_j^s\rangle \langle e_j^s|$ . Noting that  $\mathbf{P}_k |e_{j'}^{s'}\rangle = \delta_{kj'} |e_{j'}^{s'}\rangle$ , we can obtain the matrix elements of  $\mathbf{A}$  in such a basis:

$$\langle e_j^s | \mathbf{A} | e_{j'}^{s'} \rangle = \sum_{i=1}^r \lambda_i \langle e_j^s | \mathbf{P}_i | e_{j'}^{s'} \rangle = \sum_{i=1}^r \lambda_i \delta_{ij'} \langle e_j^s | e_{j'}^{s'} \rangle = \lambda_{j'} \langle e_j^s | e_{j'}^{s'} \rangle.$$

Only the diagonal elements are nonzero. We note that for each subscript  $j$  we have  $m_j$  orthonormal vectors  $|e_j^s\rangle$ , where  $m_j$  is the dimension of  $\mathcal{M}_j$ . Thus,  $\lambda_j$  occurs  $m_j$  times as a diagonal element. Therefore, in such an orthonormal basis,  $\mathbf{A}$  will be represented by

$$\text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{m_1 \text{ times}}, \underbrace{\lambda_2, \dots, \lambda_2}_{m_2 \text{ times}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{m_r \text{ times}}).$$

Let us summarize the foregoing discussion:

**4.4.7. Corollary.** *If  $\mathbf{A} \in \mathcal{L}(\mathcal{V})$  is normal, then  $\mathcal{V}$  has an orthonormal basis consisting of eigenvectors of  $\mathbf{A}$ . Therefore, a normal operator on a complex inner product space is diagonalizable.*

Using this corollary, the reader may show the following:

**4.4.8. Corollary.** *A hermitian operator is positive if and only if all its eigenvalues are positive.*

**4.4.9. Example.** COMPUTATION OF LARGEST AND SMALLEST EIGENVALUES

Computation of the largest and the smallest eigenvalues of a normal operator

There is an elegant technique that yields the largest and the smallest (in absolute value) eigenvalues of a normal operator  $\mathbf{A}$  in a straightforward way if the eigenspaces of these eigenvalues are one dimensional. For convenience, assume that the eigenvalues are labeled in order of decreasing absolute values:

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_r| \neq 0.$$

Let  $\{|a_k\rangle\}_{k=1}^N$  be a basis of  $\mathcal{V}$  consisting of eigenvectors of  $\mathbf{A}$ , and  $|x\rangle = \sum_{k=1}^N \xi_k |a_k\rangle$  an arbitrary vector in  $\mathcal{V}$ . Then

$$\mathbf{A}^m |x\rangle = \sum_{k=1}^N \xi_k \mathbf{A}^m |a_k\rangle = \sum_{k=1}^N \xi_k \lambda_k^m |a_k\rangle = \lambda_1^m \left[ \xi_1 |a_1\rangle + \sum_{k=2}^N \xi_k \left(\frac{\lambda_k}{\lambda_1}\right)^m |a_k\rangle \right].$$

In the limit  $m \rightarrow \infty$ , the summation in the brackets vanishes. Therefore,

$$\mathbf{A}^m |x\rangle \approx \lambda_1^m \xi_1 |a_1\rangle \quad \text{and} \quad \langle y | \mathbf{A}^m |x\rangle \approx \lambda_1^m \xi_1 \langle y | a_1\rangle$$

for any  $|y\rangle \in \mathcal{V}$ . Taking the ratio of this equation and the corresponding one for  $m + 1$ , we obtain

$$\lim_{m \rightarrow \infty} \frac{\langle y | \mathbf{A}^{m+1} |x\rangle}{\langle y | \mathbf{A}^m |x\rangle} = \lambda_1.$$

Note how crucially this relation depends on the fact that  $\lambda_1$  is nondegenerate, i.e., that  $\mathcal{M}_1$  is one-dimensional. By taking larger and larger values for  $m$ , we can obtain a better and better approximation to the largest eigenvalue.

Assuming that zero is not the smallest eigenvalue  $\lambda_r$ —and therefore not an eigenvalue—of  $\mathbf{A}$ , we can find the smallest eigenvalue by replacing  $\mathbf{A}$  with  $\mathbf{A}^{-1}$  and  $\lambda_1$  with  $1/\lambda_r$ . The details are left as an exercise for the reader. ■

A hermitian matrix can be diagonalized by a unitary matrix.

Any given hermitian matrix  $\mathbf{H}$  can be thought of as the representation of a hermitian operator in the standard orthonormal basis. We can find a unitary matrix  $\mathbf{U}$  that can transform the standard basis to the orthonormal basis consisting of  $|e_j^s\rangle$ , the eigenvectors of the hermitian operator. The representation of the hermitian operator in the new basis is  $\mathbf{U}\mathbf{H}\mathbf{U}^\dagger$ , as discussed in Section 3.3. However, the above argument showed that the new matrix is diagonal. We therefore have the following result.

**4.4.10. Corollary.** *A hermitian matrix can always be brought to diagonal form by means of a unitary transformation matrix.*

**4.4.11. Example.** Let us consider the diagonalization of the hermitian matrix

$$\mathbf{H} = \begin{pmatrix} 0 & 0 & -1+i & -1-i \\ 0 & 0 & -1+i & 1+i \\ -1-i & -1-i & 0 & 0 \\ -1+i & 1-i & 0 & 0 \end{pmatrix}.$$

The characteristic polynomial is  $\det(H - \lambda I) = (\lambda + 2)^2(\lambda - 2)^2$ . Thus,  $\lambda_1 = -2$  with multiplicity  $m_1 = 2$ , and  $\lambda_2 = 2$  with multiplicity  $m_2 = 2$ . To find the eigenvectors, we first look at the matrix equation  $(H + 2I)|a\rangle = 0$ , or

$$\begin{pmatrix} 2 & 0 & -1+i & -1-i \\ 0 & 2 & -1+i & 1+i \\ -1-i & -1-i & 2 & 0 \\ -1+i & 1-i & 0 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} = 0.$$

This is a system of linear equations whose "solution" is

$$\alpha_3 = \frac{1}{2}(1+i)(\alpha_1 + \alpha_2), \quad \alpha_4 = \frac{1}{2}(1-i)(\alpha_1 - \alpha_2).$$

We have two arbitrary parameters, so we expect two linearly independent solutions. For the two choices  $\alpha_1 = 2, \alpha_2 = 0$  and  $\alpha_1 = 0, \alpha_2 = 2$ , we obtain, respectively,

$$|a_1\rangle = \begin{pmatrix} 2 \\ 0 \\ 1+i \\ 1-i \end{pmatrix} \quad \text{and} \quad |a_2\rangle = \begin{pmatrix} 0 \\ 2 \\ 1+i \\ -1+i \end{pmatrix},$$

which happen to be orthogonal. We simply normalize them to obtain

$$|e_1\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} 2 \\ 0 \\ 1+i \\ 1-i \end{pmatrix} \quad \text{and} \quad |e_2\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} 0 \\ 2 \\ 1+i \\ -1+i \end{pmatrix}.$$

Similarly, the second eigenvalue equation,  $(H - 2I)|a\rangle = 0$ , gives rise to the conditions  $\alpha_3 = -\frac{1}{2}(1+i)(\alpha_1 + \alpha_2)$  and  $\alpha_4 = -\frac{1}{2}(1-i)(\alpha_1 - \alpha_2)$ , which produce the orthonormal vectors

$$|e_3\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} 2 \\ 0 \\ -1-i \\ -1+i \end{pmatrix} \quad \text{and} \quad |e_4\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} 0 \\ 2 \\ -1-i \\ 1-i \end{pmatrix}.$$

The unitary matrix that diagonalizes  $H$  can be constructed from these column vectors using the remarks before Example 3.4.4, which imply that if we simply put the vectors  $|e_i\rangle$  together as columns, the resulting matrix is  $U^\dagger$ :

$$U^\dagger = \frac{1}{2\sqrt{2}} \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 1+i & 1+i & -1-i & -1-i \\ 1-i & -1+i & -1+i & 1-i \end{pmatrix},$$

and the unitary matrix will be

$$U = (U^\dagger)^\dagger = \frac{1}{2\sqrt{2}} \begin{pmatrix} 2 & 0 & 1-i & 1+i \\ 0 & 2 & 1-i & -1-i \\ 2 & 0 & -1+i & -1-i \\ 0 & 2 & -1+i & 1+i \end{pmatrix}.$$

We can easily check that  $U$  diagonalizes  $H$ , i.e., that  $UHU^\dagger$  is diagonal. ■

application of  
diagonalization in  
electromagnetism

**4.4.12. Example.** In some physical applications the ability to diagonalize matrices can be very useful. As a simple but illustrative example, let us consider the motion of a charged particle in a constant magnetic field pointing in the  $z$  direction. The equation of motion for such a particle is

$$m \frac{d\mathbf{v}}{dt} = q\mathbf{v} \times \mathbf{B} = q \det \begin{pmatrix} \hat{\mathbf{e}}_x & \hat{\mathbf{e}}_y & \hat{\mathbf{e}}_z \\ v_x & v_y & v_z \\ 0 & 0 & B \end{pmatrix},$$

which in component form becomes

$$\frac{dv_x}{dt} = \frac{qB}{m} v_y, \quad \frac{dv_y}{dt} = -\frac{qB}{m} v_x, \quad \frac{dv_z}{dt} = 0.$$

Ignoring the uniform motion in the  $z$  direction, we need to solve the first two coupled equations, which in matrix form becomes

$$\frac{d}{dt} \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \frac{qB}{m} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} = -i\omega \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix}, \quad (4.10)$$

where we have introduced a factor of  $i$  to render the matrix hermitian, and defined  $\omega = qB/m$ . If the  $2 \times 2$  matrix were diagonal, we would get two *uncoupled* equations, which we could solve easily. Diagonalizing the matrix involves finding a matrix  $R$  such that

$$D = R \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} R^{-1} = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

If we could do such a diagonalization, we would multiply (4.10) by  $R$  to get<sup>5</sup>

$$\frac{d}{dt} R \begin{pmatrix} v_x \\ v_y \end{pmatrix} = -i\omega R \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} R^{-1} R \begin{pmatrix} v_x \\ v_y \end{pmatrix},$$

which can be written as

$$\frac{d}{dt} \begin{pmatrix} v'_x \\ v'_y \end{pmatrix} = -i\omega \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} \begin{pmatrix} v'_x \\ v'_y \end{pmatrix} = \begin{pmatrix} -i\omega\mu_1 v'_x \\ -i\omega\mu_2 v'_y \end{pmatrix}, \quad \text{where} \quad \begin{pmatrix} v'_x \\ v'_y \end{pmatrix} \equiv R \begin{pmatrix} v_x \\ v_y \end{pmatrix}.$$

<sup>5</sup>The fact that  $R$  is independent of  $t$  is crucial in this step. This fact, in turn, is a consequence of the independence from  $t$  of the original  $2 \times 2$  matrix.

We then would have a pair of uncoupled equations

$$\frac{dv'_x}{dt} = -i\omega\mu_1 v'_x, \quad \frac{dv'_y}{dt} = -i\omega\mu_2 v'_y$$

that have  $v'_x = v'_{0x} e^{-i\omega\mu_1 t}$  and  $v'_y = v'_{0y} e^{-i\omega\mu_2 t}$  as a solution set, in which  $v'_{0x}$  and  $v'_{0y}$  are integration constants.

To find  $R$ , we need the normalized eigenvectors of  $\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$ . But these are obtained in precisely the same fashion as in Example 4.4.4. There is, however, an arbitrariness in the solutions due to the choice in numbering the eigenvalues. If we choose the normalized eigenvectors

$$|e_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad |e_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \end{pmatrix},$$

then from comments at the end of Section 3.3, we get

$$R^{-1} = R^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \quad \Rightarrow \quad R = (R^\dagger)^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix}.$$

With this choice of  $R$ , we have

$$R \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} R^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

so that  $\mu_1 = 1 = -\mu_2$ . Having found  $R^\dagger$ , we can write

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = R^\dagger \begin{pmatrix} v'_x \\ v'_y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v'_{0x} e^{-i\omega t} \\ v'_{0y} e^{i\omega t} \end{pmatrix}. \tag{4.11}$$

If the  $x$  and  $y$  components of velocity at  $t = 0$  are  $v_{0x}$  and  $v_{0y}$ , respectively, then

$$\begin{pmatrix} v_{0x} \\ v_{0y} \end{pmatrix} = R^\dagger \begin{pmatrix} v'_{0x} \\ v'_{0y} \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} v'_{0x} \\ v'_{0y} \end{pmatrix} = R \begin{pmatrix} v_{0x} \\ v_{0y} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -i v_{0x} + v_{0y} \\ i v_{0x} + v_{0y} \end{pmatrix}.$$

Substituting in (4.11), we obtain

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = \frac{1}{2} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (-i v_{0x} + v_{0y}) e^{-i\omega t} \\ (i v_{0x} + v_{0y}) e^{i\omega t} \end{pmatrix} = \begin{pmatrix} v_{0x} \cos \omega t + v_{0y} \sin \omega t \\ -v_{0x} \sin \omega t + v_{0y} \cos \omega t \end{pmatrix}.$$

This gives the velocity as a function of time. Antidifferentiating once with respect to time yields the position vector. ■

In many situations of physical interest, it is desirable to know whether two operators are simultaneously diagonalizable. For instance, if there exists a basis of a Hilbert space of a quantum-mechanical system consisting of simultaneous eigenvectors of two operators, then one can measure those two operators at the same time. In particular, they are not restricted by an uncertainty relation.

simultaneous  
diagonalization  
defined

**4.4.13. Definition.** Two operators are said to be *simultaneously diagonalizable* if they can be written in terms of the same set of projection operators, as in Theorem 4.4.6.

This definition is consistent with the matrix representation of the two operators, because if we take the orthonormal basis  $B = \{e_j^s\}$  discussed right after Theorem 4.4.6, we obtain diagonal matrices for both operators. What are the conditions under which two operators can be simultaneously diagonalized? Clearly, a necessary condition is that the two operators commute. This is an immediate consequence of the orthogonality of the projection operators, which trivially implies  $P_i P_j = P_j P_i$  for all  $i$  and  $j$ . It is also apparent in the matrix representation of the operators: Any two diagonal matrices commute. What about sufficiency? Is the commutativity of the two operators sufficient for them to be simultaneously diagonalizable? To answer this question, we need the following lemma:

**4.4.14. Lemma.** *An operator  $T$  commutes with a normal operator  $A$  if and only if  $T$  commutes with all the projection operators of  $A$ .*

*Proof.* The “if” part is trivial. To prove the “only if” part, suppose  $AT = TA$ , and let  $|x\rangle$  be any vector in one of the eigenspaces of  $A$ , say  $\mathcal{M}_j$ . Then we have  $A(T|x\rangle) = T(A|x\rangle) = T(\lambda_j|x\rangle) = \lambda_j(T|x\rangle)$ ; i.e.,  $T|x\rangle$  is in  $\mathcal{M}_j$ , or  $\mathcal{M}_j$  is invariant under  $T$ . Since  $\mathcal{M}_j$  is arbitrary,  $T$  leaves all eigenspaces invariant. In particular, it leaves  $\mathcal{M}_j^\perp$ , the orthogonal complement of  $\mathcal{M}_j$  (the direct sum of all the remaining eigenspaces), invariant. By Theorems 4.2.3 and 4.2.5,  $TP_j = P_jT$ ; and this holds for all  $j$ .  $\square$

necessary and sufficient condition for simultaneous diagonalizability

**4.4.15. Theorem.** *A necessary and sufficient condition for two normal operators  $A$  and  $B$  to be simultaneously diagonalizable is  $[A, B] = 0$ .*

*Proof.* As claimed above, the “necessity” is trivial. To prove the “sufficiency,” let  $A = \sum_{j=1}^r \lambda_j P_j$  and  $B = \sum_{k=1}^s \mu_k Q_k$ , where  $\{\lambda_j\}$  and  $\{P_j\}$  are eigenvalues and projections of  $A$ , and  $\{\mu_k\}$  and  $\{Q_k\}$  are those of  $B$ . Assume  $[A, B] = 0$ . Then by Lemma 4.4.14,  $AQ_k = Q_kA$ . Since  $Q_k$  commutes with  $A$ , it must commute with the latter’s projection operators:  $P_j Q_k = Q_k P_j$ . To complete the proof, define  $R_{jk} \equiv P_j Q_k$ , and note that

$$\begin{aligned} R_{jk}^\dagger &= (P_j Q_k)^\dagger = Q_k^\dagger P_j^\dagger = Q_k P_j = P_j Q_k = R_{jk}, \\ (R_{jk})^2 &= (P_j Q_k)^2 = P_j Q_k P_j Q_k = P_j P_j Q_k Q_k = P_j Q_k = R_{jk}. \end{aligned}$$

Therefore,  $R_{jk}$  are hermitian projection operators. Furthermore,

$$\sum_{j=1}^r R_{jk} = \underbrace{\sum_{j=1}^r P_j}_{=1} Q_k = Q_k.$$

Similarly,  $\sum_{k=1}^s R_{jk} = \sum_{k=1}^s P_j Q_k = P_j \sum_{k=1}^s Q_k = P_j$ . We can now write  $A$  and  $B$  as

$$A = \sum_{j=1}^r \lambda_j P_j = \sum_{j=1}^r \sum_{k=1}^s \lambda_j R_{jk}, \quad B = \sum_{k=1}^s \mu_k Q_k = \sum_{k=1}^s \sum_{j=1}^r \mu_k R_{jk}.$$



By definition, they are simultaneously diagonalizable. □

spectral  
decomposition of a  
Pauli spin matrix

**4.4.16. Example.** Let us find the spectral decomposition of the Pauli spin matrix

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

The eigenvalues and eigenvectors have been found in Example 4.4.4. These are

$$\lambda_1 = 1, \quad |e_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{and} \quad \lambda_2 = -1, \quad |e_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

The subspaces  $\mathcal{M}_{\lambda_j}$  are one-dimensional; therefore,

$$P_1 = |e_1\rangle \langle e_1| = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \frac{1}{\sqrt{2}} (1 \quad -i) = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix},$$

$$P_2 = |e_2\rangle \langle e_2| = \frac{1}{2} \begin{pmatrix} 1 \\ -i \end{pmatrix} (1 \quad i) = \frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

We can check that  $P_1 + P_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  and

$$\lambda_1 P_1 + \lambda_2 P_2 = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_2. \quad \blacksquare$$

What restrictions are to be imposed on the most general operator  $T$  to make it diagonalizable? We saw in Chapter 2 that  $T$  can be written in terms of its so-called Cartesian components as  $T = H + iH'$  where both  $H$  and  $H'$  are hermitian and can therefore be decomposed according to Theorem 4.4.6. Can we conclude that  $T$  is also decomposable? No. Because the projection operators used in the decomposition of  $H$  may not be the same as those used for  $H'$ . However, if  $H$  and  $H'$  are simultaneously diagonalizable such that

$$H = \sum_{k=1}^r \lambda_k P_k \quad \text{and} \quad H' = \sum_{k=1}^r \lambda'_k P_k, \tag{4.12}$$

then  $T = \sum_{k=1}^r (\lambda_k + i\lambda'_k) P_k$ . It follows that  $T$  has a spectral decomposition, and therefore is diagonalizable. Theorem 4.4.15 now implies that  $H$  and  $H'$  must commute. Since,  $H = \frac{1}{2}(T + T^\dagger)$  and  $H' = \frac{1}{2i}(T - T^\dagger)$ , we have  $[H, H'] = 0$  if and only if  $[T, T^\dagger] = 0$ ; i.e.,  $T$  is normal. We thus get back to the condition with which we started the whole discussion of spectral decomposition in this section.

## 4.5 Functions of Operators

Functions of transformations were discussed in Chapter 2. With the power of spectral decomposition at our disposal, we can draw many important conclusions about them.

First, we note that if  $\mathbf{T} = \sum_{i=1}^r \lambda_i \mathbf{P}_i$ , then, because of orthogonality of the  $\mathbf{P}_i$ 's

$$\mathbf{T}^2 = \sum_{i=1}^r \lambda_i^2 \mathbf{P}_i, \dots, \mathbf{T}^n = \sum_{i=1}^r \lambda_i^n \mathbf{P}_i.$$

Thus, any polynomial  $p$  in  $\mathbf{T}$  has a spectral decomposition given by  $p(\mathbf{T}) = \sum_{i=1}^r p(\lambda_i) \mathbf{P}_i$ . Generalizing this to functions expandable in power series gives

$$f(\mathbf{T}) = \sum_{i=1}^{\infty} f(\lambda_i) \mathbf{P}_i. \tag{4.13}$$

**4.5.1. Example.** Let us investigate the spectral decomposition of the following unitary (actually orthogonal) matrix:

$$\mathbf{U} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

We find the eigenvalues

$$\det \begin{pmatrix} \cos \theta - \lambda & -\sin \theta \\ \sin \theta & \cos \theta - \lambda \end{pmatrix} = \lambda^2 - 2 \cos \theta \lambda + 1 = 0,$$

yielding  $\lambda_1 = e^{-i\theta}$  and  $\lambda_2 = e^{i\theta}$ . For  $\lambda_1$  we have (reader, provide the missing steps)

$$\begin{pmatrix} \cos \theta - e^{-i\theta} & -\sin \theta \\ \sin \theta & \cos \theta - e^{-i\theta} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = 0 \Rightarrow \alpha_2 = i\alpha_1 \Rightarrow |e_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix},$$

and for  $\lambda_2$ ,

$$\begin{pmatrix} \cos \theta - e^{i\theta} & -\sin \theta \\ \sin \theta & \cos \theta - e^{i\theta} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = 0 \Rightarrow \alpha_2 = -i\alpha_1 \Rightarrow |e_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

We note that the  $\mathcal{M}_{\lambda_j}$  are one-dimensional and spanned by  $|e_j\rangle$ . Thus,

$$\mathbf{P}_1 = |e_1\rangle \langle e_1| = \frac{1}{2} \begin{pmatrix} 1 \\ i \end{pmatrix} \begin{pmatrix} 1 & -i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix},$$

$$\mathbf{P}_2 = |e_2\rangle \langle e_2| = \frac{1}{2} \begin{pmatrix} 1 \\ -i \end{pmatrix} \begin{pmatrix} 1 & i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

Clearly,  $\mathbf{P}_1 + \mathbf{P}_2 = \mathbf{1}$ , and

$$e^{-i\theta} \mathbf{P}_1 + e^{i\theta} \mathbf{P}_2 = \frac{1}{2} \begin{pmatrix} e^{-i\theta} & -ie^{-i\theta} \\ ie^{-i\theta} & e^{-i\theta} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} e^{i\theta} & ie^{i\theta} \\ -ie^{i\theta} & e^{i\theta} \end{pmatrix} = \mathbf{U}.$$

If we take the natural log of this equation and use Equation (4.13), we obtain

$$\begin{aligned} \ln \mathbf{U} &= \ln(e^{-i\theta}) \mathbf{P}_1 + \ln(e^{i\theta}) \mathbf{P}_2 = -i\theta \mathbf{P}_1 + i\theta \mathbf{P}_2 \\ &= i(-\theta \mathbf{P}_1 + \theta \mathbf{P}_2) \equiv i\mathbf{H}, \end{aligned} \tag{4.14}$$

where  $\mathbf{H} \equiv -\theta \mathbf{P}_1 + \theta \mathbf{P}_2$  is a hermitian operator because  $\theta$  is real and  $\mathbf{P}_1$  and  $\mathbf{P}_2$  are hermitian. Inverting Equation (4.14) gives  $\mathbf{U} = e^{i\mathbf{H}}$ , where

$$\mathbf{H} = \theta(-\mathbf{P}_1 + \mathbf{P}_2) = \theta \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \quad \blacksquare$$

The example above shows that the unitary  $2 \times 2$  matrix  $U$  can be written as an exponential of an anti-hermitian operator. This is a general result. In fact, we have the following theorem, whose proof is left as an exercise for the reader (see Problem 4.22).

**4.5.2. Theorem.** *A unitary operator  $U$  on a finite-dimensional complex inner product space can be written as  $U = e^{iH}$  where  $H$  is hermitian. Furthermore, a unitary matrix can be brought to diagonal form by a unitary transformation matrix.*

The last statement follows from Corollary 4.4.10 and the fact that

$$f(\mathbf{RHR}^{-1}) = \mathbf{R}f(\mathbf{H})\mathbf{R}^{-1}$$

for any function  $f$  that can be expanded in a Taylor series.

The square root of an operator is plagued by multivaluedness. In the real numbers, we have only two-valuedness!

A useful function of an operator is its square root. A natural way to define the square root of an operator  $A$  is  $\sqrt{A} = \sum_{i=1}^r (\pm\sqrt{\lambda_i})P_i$ . This clearly gives many candidates for the root, because each term in the sum can have either the plus sign or the minus sign.

**4.5.3. Definition.** *The positive square root of a positive operator  $A = \sum_{i=1}^r \lambda_i P_i$  is  $\sqrt{A} = \sum_{i=1}^r \sqrt{\lambda_i} P_i$ .*

The uniqueness of the spectral decomposition implies that the positive square root of a positive operator is unique.

**4.5.4. Example.** Let us evaluate  $\sqrt{A}$  where

$$A = \begin{pmatrix} 5 & 3i \\ -3i & 5 \end{pmatrix}.$$

First, we have to spectrally decompose  $A$ . Its characteristic equation is

$$\lambda^2 - 10\lambda + 16 = 0,$$

with roots  $\lambda_1 = 8$  and  $\lambda_2 = 2$ . Since both eigenvalues are positive and  $A$  is hermitian, we conclude that  $A$  is indeed positive (Corollary 4.4.8). We can also easily find its normalized eigenvectors:

$$|e_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix} \quad \text{and} \quad |e_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

Thus,

$$P_1 = |e_1\rangle\langle e_1| = \frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}, \quad P_2 = |e_2\rangle\langle e_2| = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix},$$

and

$$\sqrt{A} = \sqrt{\lambda_1}P_1 + \sqrt{\lambda_2}P_2 = \sqrt{8}\frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} + \sqrt{2}\frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix}.$$

We can easily check that  $(\sqrt{A})^2 = A$ . ■

Intuitively, higher and higher powers of  $\mathbf{T}$ , when acting on a few vectors of the space, eventually exhaust all vectors, and further increase in power will be a repetition of lower powers. This intuitive idea can be made more precise by looking at the projection operators. We have already seen that  $\mathbf{T}^n = \sum_{j=1}^r \lambda_j^n \mathbf{P}_j$ ,  $n = 1, 2, \dots$ . For various  $n$ 's one can "solve" for  $\mathbf{P}_j$  in terms of powers of  $\mathbf{T}$ . Since there are only a finite number of  $\mathbf{P}_j$ 's, only a finite number of powers of  $\mathbf{T}$  will suffice. In fact, we can explicitly construct the polynomial in  $\mathbf{T}$  for  $\mathbf{P}_j$ . If there is such a polynomial, by Equation (4.13) it must satisfy  $\mathbf{P}_j = p_j(\mathbf{T}) = \sum_{k=1}^r p_j(\lambda_k) \mathbf{P}_k$ , where  $p_j$  is some polynomial to be determined. By orthogonality of the projection operators,  $p_j(\lambda_k)$  must be zero unless  $k = j$ , in which case it must be 1. In other words,  $p_j(\lambda_k) = \delta_{kj}$ . Such a polynomial can be explicitly constructed:

$$p_j(x) = \left( \frac{x - \lambda_1}{\lambda_j - \lambda_1} \right) \left( \frac{x - \lambda_2}{\lambda_j - \lambda_2} \right) \cdots \left( \frac{x - \lambda_r}{\lambda_j - \lambda_r} \right) \equiv \prod_{k \neq j} \frac{x - \lambda_k}{\lambda_j - \lambda_k}.$$

Therefore,

$$\mathbf{P}_j = p_j(\mathbf{T}) \equiv \prod_{k \neq j} \frac{\mathbf{T} - \lambda_k \mathbf{1}}{\lambda_j - \lambda_k}, \quad (4.15)$$

and we have the following result.

**4.5.5. Proposition.** *Every function of a normal operator on a finite-dimensional vector space can be expressed as a polynomial. In fact, from Equations (4.13) and (4.15),*

$$f(\mathbf{T}) = \sum_{j=1}^r f(\lambda_j) \mathbf{P}_j = \sum_{j=1}^r f(\lambda_j) \prod_{k \neq j} \frac{\mathbf{T} - \lambda_k \mathbf{1}}{\lambda_j - \lambda_k}. \quad (4.16)$$

**4.5.6. Example.** Let us write  $\sqrt{A}$  of the last example as a polynomial in  $A$ . We have

$$p_1(A) = \prod_{k \neq 1} \frac{A - \lambda_k \mathbf{1}}{\lambda_1 - \lambda_k} = \frac{A - \lambda_2 \mathbf{1}}{\lambda_1 - \lambda_2} = \frac{1}{6}(A - 2),$$

$$p_2(A) = \prod_{k \neq 2} \frac{A - \lambda_k \mathbf{1}}{\lambda_2 - \lambda_k} = \frac{A - \lambda_1 \mathbf{1}}{\lambda_2 - \lambda_1} = -\frac{1}{6}(A - 8).$$

Substituting in Equation (4.16), we obtain

$$\sqrt{A} = \sqrt{\lambda_1} p_1(A) + \sqrt{\lambda_2} p_2(A) = \frac{\sqrt{8}}{6}(A - 2) - \frac{\sqrt{2}}{6}(A - 8) = \frac{\sqrt{2}}{6}A + \frac{\sqrt{8}}{3}.$$

The RHS is clearly a (first-degree) polynomial in  $A$ , and it is easy to verify that it is the matrix of  $\sqrt{A}$  obtained in the previous example. ■

## 4.6 Polar Decomposition

We have seen many similarities between operators and complex numbers. For instance, hermitian operators behave very much like the real numbers: they have real eigenvalues; their squares are positive; every operator can be written as  $\mathbf{H} + i\mathbf{H}'$ , where both  $\mathbf{H}$  and  $\mathbf{H}'$  are hermitian; and so forth. Also, unitary operators can be written as  $\exp i\mathbf{H}$ , where  $\mathbf{H}$  is hermitian. So unitary operators are the analogue of complex numbers of unit magnitude such as  $e^{i\theta}$ . A general complex number can be written as  $re^{i\theta}$ . Can we write an arbitrary operator in an analogous way? The following theorem provides the answer.

polar decomposition theorem

**4.6.1. Theorem.** (polar decomposition theorem) *An operator  $\mathbf{A}$  on a finite-dimensional complex inner product space can be written as  $\mathbf{A} = \mathbf{U}\mathbf{R}$  where  $\mathbf{R}$  is a (unique) positive operator and  $\mathbf{U}$  a unitary operator. If  $\mathbf{A}$  is invertible, then  $\mathbf{U}$  is also unique.*

*Proof.* We will prove the theorem for the case where the operator is invertible. The proof of the general case can be found in books on linear algebra (such as [Halm 58]).

The reader may show that the operator  $\mathbf{A}^\dagger\mathbf{A}$  is positive. Therefore, it has a unique positive square root  $\mathbf{R}$ . We let  $\mathbf{V} = \mathbf{R}\mathbf{A}^{-1}$ , or  $\mathbf{V}\mathbf{A} = \mathbf{R}$ . Then

$$\begin{aligned} \mathbf{V}\mathbf{V}^\dagger &= \mathbf{R}\mathbf{A}^{-1}(\mathbf{R}\mathbf{A}^{-1})^\dagger = \mathbf{R}\mathbf{A}^{-1}(\mathbf{A}^{-1})^\dagger\mathbf{R}^\dagger = \mathbf{R}(\mathbf{A}^\dagger\mathbf{A})^{-1}\mathbf{R}^\dagger \\ &= \mathbf{R}(\mathbf{R}^2)^{-1}\mathbf{R}^\dagger = \mathbf{R}(\mathbf{R}^\dagger\mathbf{R})^{-1}\mathbf{R}^\dagger = \mathbf{R}\mathbf{R}^{-1}(\mathbf{R}^\dagger)^{-1}\mathbf{R}^\dagger = \mathbf{1} \end{aligned}$$

and

$$\mathbf{V}^\dagger\mathbf{V} = (\mathbf{R}\mathbf{A}^{-1})^\dagger\mathbf{R}\mathbf{A}^{-1} = (\mathbf{A}^\dagger)^{-1}\underbrace{\mathbf{R}^\dagger\mathbf{R}}_{\mathbf{R}^2}\mathbf{A}^{-1} = (\mathbf{A}^\dagger)^{-1}\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^{-1} = \mathbf{1},$$

and  $\mathbf{V}$  is indeed a unitary operator. Now choose  $\mathbf{U} = \mathbf{V}^\dagger$  to get the desired decomposition.

To prove uniqueness we note that  $\mathbf{U}\mathbf{R} = \mathbf{U}'\mathbf{R}'$  implies that  $\mathbf{R} = \mathbf{U}^\dagger\mathbf{U}'\mathbf{R}'$  and  $\mathbf{R}^2 = \mathbf{R}^\dagger\mathbf{R} = (\mathbf{U}^\dagger\mathbf{U}'\mathbf{R}')^\dagger(\mathbf{U}^\dagger\mathbf{U}'\mathbf{R}') = \mathbf{R}'^\dagger\mathbf{U}'^\dagger\mathbf{U}\mathbf{U}'\mathbf{R}' = \mathbf{R}'^\dagger\mathbf{R}' = \mathbf{R}'^2$ . Since the positive transformation  $\mathbf{R}^2$  (or  $\mathbf{R}'^2$ ) has only one positive square root, it follows that  $\mathbf{R} = \mathbf{R}'$ .

If  $\mathbf{A}$  is invertible, then so is  $\mathbf{R} = \mathbf{U}^\dagger\mathbf{A}$ . Therefore,

$$\mathbf{U}\mathbf{R} = \mathbf{U}'\mathbf{R} \Rightarrow \mathbf{U}\mathbf{R}\mathbf{R}^{-1} = \mathbf{U}'\mathbf{R}\mathbf{R}^{-1} \Rightarrow \mathbf{U} = \mathbf{U}',$$

and  $\mathbf{U}$  is also unique. □

It is interesting to note that the positive definiteness of  $\mathbf{R}$  and the nonuniqueness of  $\mathbf{U}$  are the analogue of the positivity of  $r$  and the nonuniqueness of  $e^{i\theta}$  in the polar representation of complex numbers:

$$z = re^{i\theta} = re^{i(\theta+2n\pi)} \quad \forall n \in \mathbb{Z}.$$

In practice,  $\mathbf{R}$  is found by spectrally decomposing  $\mathbf{A}^\dagger\mathbf{A}$  and taking its positive square root.<sup>6</sup> Once  $\mathbf{R}$  is found,  $\mathbf{U}$  can be calculated from the definition  $\mathbf{A} = \mathbf{U}\mathbf{R}$ .

**4.6.2. Example.** Let us find the polar decomposition of

$$\mathbf{A} = \begin{pmatrix} -2i & \sqrt{7} \\ 0 & 3 \end{pmatrix}.$$

We have

$$\mathbf{R}^2 = \mathbf{A}^\dagger\mathbf{A} = \begin{pmatrix} 2i & 0 \\ \sqrt{7} & 3 \end{pmatrix} \begin{pmatrix} -2i & \sqrt{7} \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 2i\sqrt{7} \\ -2i\sqrt{7} & 16 \end{pmatrix}.$$

The eigenvalues and eigenvectors of  $\mathbf{R}^2$  are routinely found to be

$$\lambda_1 = 18, \quad \lambda_2 = 2, \quad |e_1\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} i \\ \sqrt{7} \end{pmatrix}, \quad |e_2\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} \sqrt{7} \\ i \end{pmatrix}.$$

The projection matrices are

$$\mathbf{P}_1 = |e_1\rangle\langle e_1| = \frac{1}{8} \begin{pmatrix} 1 & i\sqrt{7} \\ -i\sqrt{7} & 7 \end{pmatrix}, \quad \mathbf{P}_2 = |e_2\rangle\langle e_2| = \frac{1}{8} \begin{pmatrix} 7 & -i\sqrt{7} \\ i\sqrt{7} & 1 \end{pmatrix}.$$

Thus,

$$\mathbf{R} = \sqrt{\lambda_1}\mathbf{P}_1 + \sqrt{\lambda_2}\mathbf{P}_2 = \frac{1}{4} \begin{pmatrix} 5\sqrt{2} & i\sqrt{14} \\ -i\sqrt{14} & 11\sqrt{2} \end{pmatrix}.$$

To find  $\mathbf{U}$ , we note that  $\det \mathbf{A}$  is nonzero. Hence,  $\mathbf{A}$  is invertible, which implies that  $\mathbf{R}$  is also invertible. The inverse of  $\mathbf{R}$  is

$$\mathbf{R}^{-1} = \frac{1}{24} \begin{pmatrix} 11\sqrt{2} & -i\sqrt{14} \\ i\sqrt{14} & 5\sqrt{2} \end{pmatrix}.$$

The unitary matrix is simply

$$\mathbf{U} = \mathbf{A}\mathbf{R}^{-1} = \frac{1}{24} \begin{pmatrix} -i15\sqrt{2} & 3\sqrt{14} \\ 3i\sqrt{14} & 15\sqrt{2} \end{pmatrix}.$$

It is left for the reader to verify that  $\mathbf{U}$  is indeed unitary. ■

## 4.7 Real Vector Spaces

The treatment so far in this chapter has focused on complex inner product spaces. The complex number system is far more complete than the real numbers. For example, in preparation for the proof of the spectral decomposition theorem, we used the existence of  $n$  roots of a polynomial of degree  $n$  over the complex field

<sup>6</sup>It is important to pay attention to the order of the two operators: One decomposes  $\mathbf{A}^\dagger\mathbf{A}$ , not  $\mathbf{A}\mathbf{A}^\dagger$ .

(this is the fundamental theorem of algebra). A polynomial over the reals, on the other hand, does not necessarily have all its roots in the real number system.

It may therefore seem that vector spaces over the reals will not satisfy the useful theorems and results developed for complex spaces. However, through a process called *complexification* of a real vector space, in which an imaginary part is added to such a space, it is possible to prove (see, for example, [Halm 58]) practically all the results obtained for complex vector spaces. Only the results are given here.

**4.7.1. Theorem.** *A real symmetric operator has a spectral decomposition as stated in Theorem 4.4.6.*

This theorem is especially useful in applications of classical physics, which deal mostly with real vector spaces. A typical situation involves a vector that is related to another vector by a symmetric matrix. It is then convenient to find a coordinate system in which the two vectors are related in a simple manner. This involves diagonalizing the symmetric matrix by a rotation (a real orthogonal matrix). Theorem 4.7.1 reassures us that such a diagonalization is possible.

**4.7.2. Example.** For a system of  $N$  point particles constituting a rigid body, the total angular momentum  $\mathbf{L} = \sum_{i=1}^N m_i (\mathbf{r}_i \times \mathbf{v}_i)$  is related to the angular frequency via  $\mathbf{L} = \sum_{i=1}^N m_i [\mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i)] = \sum_{i=1}^N m_i [\boldsymbol{\omega} \mathbf{r}_i \cdot \mathbf{r}_i - \mathbf{r}_i (\mathbf{r}_i \cdot \boldsymbol{\omega})]$ , or

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix},$$

where

$$\begin{aligned} I_{xx} &= \sum_{i=1}^N m_i (r_i^2 - x_i^2), & I_{yy} &= \sum_{i=1}^N m_i (r_i^2 - y_i^2), & I_{zz} &= \sum_{i=1}^N m_i (r_i^2 - z_i^2), \\ I_{xy} &= -\sum_{i=1}^N m_i x_i y_i, & I_{xz} &= -\sum_{i=1}^N m_i x_i z_i, & I_{yz} &= -\sum_{i=1}^N m_i y_i z_i, \end{aligned}$$

with  $I_{xy} = I_{yx}$ ,  $I_{xz} = I_{zx}$ , and  $I_{yz} = I_{zy}$ .

The  $3 \times 3$  matrix is denoted by  $I$  and is called the *moment of inertia* matrix. It is symmetric, and Theorem 4.7.1 permits its diagonalization by an orthogonal transformation (the counterpart of a unitary transformation in a real vector space). But an orthogonal transformation in three dimensions is merely a rotation of coordinates.<sup>7</sup> Thus, Theorem 4.7.1 says that it is always possible to choose coordinate systems in which the moment of inertia matrix is diagonal. In such a coordinate system we have  $L_x = I_{xx}\omega_x$ ,  $L_y = I_{yy}\omega_y$ , and  $L_z = I_{zz}\omega_z$ , simplifying the equations considerably.

Similarly, the kinetic energy of the rigid rotating body,

$$T = \sum_{i=1}^N \frac{1}{2} m_i v_i^2 = \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i \cdot (\boldsymbol{\omega} \times \mathbf{r}_i) = \sum_{i=1}^N \frac{1}{2} m_i \boldsymbol{\omega} \cdot (\mathbf{r}_i \times \mathbf{v}_i) = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{L} = \frac{1}{2} \omega^t I \boldsymbol{\omega},$$

<sup>7</sup>This is not entirely true! There are orthogonal transformations that are composed of a rotation followed by a reflection about the origin. See Example 3.5.8.

which in general has off-diagonal terms involving  $I_{xy}$  and so forth, reduces to a simple form:  $T = \frac{1}{2}I_{xx}\omega_x^2 + \frac{1}{2}I_{yy}\omega_y^2 + \frac{1}{2}I_{zz}\omega_z^2$ . ■

**4.7.3. Example.** Another application of Theorem 4.7.1 is in the study of conic sections. The most general form of the equation of a conic section is

$$a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6 = 0,$$

where  $a_1, \dots, a_6$  are constants. If the coordinate axes coincide with the principal axes of the conic section, the  $xy$  term will be absent, and the equation of the conic section takes the familiar form. On geometrical grounds we have to be able to rotate  $xy$ -coordinates to coincide with the principal axes. We shall do this using the ideas discussed in this chapter.

First, we note that the general equation for a conic section can be written in matrix form as

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a_1 & a_3/2 \\ a_3/2 & a_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_4 & a_5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + a_6 = 0.$$

The  $2 \times 2$  matrix is symmetric and can therefore be diagonalized by means of an orthogonal matrix  $R$ . Then  $R^t R = 1$ , and we can write

$$\begin{pmatrix} x & y \end{pmatrix} R^t R \begin{pmatrix} a_1 & a_3/2 \\ a_3/2 & a_2 \end{pmatrix} R R^t \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_4 & a_5 \end{pmatrix} R^t R \begin{pmatrix} x \\ y \end{pmatrix} + a_6 = 0.$$

Let

$$R \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}, \quad R \begin{pmatrix} a_1 & a_3/2 \\ a_3/2 & a_2 \end{pmatrix} R^t = \begin{pmatrix} a'_1 & 0 \\ 0 & a'_2 \end{pmatrix}, \quad R \begin{pmatrix} a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} a'_4 \\ a'_5 \end{pmatrix}.$$

Then we get

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} a'_1 & 0 \\ 0 & a'_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} a'_4 & a'_5 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + a_6 = 0;$$

or

$$a'_1x'^2 + a'_2y'^2 + a'_4x' + a'_5y' + a_6 = 0.$$

The cross term has disappeared. The orthogonal matrix  $R$  is simply a rotation. In fact, it rotates the original coordinate system to coincide with the principal axes of the conic section. ■

**4.7.4. Example.** In this example we investigate conditions under which a multivariable function has a maximum or a minimum.

A point  $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$  is a maximum (minimum) of a function

$$f(x_1, x_2, \dots, x_n) \equiv f(\mathbf{r})$$

if

$$\nabla f|_{x_i=a_i} = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)_{x_i=a_i} = 0$$



and for small  $x_i - a_i$ , the difference  $f(\mathbf{r}) - f(\mathbf{a})$  is negative (positive). To relate this difference to the topics of this section, write the Taylor expansion of the function around  $\mathbf{a}$  keeping terms up to the second order:

$$f(\mathbf{r}) = f(\mathbf{a}) + \sum_{i=1}^n (x_i - a_i) \left( \frac{\partial f}{\partial x_i} \right)_{\mathbf{r}=\mathbf{a}} + \frac{1}{2} \sum_{i,j} (x_i - a_i)(x_j - a_j) \left( \frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{\mathbf{r}=\mathbf{a}} + \dots,$$

or, constructing a column vector out of  $\delta_i \equiv x_i - a_i$  and a symmetric matrix  $D_{ij}$  out of the second derivatives, we can write

$$f(\mathbf{r}) = f(\mathbf{a}) + \frac{1}{2} \sum_{i,j} \delta_i \delta_j D_{ij} + \dots \Rightarrow f(\mathbf{r}) - f(\mathbf{a}) = \frac{1}{2} \delta^t D \delta + \dots$$

because the first derivatives vanish. For  $\mathbf{a}$  to be a minimum point of  $f$ , the RHS of the last equation must be positive for arbitrary  $\delta$ . This means that  $D$  must be a positive matrix.<sup>8</sup> Thus, all its eigenvalues must be positive (Corollary 4.4.8). Similarly, we can show that for  $\mathbf{a}$  to be a maximum point of  $f$ ,  $-D$  must be positive definite. This means that  $D$  must have negative eigenvalues.

When we specialize the foregoing discussion to two dimensions, we obtain results that are familiar from calculus. For the function  $f(x, y)$  to have a minimum, the eigenvalues of the matrix

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

must be positive. The characteristic polynomial

$$\det \begin{pmatrix} f_{xx} - \lambda & f_{xy} \\ f_{yx} & f_{yy} - \lambda \end{pmatrix} = 0 \Rightarrow \lambda^2 - (f_{xx} + f_{yy})\lambda + f_{xx}f_{yy} - f_{xy}^2 = 0$$

yields two eigenvalues:

$$\lambda_1 = \frac{f_{xx} + f_{yy} + \sqrt{(f_{xx} - f_{yy})^2 + 4f_{xy}^2}}{2},$$

$$\lambda_2 = \frac{f_{xx} + f_{yy} - \sqrt{(f_{xx} - f_{yy})^2 + 4f_{xy}^2}}{2}.$$

These eigenvalues will be both positive if

$$f_{xx} + f_{yy} > \sqrt{(f_{xx} - f_{yy})^2 + 4f_{xy}^2},$$

and both negative if

$$f_{xx} + f_{yy} < -\sqrt{(f_{xx} - f_{yy})^2 + 4f_{xy}^2}.$$

Squaring these inequalities and simplifying yields

$$f_{xx}f_{yy} > f_{xy}^2,$$

<sup>8</sup>Note that  $D$  is already symmetric—the real analogue of hermitian.

which shows that  $f_{xx}$  and  $f_{yy}$  must have *the same sign*. If they are both positive (negative), we have a minimum (maximum). This is the familiar condition for the attainment of extrema by a function of two variables. ■

Although the establishment of spectral decomposition for symmetric operators is fairly straightforward, the case of orthogonal operators (the counterpart of unitary operators in a real vector space) is more complicated. In fact, we have already seen in Example 4.5.1 that the eigenvalues of an orthogonal transformation in two dimensions are, in general, complex. This is in contrast to symmetric transformations.

Think of the orthogonal operator  $\mathbf{O}$  as a unitary operator.<sup>9</sup> Since the absolute value of the eigenvalues of a unitary operator is 1, the only real possibilities are  $\pm 1$ . To find the other eigenvalues we note that as a unitary operator,  $\mathbf{O}$  can be written as  $e^{\mathbf{A}}$ , where  $\mathbf{A}$  is anti-hermitian (see Problem 4.22). Since hermitian conjugation and transposition coincide for real vector spaces, we conclude that  $\mathbf{A} = -\mathbf{A}^t$ , and  $\mathbf{A}$  is antisymmetric. It is also real, because  $\mathbf{O}$  is.

Let us now consider the eigenvalues of  $\mathbf{A}$ . If  $\lambda$  is an eigenvalue of  $\mathbf{A}$  corresponding to the eigenvector  $|a\rangle$ , then  $\langle a|\mathbf{A}|a\rangle = \lambda \langle a|a\rangle$ . Taking the complex conjugate of both sides gives  $\langle a|\mathbf{A}^\dagger|a\rangle = \lambda^* \langle a|a\rangle$ ; but  $\mathbf{A}^\dagger = \mathbf{A}^t = -\mathbf{A}$ , because  $\mathbf{A}$  is real and antisymmetric. We therefore have  $\langle a|\mathbf{A}|a\rangle = -\lambda^* \langle a|a\rangle$ , which gives  $\lambda^* = -\lambda$ . It follows that if we restrict  $\lambda$  to be real, then it can only be zero; otherwise, it must be *purely imaginary*. Furthermore, the reader may verify that if  $\lambda$  is an eigenvalue of  $\mathbf{A}$ , so is  $-\lambda$ . Therefore, the diagonal form of  $\mathbf{A}$  looks like this:

$$\mathbf{A}_{\text{diag}} = \text{diag}(0, 0, \dots, 0, i\theta_1, -i\theta_1, i\theta_2, -i\theta_2, \dots, i\theta_k, -i\theta_k),$$

which gives  $\mathbf{O}$  the following diagonal form:

$$\mathbf{O}_{\text{diag}} = e^{\mathbf{A}_{\text{diag}}} = \text{diag}(e^0, e^0, \dots, e^0, e^{i\theta_1}, e^{-i\theta_1}, e^{i\theta_2}, e^{-i\theta_2}, \dots, e^{i\theta_k}, e^{-i\theta_k})$$

with  $\theta_1, \theta_2, \dots, \theta_k$  all real. It is clear that if  $\mathbf{O}$  has  $-1$  as an eigenvalue, then some of the  $\theta$ 's must equal  $\pm\pi$ . Separating the  $\pi$ 's from the rest of  $\theta$ 's and putting all of the above arguments together, we get

$$\mathbf{O}_{\text{diag}} = \text{diag}(\underbrace{1, 1, \dots, 1}_{N_+}, \underbrace{-1, -1, \dots, -1}_{N_-}, e^{i\theta_1}, e^{-i\theta_1}, e^{i\theta_2}, e^{-i\theta_2}, \dots, e^{i\theta_m}, e^{-i\theta_m})$$

where  $N_+ + N_- + 2m = \dim \mathbf{O}$ .

Getting insight from Example 4.5.1, we can argue, admittedly in a nonrigorous way, that corresponding to each pair  $e^{\pm i\theta_j}$  is a  $2 \times 2$  matrix of the form

$$\begin{pmatrix} \cos \theta_j & -\sin \theta_j \\ \sin \theta_j & \cos \theta_j \end{pmatrix} \equiv \mathbf{R}_2(\theta_j) \quad (4.17)$$

<sup>9</sup>This can always be done by formally identifying transposition with hermitian conjugation, an identification that holds when the underlying field of numbers is real.

We therefore have the following theorem (refer to [Halm 58] for a rigorous treatment).

**4.7.5. Theorem.** *A real orthogonal operator on a real inner product space  $\mathcal{V}$  cannot, in general, be completely diagonalized. The closest it can get to a diagonal form is*

$$O_{\text{diag}} = \text{diag}(\underbrace{1, 1, \dots, 1}_{N_+}, \underbrace{-1, -1, \dots, -1}_{N_-}, R_2(\theta_1), R_2(\theta_2), \dots, R_2(\theta_m)),$$

where  $N_+ + N_- + 2m = \dim \mathcal{V}$  and  $R_2(\theta_j)$  is as given in (4.17). Furthermore, the matrix that transforms an orthogonal matrix into the form above is itself an orthogonal matrix.

The last statement follows from Theorem 4.5.2 and the fact that an orthogonal matrix is the real analogue of a unitary matrix.

**4.7.6. Example.** An interesting application of Theorem 4.7.5 occurs in classical mechanics, where it is shown that the motion of a rigid body consists of a translation and a rotation. The rotation is represented by a  $3 \times 3$  orthogonal matrix. Theorem 4.7.5 states that by an appropriate choice of coordinate systems (i.e., by applying the same orthogonal transformation that diagonalizes the rotation matrix of the rigid body), one can “diagonalize” the  $3 \times 3$  orthogonal matrix. The “diagonal” form is

$$\begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

Excluding the reflections (corresponding to  $-1$ 's) and the trivial identity rotation, we conclude that any rotation of a rigid body can be written as

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix},$$

which is a rotation through the angle  $\theta$  about the (new)  $x$ -axis. ■

Combining the rotation of the example above with the translations, we obtain the following theorem.

**4.7.7. Theorem.** (Euler) *The general motion of a rigid body consists of the translation of one point of that body and a rotation about a single axis through that point.*

Finally, we quote the polar decomposition for real inner product spaces.

**4.7.8. Theorem.** *Any operator  $\mathbf{A}$  on a real inner product space can be written as  $\mathbf{A} = \mathbf{O}\mathbf{R}$ , where  $\mathbf{R}$  is a (unique) symmetric positive operator and  $\mathbf{O}$  is orthogonal.*

**4.7.9. Example.** Let us decompose the following matrix into its polar form:

$$A = \begin{pmatrix} 2 & 0 \\ 3 & -2 \end{pmatrix}.$$

The procedure is the same as in the complex case. We have

$$R^2 = A^t A = \begin{pmatrix} 2 & 3 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 13 & -6 \\ -6 & 4 \end{pmatrix}$$

with eigenvalues  $\lambda_1 = 1$  and  $\lambda_2 = 16$  and normalized eigenvectors

$$|e_1\rangle = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{and} \quad |e_2\rangle = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

The projection operators are

$$P_1 = |e_1\rangle \langle e_1| = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad P_2 = |e_2\rangle \langle e_2| = \frac{1}{5} \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}$$

Thus, we have

$$R = \sqrt{R^2} = \sqrt{\lambda_1} P_1 + \sqrt{\lambda_2} P_2 = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} + \frac{4}{5} \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 17 & -6 \\ -6 & 8 \end{pmatrix}.$$

We note that  $A$  is invertible. Thus,  $R$  is also invertible, and

$$R^{-1} = \frac{1}{20} \begin{pmatrix} 8 & 6 \\ 6 & 17 \end{pmatrix}.$$

This gives  $O = AR^{-1}$ , or

$$O = \frac{1}{5} \begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}.$$

It is readily verified that  $O$  is indeed orthogonal. ■

Our excursion through operator algebra and matrix theory has revealed to us the diversity of diagonalizable operators. Could it be perhaps that all operators are diagonalizable? In other words, given any operator, can we find a basis in which the matrix representing that operator is diagonal? The answer is, in general, no! (See Problem 4.27.) Discussion of this topic entails a treatment of the Hamilton–Cayley theorem and the Jordan canonical form of a matrix, in which the so-called *generalized eigenvectors* are introduced. A generalized eigenvector belongs to the kernel of  $(A - \lambda I)^m$  for some positive integer  $m$ . Then  $\lambda$  is called a *generalized eigenvalue*. We shall not pursue this matter here. The interested reader can find such a discussion in books on linear algebra and matrix theory. We shall, however, see the application of this notion to special operators on infinite-dimensional vector spaces in Chapter 16. One result is worth mentioning at this point.

**4.7.10. Proposition.** *If the roots of the characteristic polynomial of a matrix are all simple, then the matrix can be brought to diagonal form by a similarity (not necessarily a unitary) transformation.*

**4.7.11. Example.** As a final example of the application of the results of this section, let us evaluate the  $n$ -fold integral

$$I_n = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \cdots \int_{-\infty}^{\infty} dx_n e^{-\sum_{i,j=1}^n m_{ij} x_i x_j}, \tag{4.18}$$

where the  $m_{ij}$  are elements of a real, symmetric, positive definite matrix, say  $M$ . Because it is symmetric,  $M$  can be diagonalized by an orthogonal matrix  $R$  so that  $RM R^t = D$  is a diagonal matrix whose diagonal entries are the eigenvalues,  $\lambda_1, \lambda_2, \dots, \lambda_n$ , of  $M$ , whose positive definiteness ensures that none of these eigenvalues is zero or negative.

The exponent in (4.18) can be written as

$$\sum_{i,j=1}^n m_{ij} x_i x_j = \mathbf{x}^t M \mathbf{x} = \mathbf{x}^t R^t R M R^t R \mathbf{x} = \mathbf{x}^t D \mathbf{x}' = \lambda_1 x_1'^2 + \cdots + \lambda_n x_n'^2,$$

where

$$\mathbf{x}' = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = R \mathbf{x} = R \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

or, in component form,  $x_i' = \sum_{j=1}^n r_{ij} x_j$  for  $i = 1, 2, \dots, n$ . Similarly, since  $\mathbf{x} = R^t \mathbf{x}'$ , it follows that  $x_i = \sum_{j=1}^n r_{ji} x_j'$  for  $i = 1, 2, \dots, n$ .

The “volume element”  $dx_1 \cdots dx_n$  is related to the primed volume element as follows:

$$dx_1 \cdots dx_n = \left| \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(x_1', x_2', \dots, x_n')} \right| dx_1' \cdots dx_n' \equiv |\det J| dx_1' \cdots dx_n',$$

where  $J$  is the Jacobian matrix whose  $ij$ th element is  $\partial x_i / \partial x_j'$ . But

$$\frac{\partial x_i}{\partial x_j'} = r_{ji} \quad \Rightarrow \quad J = R^t \quad \Rightarrow \quad |\det J| = |\det R^t| = 1.$$

Therefore, in terms of  $x'$ , the integral  $I_n$  becomes

$$\begin{aligned} I_n &= \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_n' e^{-\lambda_1 x_1'^2 - \lambda_2 x_2'^2 - \cdots - \lambda_n x_n'^2} \\ &= \left( \int_{-\infty}^{\infty} dx_1' e^{-\lambda_1 x_1'^2} \right) \left( \int_{-\infty}^{\infty} dx_2' e^{-\lambda_2 x_2'^2} \right) \cdots \left( \int_{-\infty}^{\infty} dx_n' e^{-\lambda_n x_n'^2} \right) \\ &= \sqrt{\frac{\pi}{\lambda_1}} \sqrt{\frac{\pi}{\lambda_2}} \cdots \sqrt{\frac{\pi}{\lambda_n}} = \pi^{n/2} \frac{1}{\sqrt{\lambda_1 \lambda_2 \cdots \lambda_n}} = \pi^{n/2} (\det M)^{-1/2}, \end{aligned}$$

because the determinant of a matrix is the product of its eigenvalues. This result can be written as

analytic definition of the determinant of a matrix

$$\int_{-\infty}^{\infty} d^n x e^{-\mathbf{x}^t M \mathbf{x}} = \pi^{n/2} (\det M)^{-1/2} \Rightarrow \det M = \frac{\pi^n}{\left( \int_{-\infty}^{\infty} d^n x e^{-\mathbf{x}^t M \mathbf{x}} \right)^2},$$

which gives an analytic definition of the determinant. ■

## 4.8 Problems

4.1. Let  $\mathcal{U}_1$  and  $\mathcal{U}_2$  be subspaces of  $\mathcal{V}$ . Show that

(a)  $\dim(\mathcal{U}_1 + \mathcal{U}_2) = \dim \mathcal{U}_1 + \dim \mathcal{U}_2 - \dim(\mathcal{U}_1 \cap \mathcal{U}_2)$ . Hint: Extend a basis of  $\mathcal{U}_1 \cap \mathcal{U}_2$  to both  $\mathcal{U}_1$  and  $\mathcal{U}_2$ .

(b) If  $\mathcal{U}_1 + \mathcal{U}_2 = \mathcal{V}$  and  $\dim \mathcal{U}_1 + \dim \mathcal{U}_2 = \dim \mathcal{V}$ , then  $\mathcal{V} = \mathcal{U}_1 \oplus \mathcal{U}_2$ .

(b) If  $\dim \mathcal{U}_1 + \dim \mathcal{U}_2 > \dim \mathcal{V}$ , then  $\mathcal{U}_1 \cap \mathcal{U}_2 \neq \{0\}$ .

4.2. Let  $\mathbf{P}$  be the (hermitian) projection operator onto a subspace  $\mathcal{M}$ . Show that  $\mathbf{1} - \mathbf{P}$  projects onto  $\mathcal{M}^\perp$ . Hint: You need to show that  $\langle m | \mathbf{P} | a \rangle = \langle m | a \rangle$  for arbitrary  $|a\rangle$  and  $|m\rangle \in \mathcal{M}$ ; therefore, consider  $\langle m | \mathbf{P} | a \rangle^*$ , and use the hermiticity of  $\mathbf{P}$ .

4.3. Show that a subspace  $\mathcal{M}$  of an inner product space  $\mathcal{V}$  is invariant under the linear operator  $\mathbf{A}$  if and only if  $\mathcal{M}^\perp$  is invariant under  $\mathbf{A}^\dagger$ .

4.4. Show that the intersection of two invariant subspaces of an operator is also an invariant subspace.

4.5. Let  $\pi$  be a permutation of the integers  $\{1, 2, \dots, n\}$ . Find the spectrum of  $\mathbf{A}_\pi$ , if for  $|x\rangle = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{C}^n$ , we define

$$\mathbf{A}_\pi |x\rangle = (\alpha_{\pi(1)}, \dots, \alpha_{\pi(n)}).$$

4.6. Show that

(a) the coefficient of  $\lambda^N$  in the characteristic polynomial is  $(-1)^N$ , where  $N = \dim \mathcal{V}$ , and

(b) the constant in the characteristic polynomial of an operator is its determinant.

4.7. Operators  $\mathbf{A}$  and  $\mathbf{B}$  satisfy the commutation relation  $[\mathbf{A}, \mathbf{B}] = \mathbf{1}$ . Let  $|b\rangle$  be an eigenvector of  $\mathbf{B}$  with eigenvalue  $\lambda$ . Show that  $e^{-\tau\mathbf{A}} |b\rangle$  is also an eigenvector of  $\mathbf{B}$ , but with eigenvalue  $\lambda + \tau$ . This is why  $e^{-\tau\mathbf{A}}$  is called the **translation operator** for  $\mathbf{B}$ . Hint: First find  $[\mathbf{B}, e^{-\tau\mathbf{A}}]$ .

translation operator

4.8. Find the eigenvalues of an *involutive* operator, that is, an operator  $\mathbf{A}$  with the property  $\mathbf{A}^2 = \mathbf{1}$ .

4.9. Assume that  $\mathbf{A}$  and  $\mathbf{A}'$  are similar matrices. Show that they have the same eigenvalues.

4.10. In each of the following cases, determine the counterclockwise rotation of the  $xy$ -axes that brings the conic section into the standard form and determine the conic section.

$$(a) 11x^2 + 3y^2 + 6xy - 12 = 0$$

$$(b) 5x^2 - 3y^2 + 6xy + 6 = 0$$

$$(c) 2x^2 - y^2 - 4xy - 3 = 0$$

$$(d) 6x^2 + 3y^2 - 4xy - 7 = 0$$

$$(e) 2x^2 + 5y^2 - 4xy - 36 = 0$$

**4.11.** Show that if  $\mathbf{A}$  is invertible, then the eigenvectors of  $\mathbf{A}^{-1}$  are the same as those of  $\mathbf{A}$  and the eigenvalues of  $\mathbf{A}^{-1}$  are the reciprocals of those of  $\mathbf{A}$ .

**4.12.** Find all eigenvalues and eigenvectors of the following matrices:

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 1 \\ 0 & i \end{pmatrix} \quad \mathbf{B}_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \mathbf{C}_1 = \begin{pmatrix} 2 & -2 & -1 \\ -1 & 3 & 1 \\ 2 & -4 & -1 \end{pmatrix}$$

$$\mathbf{A}_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \mathbf{B}_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \mathbf{C}_2 = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

$$\mathbf{A}_3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{B}_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{C}_3 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

**4.13.** Show that a  $2 \times 2$  rotation matrix does not have a real eigenvalue (and, therefore, eigenvector) when the rotation angle is not an integer multiple of  $\pi$ . What is the physical interpretation of this?

**4.14.** Three equal point masses are located at  $(a, a, 0)$ ,  $(a, 0, a)$ , and  $(0, a, a)$ . Find the moment of inertia matrix as well as its eigenvalues and the corresponding eigenvectors.

**4.15.** Consider  $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{C}^n$  and define  $\mathbf{E}_{ij}$  as the operator that interchanges  $\alpha_i$  and  $\alpha_j$ . Find the eigenvalues of this operator.

**4.16.** Find the eigenvalues and eigenvectors of the operator  $-id/dx$  acting in the vector space of differentiable functions  $\mathcal{C}^1(-\infty, \infty)$ .

**4.17.** Show that a hermitian operator is positive if and only if its eigenvalues are positive.

**4.18.** What are the spectral decompositions of  $\mathbf{A}^\dagger$ ,  $\mathbf{A}^{-1}$ , and  $\mathbf{A}\mathbf{A}^\dagger$  for an invertible normal operator  $\mathbf{A}$ ?

**4.19.** Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}.$$

- Find the eigenvalues and the orthonormal eigenvectors of  $\mathbf{A}$ .
- Calculate the projection operators (matrices)  $\mathbf{P}_1$  and  $\mathbf{P}_2$  and verify that  $\sum_i \mathbf{P}_i = \mathbf{1}$  and  $\sum_i \lambda_i \mathbf{P}_i = \mathbf{A}$ .
- Find the matrices  $\sqrt{\mathbf{A}}$ ,  $\sin(\pi \mathbf{A}/6)$ , and  $\cos(\pi \mathbf{A}/6)$ .
- Is  $\mathbf{A}$  invertible? If so, find the eigenvalues and eigenvectors of  $\mathbf{A}^{-1}$ .

4.20. Consider the matrix

$$A = \begin{pmatrix} 4 & i & 1 \\ -i & 4 & -i \\ 1 & i & 4 \end{pmatrix}.$$

- (a) Find the eigenvalues of  $A$ . Hint: Try  $\lambda = 3$  in the characteristic polynomial of  $A$ .  
 (b) For each  $\lambda$ , find a basis for  $\mathcal{M}_\lambda$  the eigenspace associated with the eigenvalue  $\lambda$ .  
 (c) Use the Gram-Schmidt process to orthonormalize the above basis vectors.  
 (d) Calculate the projection operators (matrices)  $P_i$  for each subspace and verify that  $\sum_i P_i = 1$  and  $\sum_i \lambda_i P_i = A$ .  
 (e) Find the matrices  $\sqrt{A}$ ,  $\sin(\pi A/2)$ , and  $\cos(\pi A/2)$ .  
 (f) Is  $A$  invertible? If so, find the eigenvalues and eigenvectors of  $A^{-1}$ .

4.21. Show that if two hermitian matrices have the same set of eigenvalues, then they are unitarily related.

4.22. Prove that corresponding to every unitary operator  $U$  acting on a finite-dimensional vector space, there is a hermitian operator  $H$  such that  $U = \exp iH$ .

4.23. Find the polar decomposition of the following matrices:

$$A = \begin{pmatrix} 2i & 0 \\ \sqrt{7} & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 41 & -12i \\ 12i & 34 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -i \\ 1 & i & 0 \end{pmatrix}.$$

4.24. Show that an arbitrary matrix  $A$  can be “diagonalized” as  $D = UAV$ , where  $U$  is unitary and  $D$  is a real diagonal matrix with only nonnegative eigenvalues. Hint: Consider  $AA^\dagger$ .

4.25. Show that (a) if  $\lambda$  is an eigenvalue of an antisymmetric operator, then so is  $-\lambda$ , and (b) antisymmetric operators (matrices) of odd dimension cannot be invertible.

4.26. Find the unitary matrices that diagonalize the following hermitian matrices:

$$A_1 = \begin{pmatrix} 2 & -1+i \\ -1-i & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & -i \\ i & 0 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} 1 & -1 & -i \\ -1 & 0 & i \\ i & -i & -1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 2 & 0 & i \\ 0 & -1 & -i \\ -i & i & 0 \end{pmatrix}.$$

Warning! You may have to resort to numerical approximations for some of these.

4.27. For  $A = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ , where  $x \neq 0$ , show that it is impossible to find an invertible  $2 \times 2$  matrix  $R$  such that  $RAR^{-1}$  is diagonal. (This shows that not all operators are diagonalizable.)



## Additional Reading

1. Axler, S. *Linear Algebra Done Right*, Springer-Verlag, 1996. Concise but useful discussion of real and complex spectral theory.
2. DeVito, C. *Functional Analysis and Linear Operator Theory*, Addison-Wesley, 1990. Has a good discussion of spectral theory for finite and infinite dimensions.
3. Halmos, P. *Finite Dimensional Vector Spaces*, 2nd ed., Van Nostrand, 1958. Comprehensive treatment of real and complex spectral theory for operators on finite dimensional vector spaces.



Part II

---

# Infinite-Dimensional Vector Spaces



# Hilbert Spaces

The basic concepts of finite-dimensional vector spaces introduced in Chapter 1 can readily be generalized to infinite dimensions. The definition of a vector space and concepts of linear combination, linear independence, basis, subspace, span, and so forth all carry over to infinite dimensions. However, one thing is crucially different in the new situation, and this difference makes the study of infinite-dimensional vector spaces both richer and more nontrivial: In a finite-dimensional vector space we dealt with finite sums; in infinite dimensions we encounter infinite sums. Thus, we have to investigate the convergence of such sums.

## 5.1 The Question of Convergence

The intuitive notion of convergence acquired in calculus makes use of the idea of closeness. This, in turn, requires the notion of distance.<sup>1</sup> We considered such a notion in Chapter 1 in the context of a norm, and saw that the inner product had an associated norm. However, it is possible to introduce a norm on a vector space without an inner product.

One such norm, applicable to  $\mathbb{C}^n$  and  $\mathbb{R}^n$ , was

$$\|a\|_p \equiv \left( \sum_{i=1}^n |\alpha_i|^p \right)^{1/p},$$

where  $p$  is an integer. The “natural” norm, i.e., that induced on  $\mathbb{C}^n$  (or  $\mathbb{R}^n$ ) by the usual inner product, corresponds to  $p = 2$ . The distance between two points

---

<sup>1</sup>It is possible to introduce the idea of closeness abstractly, without resort to the notion of distance, as is done in topology. However, distance, as applied in vector spaces, is as abstract as we want to get.

Closeness is a relative concept!

depends on the particular norm used. For example, consider the “point” (or vector)  $|b\rangle = (0.1, 0.1, \dots, 0.1)$  in a 1000-dimensional space ( $n = 1000$ ). One can easily check that the distance of this vector from the origin varies considerably with  $p$ :  $\|b\|_1 = 100$ ,  $\|b\|_2 = 3.16$ ,  $\|b\|_{10} = 0.2$ . This variation may give the impression that there is no such thing as “closeness”, and it all depends on how one defines the norm. This is not true, because closeness is a relative concept: One always *compares* distances. A norm with large  $p$  shrinks *all* distances of a space, and a norm with small  $p$  stretches them. Thus, although it is impossible (and meaningless) to say that “ $|a\rangle$  is close to  $|b\rangle$ ” because of the dependence of distance on  $p$ , one can always say “ $|a\rangle$  is closer to  $|b\rangle$  than  $|c\rangle$  is to  $|d\rangle$ ,” regardless of the value of  $p$ .

Now that we have a way of telling whether vectors are close together or far apart, we can talk about limits and the convergence of sequences of vectors. Let us begin by recalling the definition of a Cauchy sequence

Cauchy sequence defined

**5.1.1. Definition.** An infinite sequence of vectors  $\{|a_i\rangle\}_{i=1}^\infty$  in a normed linear space  $\mathcal{V}$  is called a **Cauchy sequence** if  $\lim_{i,j \rightarrow \infty} \|a_i - a_j\| = 0$ .

A convergent sequence is necessarily Cauchy. This can be shown using the triangle inequality (see Problem 5.2). However, there may be Cauchy sequences in a given vector space that do not converge to any vector in that space (see the example below). Such a convergence requires additional properties of a vector space summarized in the following definition.

complete vector space defined

**5.1.2. Definition.** A **complete vector space**  $\mathcal{V}$  is a normed linear space for which every Cauchy sequence of vectors in  $\mathcal{V}$  has a limit vector in  $\mathcal{V}$ . In other words, if  $\{|a_i\rangle\}_{i=1}^\infty$  is a Cauchy sequence, then there exists a vector  $|a\rangle \in \mathcal{V}$  such that  $\lim_{i \rightarrow \infty} \|a_i - a\rangle = 0$ .

**5.1.3. Example.** 1.  $\mathbb{R}$  is complete with respect to the absolute-value norm  $\|\alpha\| = |\alpha|$ . In other words, every Cauchy sequence of real numbers has a limit in  $\mathbb{R}$ . This is proved in real analysis.

2.  $\mathbb{C}$  is complete with respect to the norm  $\|\alpha\| = |\alpha| = \sqrt{(\text{Re } \alpha)^2 + (\text{Im } \alpha)^2}$ . Using  $|\alpha| \leq |\text{Re } \alpha| + |\text{Im } \alpha|$ , one can show that the completeness of  $\mathbb{C}$  follows from that of  $\mathbb{R}$ . Details are left as an exercise for the reader.

3. The set of rational numbers  $\mathbb{Q}$  is *not* complete with respect to the absolute-value norm. In fact,  $\{(1 + 1/k)^k\}_{k=1}^\infty$  is a sequence of rational numbers that is Cauchy but does not converge to a rational number; it converges to  $e$ , the base of the natural logarithm, which is known to be an irrational number. ■

Let  $\{|a_i\rangle\}_{i=1}^\infty$  be a Cauchy sequence of vectors in a finite-dimensional vector space  $\mathcal{V}_N$ . Choose an orthonormal basis  $\{|e_k\rangle\}_{k=1}^N$  in  $\mathcal{V}_N$  such that<sup>2</sup>  $|a_i\rangle =$

<sup>2</sup>Recall that one can always define an inner product on a finite-dimensional vector space. So, the existence of orthonormal bases is guaranteed.

$\sum_{k=1}^N \alpha_k^{(i)} |e_k\rangle$  and  $|a_j\rangle = \sum_{k=1}^N \alpha_k^{(j)} |e_k\rangle$ . Then

$$\begin{aligned} \|a_i - a_j\|^2 &= \langle a_i - a_j | a_i - a_j \rangle = \left\| \sum_{k=1}^N (\alpha_k^{(i)} - \alpha_k^{(j)}) |e_k\rangle \right\|^2 \\ &= \sum_{k,l=1}^N (\alpha_k^{(i)} - \alpha_k^{(j)})^* (\alpha_l^{(i)} - \alpha_l^{(j)}) \langle e_k | e_l \rangle = \sum_{k=1}^N |\alpha_k^{(i)} - \alpha_k^{(j)}|^2. \end{aligned}$$

The LHS goes to zero, because the sequence is assumed Cauchy. Furthermore, all terms on the RHS are positive. Thus, they too must go to zero as  $i, j \rightarrow \infty$ . By the completeness of  $\mathbb{C}$ , there must exist  $\alpha_k \in \mathbb{C}$  such that  $\lim_{n \rightarrow \infty} \alpha_k^{(n)} = \alpha_k$  for  $k = 1, 2, \dots, N$ . Now consider  $|a\rangle \in \mathcal{V}_N$  given by  $|a\rangle = \sum_{k=1}^N \alpha_k |e_k\rangle$ . We claim that  $|a\rangle$  is the limit of the above sequence of vectors in  $\mathcal{V}_N$ . Indeed,

$$\lim_{i \rightarrow \infty} \|a_i - a\|^2 = \lim_{i \rightarrow \infty} \sum_{k=1}^N |\alpha_k^{(i)} - \alpha_k|^2 = \sum_{k=1}^N \lim_{i \rightarrow \infty} |\alpha_k^{(i)} - \alpha_k|^2 = 0$$

We have proved the following:

all finite-dimensional vector spaces are complete

**5.1.4. Proposition.** *Every Cauchy sequence in a finite-dimensional inner product space over  $\mathbb{C}$  (or  $\mathbb{R}$ ) is convergent. In other words, every finite-dimensional complex (or real) inner product space is complete with respect to the norm induced by its inner product.*

The next example shows how important the word “finite” is.

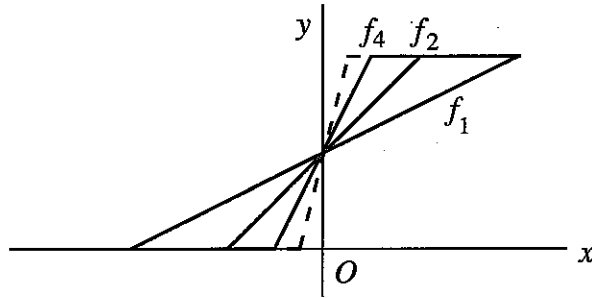
**5.1.5. Example.** Consider  $\{f_k\}_{k=1}^\infty$ , the infinite sequence of *continuous* functions defined in the interval  $[-1, +1]$  by

$$f_k(x) = \begin{cases} 1 & \text{if } 1/k \leq x \leq 1, \\ (kx + 1)/2 & \text{if } -1/k \leq x \leq 1/k, \\ 0 & \text{if } -1 \leq x \leq -1/k. \end{cases}$$

This sequence belongs to  $\mathcal{C}^0(-1, 1)$ , the inner product space of continuous functions with its usual inner product:  $\langle f | g \rangle = \int_{-1}^1 f^*(x)g(x) dx$ . It is straightforward to verify that  $\|f_k - f_j\|^2 = \int_{-1}^1 |f_k(x) - f_j(x)|^2 dx \xrightarrow[k, j \rightarrow \infty]{} 0$ . Therefore, the sequence is Cauchy. However, the limit of this sequence is (see Figure 5.1)

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1, \\ 0 & \text{if } -1 < x < 0, \end{cases}$$

which is discontinuous at  $x = 0$  and therefore does not belong to the space in which the original sequence lies. ■



**Figure 5.1** The limit of the sequence of the *continuous* functions  $f_k$  is a discontinuous function that is 1 for  $x > 0$  and 0 for  $x < 0$ .

We see that infinite-dimensional vector spaces are not generally complete. It is a nontrivial task to show whether or not a given infinite-dimensional vector space is complete.

Any vector space (finite- or infinite-dimensional) contains all finite linear combinations of the form  $\sum_{i=1}^n \alpha_i |a_i\rangle$  when it contains all the  $|a_i\rangle$ 's. This follows from the very definition of a vector space. However, the situation is different when  $n$  goes to infinity. For the vector space to contain the infinite sum, firstly, the meaning of such a sum has to be clarified, i.e., a norm and an associated convergence criterion needs to be put in place. Secondly, the vector space has to be complete with respect to that norm. A complete normed vector space is called a **Banach space**. We shall not deal with a general Banach space, but only with those spaces whose norms arise naturally from an inner product. This leads to the following definition:

Banach space

Hilbert space defined

**5.1.6. Definition.** A complete inner product space, commonly denoted by  $\mathcal{H}$ , is called a **Hilbert space**.

Thus, all finite-dimensional real or complex vector spaces are Hilbert spaces. However, when we speak of a Hilbert space, we shall usually assume that it is infinite-dimensional.

It is convenient to use orthonormal vectors in studying Hilbert spaces. So, let us consider an infinite sequence  $\{|e_i\rangle\}_{i=1}^{\infty}$  of orthonormal vectors all belonging to a Hilbert space  $\mathcal{H}$ . Next, take any vector  $|f\rangle \in \mathcal{H}$ , construct the complex numbers  $f_i = \langle e_i | f \rangle$ , and form the sequence of vectors<sup>3</sup>

$$|f_n\rangle = \sum_{i=1}^n f_i |e_i\rangle \quad \text{for } n = 1, 2, \dots \tag{5.1}$$

<sup>3</sup>We can consider  $|f_n\rangle$  as an "approximation" to  $|f\rangle$ , because both share the same components along the same set of orthonormal vectors. The sequence of orthonormal vectors acts very much as a basis. However, to be a basis, an extra condition must be met. We shall discuss this condition shortly.



For the pair of vectors  $|f\rangle$  and  $|f_n\rangle$ , the Schwarz inequality gives

$$|\langle f|f_n\rangle|^2 \leq \langle f|f\rangle \langle f_n|f_n\rangle = \langle f|f\rangle \left(\sum_{i=1}^n |f_i|^2\right), \tag{5.2}$$

where Equation (5.1) has been used to evaluate  $\langle f_n|f_n\rangle$ . On the other hand, taking the inner product of (5.1) with  $\langle f|$  yields

$$\langle f|f_n\rangle = \sum_{i=1}^n f_i \langle f|e_i\rangle = \sum_{i=1}^n f_i f_i^* = \sum_{i=1}^n |f_i|^2.$$

Parseval inequality Substitution of this in Equation (5.2) yields the **Parseval inequality**:

$$\sum_{i=1}^n |f_i|^2 \leq \langle f|f\rangle. \tag{5.3}$$

This conclusion is true for arbitrarily large  $n$  and can be stated as follows:

**5.1.7. Proposition.** Let  $\{|e_i\rangle\}_{i=1}^\infty$  be an infinite set of orthonormal vectors in a Hilbert space,  $\mathcal{H}$ . Let  $|f\rangle \in \mathcal{H}$  and define complex numbers  $f_i = \langle e_i|f\rangle$ . Then the **Bessel inequality** holds:  $\sum_{i=1}^\infty |f_i|^2 \leq \langle f|f\rangle$ .

The Bessel inequality shows that the vector

$$\sum_{i=1}^\infty f_i |e_i\rangle \equiv \lim_{n \rightarrow \infty} \sum_{i=1}^n f_i |e_i\rangle$$

converges; that is, it has a finite norm. However, the inequality does not say whether the vector converges to  $|f\rangle$ . To make such a statement we need completeness:

**5.1.8. Definition.** A sequence of orthonormal vectors  $\{|e_i\rangle\}_{i=1}^\infty$  in a Hilbert space  $\mathcal{H}$  is called **complete** if the only vector in  $\mathcal{H}$  that is orthogonal to all the  $|e_i\rangle$  is the zero vector.

This completeness property is the extra condition alluded to (in the footnote) above, and is what is required to make a basis.

**5.1.9. Proposition.** Let  $\{|e_i\rangle\}_{i=1}^\infty$  be an orthonormal sequence in  $\mathcal{H}$ . Then the following statements are equivalent:

1.  $\{|e_i\rangle\}_{i=1}^\infty$  is complete.
2.  $|f\rangle = \sum_{i=1}^\infty |e_i\rangle \langle e_i|f\rangle \quad \forall |f\rangle \in \mathcal{H}$ .
3.  $\sum_{i=1}^\infty |e_i\rangle \langle e_i| = \mathbf{1}$ .
4.  $\langle f|g\rangle = \sum_{i=1}^\infty \langle f|e_i\rangle \langle e_i|g\rangle \quad \forall |f\rangle, |g\rangle \in \mathcal{H}$ .

$$5. \|f\|^2 = \sum_{i=1}^{\infty} |\langle e_i | f \rangle|^2 \quad \forall |f\rangle \in \mathcal{H}.$$

*Proof.* We shall prove the implications  $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 5 \Rightarrow 1$ .

$1 \Rightarrow 2$ : It is sufficient to show that the vector  $|\psi\rangle \equiv |f\rangle - \sum_{i=1}^{\infty} |e_i\rangle \langle e_i | f \rangle$  is orthogonal to all the  $|e_j\rangle$ :

$$\langle e_j | \psi \rangle = \langle e_j | f \rangle - \sum_{i=1}^{\infty} \overbrace{\langle e_j | e_i \rangle}^{\delta_{ij}} \langle e_i | f \rangle = 0.$$

$2 \Rightarrow 3$ : Since  $|f\rangle = \mathbf{1} |f\rangle = \sum_{i=1}^{\infty} \langle e_i | \langle e_i | |f\rangle$  is true for all  $|f\rangle \in \mathcal{H}$ , we must have  $\mathbf{1} = \sum_{i=1}^{\infty} |e_i\rangle \langle e_i|$ .

$3 \Rightarrow 4$ :  $\langle f | g \rangle = \langle f | \mathbf{1} | g \rangle = \langle f | (\sum_{i=1}^{\infty} |e_i\rangle \langle e_i|) | g \rangle = \sum_{i=1}^{\infty} \langle f | e_i \rangle \langle e_i | g \rangle$ .

$4 \Rightarrow 5$ : Let  $|g\rangle = |f\rangle$  in statement 4 and recall that  $\langle f | e_i \rangle = \langle e_i | f \rangle^*$ .

$5 \Rightarrow 1$ : Let  $|f\rangle$  be orthogonal to all the  $|e_i\rangle$ . Then all the terms in the sum are zero implying that  $\|f\|^2 = 0$ , which in turn gives  $|f\rangle = 0$ , because only the zero vector has a zero norm.  $\square$

Parseval equality;  
generalized Fourier  
coefficients

The equality

$$\|f\|^2 = \langle f | f \rangle = \sum_{i=1}^{\infty} |\langle e_i | f \rangle|^2 = \sum_{i=1}^{\infty} |f_i|^2, \quad f_i = \langle e_i | f \rangle, \quad (5.4)$$

completeness  
relation

is called the **Parseval equality**, and the complex numbers  $f_i$  are called **generalized Fourier coefficients**. The relation

$$\mathbf{1} = \sum_{i=1}^{\infty} |e_i\rangle \langle e_i| \quad (5.5)$$

is called the **completeness relation**.

basis for Hilbert  
spaces

**5.1.10. Definition.** A complete orthonormal sequence  $\{|e_i\rangle\}_{i=1}^{\infty}$  in a Hilbert space  $\mathcal{H}$  is called a **basis** of  $\mathcal{H}$ .

## 5.2 The Space of Square-Integrable Functions

Chapter 1 showed that the collection of all continuous functions defined on an interval  $[a, b]$  forms a linear vector space. Example 5.1.5 showed that this space is not complete. Can we enlarge this space to make it complete? Since we are interested in an inner product as well, and since a natural inner product for functions is defined in terms of integrals, we want to make sure that our functions are integrable. However, integrability does not require continuity, it only requires *piecewise* continuity. In this section we shall discuss conditions under which the

space of functions becomes complete. An important class of functions has already been mentioned in Chapter 1. These functions satisfy the inner product given by

$$\langle g | f \rangle = \int_a^b g^*(x) f(x) w(x) dx.$$

square-integrable  
functions

If  $g(x) = f(x)$ , we obtain

$$\langle f | f \rangle = \int_a^b |f(x)|^2 w(x) dx. \quad (5.6)$$

Functions for which such an integral is defined are said to be **square-integrable**.

**David Hilbert** (1862–1943), the greatest mathematician of this century, received his Ph.D. from the University of Königsberg and was a member of the staff there from 1886 to 1895. In 1895 he was appointed to the chair of mathematics at the University of Göttingen, where he continued to teach for the rest of his life.

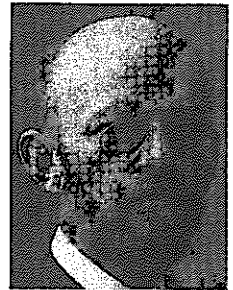
Hilbert is one of that rare breed of late 19th-century mathematicians whose spectrum of expertise covered a wide range, with formal set theory at one end and mathematical physics at the other. He did superb work in geometry, algebraic geometry, algebraic number theory, integral equations, and operator theory. The seminal two-volume book *Methoden der mathematische Physik* by R. Courant, still one of the best books on the subject, was greatly influenced by Hilbert.

Hilbert's work in geometry had the greatest influence in that area since Euclid. A systematic study of the axioms of Euclidean geometry led Hilbert to propose 21 such axioms, and he analyzed their significance. He published *Grundlagen der Geometrie* in 1899, putting geometry on a formal axiomatic foundation. His famous 23 Paris problems challenged (and still today challenge) mathematicians to solve fundamental questions.

It was late in his career that Hilbert turned to the subject for which he is most famous among physicists. A lecture by Erik Holmgren in 1901 on Fredholm's work on integral equations, which had already been published in Sweden, aroused Hilbert's interest in the subject. David Hilbert, having established himself as the leading mathematician of his time by his work on algebraic numbers, algebraic invariants, and the foundations of geometry, now turned his attention to **integral equations**. He says that an investigation of the subject showed him that it was important for the theory of definite integrals, for the development of arbitrary functions in series (of special functions or trigonometric functions), for the theory of linear differential equations, for potential theory, and for the calculus of variations. He wrote a series of six papers from 1904 to 1910 and reproduced them in his book *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* (1912). During the latter part of this work he applied integral equations to problems of mathematical physics.

It is said that Hilbert discovered the correct field equation for general relativity in 1915 (one year before Einstein) using the variational principle, but never claimed priority.

Hilbert claimed that he worked best out-of-doors. He accordingly attached an 18-foot blackboard to his neighbor's wall and built a covered walkway there so that he could work outside in any weather. He would intermittently interrupt his pacing and his blackboard



computations with a few turns around the rest of the yard on his bicycle, or he would pull some weeds, or do some garden trimming. Once, when a visitor called, the maid sent him to the backyard and advised that if the master wasn't readily visible at the blackboard to look for him up in one of the trees.

Highly gifted and highly versatile, David Hilbert radiated over mathematics a catching optimism and a stimulating vitality that can only be called "the spirit of Hilbert." Engraved on a stone marker set over Hilbert's grave in Göttingen are the master's own optimistic words: "Wir müssen wissen. Wir werden wissen." ("We must know. We shall know.")

The space of square-integrable functions over the interval  $[a, b]$  is denoted by  $\mathcal{L}_w^2(a, b)$ . In this notation  $\mathcal{L}$  stands for *Lebesgue*, who generalized the notion of the ordinary Riemann integral to cases for which the integrand could be highly discontinuous; 2 stands for the power of  $f(x)$  in the integral;  $a$  and  $b$  denote the limits of integration; and  $w$  refers to the weight function (a strictly positive real-valued function). When  $w(x) = 1$ , we use the notation  $\mathcal{L}^2(a, b)$ . The significance of  $\mathcal{L}_w^2(a, b)$  lies in the following theorem (for a proof, see [Reed 80, Chapter III]):

$\mathcal{L}_w^2(a, b)$  is complete

**5.2.1. Theorem.** (Riesz-Fischer theorem) *The space  $\mathcal{L}_w^2(a, b)$  is complete.*

A complete infinite-dimensional inner product space was earlier defined to be a Hilbert space. The following theorem shows that the number of Hilbert spaces is severely restricted. (For a proof, see [Frie 82, p. 216].)

all Hilbert spaces are alike

**5.2.2. Theorem.** *All infinite-dimensional complete inner product spaces are isomorphic to  $\mathcal{L}_w^2(a, b)$ .*

$\mathcal{L}_w^2(a, b)$  is defined in terms of functions that satisfy Equation (5.6). Yet an inner product involves integrals of the form  $\int_a^b g^*(x)f(x)w(x)dx$ : Are such integrals well-defined and finite? Using the Schwarz inequality, which holds for any inner product space, finite or infinite, one can show that the integral is defined. The isomorphism of Theorem 5.2.2 makes the Hilbert space more tangible, because it identifies the space with a space of functions, objects that are more familiar than abstract vectors. Nonetheless, a faceless function is very little improvement over an abstract vector. What is desirable is a set of concrete functions with which we can calculate. The following theorem provides such functions (for a proof, see [Simm 83, pp. 154–161]).

**5.2.3. Theorem.** (Stone-Weierstrass approximation theorem) *The sequence of functions (monomials)  $\{x^k\}$ , where  $k = 0, 1, 2, \dots$ , forms a basis of  $\mathcal{L}_w^2(a, b)$ .*

Thus, any function  $f$  can be written as  $f(x) = \sum_{k=0}^{\infty} \alpha_k x^k$ . Note that the  $\{x^k\}$  are not orthonormal but are linearly independent. If we wish to obtain an orthonormal—or simply orthogonal—linear combination of these vectors, we can use the Gram-Schmidt process. The result will be certain polynomials, denoted by  $C_n(x)$ , that are orthogonal to one another and span  $\mathcal{L}_w^2(a, b)$ .

Such orthogonal polynomials satisfy very useful **recurrence relations**, which we now derive. In the following discussion  $p_{\leq k}(x)$  denotes a generic polynomial of degree less than or equal to  $k$ . For example,  $3x^5 - 4x^2 + 5$ ,  $2x + 1$ ,  $-2.4x^4 + 3x^3 - x^2 + 6$ , and  $2$  are all denoted by  $p_{\leq 5}(x)$  or  $p_{\leq 8}(x)$  or  $p_{\leq 59}(x)$  because they all have degrees less than or equal to  $5$ ,  $8$ , and  $59$ . Since a polynomial of degree less than  $n$  can be written as a linear combination of  $C_k(x)$  with  $k < n$ , we have the obvious property

$$\int_a^b C_n(x) p_{\leq n-1}(x) w(x) dx = 0. \quad (5.7)$$

Let  $k_m$  and  $k'_m$  denote, respectively, the coefficients of  $x^m$  and  $x^{m-1}$  in  $C_m(x)$ , and let

$$h_m = \int_a^b [C_m(x)]^2 w(x) dx. \quad (5.8)$$

The polynomial  $C_{n+1}(x) - (k_{n+1}/k_n)x C_n(x)$  has degree less than or equal to  $n$ , and therefore can be expanded as a linear combination of the  $C_j(x)$ :

$$C_{n+1}(x) - \frac{k_{n+1}}{k_n} x C_n(x) = \sum_{j=0}^n a_j C_j(x). \quad (5.9)$$

Take the inner product of both sides of this equation with  $C_m(x)$ :

$$\begin{aligned} \int_a^b C_{n+1}(x) C_m(x) w(x) dx - \frac{k_{n+1}}{k_n} \int_a^b x C_n(x) C_m(x) w(x) dx \\ = \sum_{j=0}^n a_j \int_a^b C_j(x) C_m(x) w(x) dx. \end{aligned}$$

The first integral on the LHS vanishes as long as  $m \leq n$ ; the second integral vanishes if  $m \leq n - 2$  [if  $m \leq n - 2$ , then  $x C_m(x)$  is a polynomial of degree  $n - 1$ ]. Thus, we have

$$\sum_{j=0}^n a_j \int_a^b C_j(x) C_m(x) w(x) dx = 0 \quad \text{for } m \leq n - 2.$$

The integral in the sum is zero unless  $j = m$ , by orthogonality. Therefore, the sum reduces to

$$a_m \int_a^b [C_m(x)]^2 w(x) dx = 0 \quad \text{for } m \leq n - 2.$$

Since the integral is nonzero, we conclude that  $a_m = 0$  for  $m = 0, 1, 2, \dots, n - 2$ , and Equation (5.9) reduces to

$$C_{n+1}(x) - \frac{k_{n+1}}{k_n} x C_n(x) = a_{n-1} C_{n-1}(x) + a_n C_n(x). \quad (5.10)$$

It can be shown that if we define

$$\alpha_n = \frac{k_{n+1}}{k_n}, \quad \beta_n = \alpha_n \left( \frac{k'_{n+1}}{k_{n+1}} - \frac{k'_n}{k_n} \right), \quad \gamma_n = -\frac{h_n}{h_{n-1}} \frac{\alpha_n}{\alpha_{n-1}}, \quad (5.11)$$

a recurrence relation for orthogonal polynomials

then Equation (5.10) can be expressed as

$$C_{n+1}(x) = (\alpha_n x + \beta_n) C_n(x) + \gamma_n C_{n-1}(x), \quad (5.12)$$

or

$$x C_n(x) = \frac{1}{\alpha_n} C_{n+1}(x) - \frac{\beta_n}{\alpha_n} C_n(x) - \frac{\gamma_n}{\alpha_n} C_{n-1}(x). \quad (5.13)$$

Other recurrence relations, involving higher powers of  $x$ , can be obtained from the one above. For example, a recurrence relation involving  $x^2$  can be obtained by multiplying both sides of Equation (5.13) by  $x$  and expanding each term of the RHS using that same equation. The result will be

$$\begin{aligned} x^2 C_n(x) &= \frac{1}{\alpha_n \alpha_{n+1}} C_{n+2}(x) - \left( \frac{\beta_{n+1}}{\alpha_n \alpha_{n+1}} + \frac{\beta_n}{\alpha_n^2} \right) C_{n+1}(x) \\ &\quad - \left( \frac{\gamma_{n+1}}{\alpha_n \alpha_{n+1}} - \frac{\beta_n^2}{\alpha_n^2} + \frac{\gamma_n}{\alpha_n \alpha_{n-1}} \right) C_n(x) \\ &\quad + \left( \frac{\beta_n \gamma_n}{\alpha_n^2} + \frac{\beta_{n-1} \gamma_n}{\alpha_n \alpha_{n-1}} \right) C_{n-1}(x) + \frac{\gamma_{n-1} \gamma_n}{\alpha_n \alpha_{n-1}} C_{n-2}(x). \end{aligned} \quad (5.14)$$

**5.2.4. Example.** As an application of the recurrence relations above, let us evaluate

$$I_1 \equiv \int_a^b x C_m(x) C_n(x) w(x) dx.$$

Substituting (5.13) in the integral gives

$$\begin{aligned} I_1 &= \frac{1}{\alpha_n} \int_a^b C_m(x) C_{n+1}(x) w(x) dx - \frac{\beta_n}{\alpha_n} \int_a^b C_m(x) C_n(x) w(x) dx \\ &\quad - \frac{\gamma_n}{\alpha_n} \int_a^b C_m(x) C_{n-1}(x) w(x) dx. \end{aligned}$$

We now use the orthogonality relations among the  $C_k(x)$  to obtain

$$\begin{aligned} I_1 &= \frac{1}{\alpha_n} \delta_{m,n+1} \overbrace{\int_a^b C_m^2(x) w(x) dx}^{=h_m} - \frac{\beta_n}{\alpha_n} \delta_{mn} \int_a^b C_m^2(x) w(x) dx \\ &\quad - \frac{\gamma_n}{\alpha_n} \delta_{m,n-1} \int_a^b C_m^2(x) w(x) dx \\ &= \left( \frac{1}{\alpha_{m-1}} \delta_{m,n+1} - \frac{\beta_m}{\alpha_m} \delta_{mn} - \frac{\gamma_{m+1}}{\alpha_{m+1}} \delta_{m,n-1} \right) h_m, \end{aligned}$$

or

$$I_1 = \begin{cases} h_m/\alpha_{m-1} & \text{if } m = n + 1, \\ -\beta_m h_m/\alpha_m & \text{if } m = n, \\ -\gamma_{m+1} h_m/\alpha_{m+1} & \text{if } m = n - 1, \\ 0 & \text{otherwise.} \end{cases}$$

**5.2.5. Example.** Let us find the orthogonal polynomials forming a basis of  $\mathcal{L}^2(-1, +1)$ , which we denote by  $P_k(x)$ , where  $k$  is the degree of the polynomial. Let  $P_0(x) = 1$ . To find  $P_1(x)$ , write  $P_1(x) = ax + b$ , and determine  $a$  and  $b$  in such a way that  $P_1(x)$  is orthogonal to  $P_0(x)$ :

$$0 = \int_{-1}^1 P_1(x)P_0(x) dx = \int_{-1}^1 (ax + b) dx = \frac{1}{2}ax^2 \Big|_{-1}^1 + 2b = 2b.$$

So one of the coefficients,  $b$ , is zero. To find the other one, we need some standardization procedure. We “standardize”  $P_k(x)$  by requiring that  $P_k(1) = 1 \forall k$ . For  $k = 1$  this yields  $a \times 1 = 1$ , or  $a = 1$ , so that  $P_1(x) = x$ .

We can calculate  $P_2(x)$  similarly: Write  $P_2(x) = ax^2 + bx + c$ , impose the condition that it be orthogonal to both  $P_1(x)$  and  $P_0(x)$ , and enforce the standardization procedure. All this will yield

$$0 = \int_{-1}^1 P_2(x)P_0(x) dx = \frac{2}{3}a + 2c, \quad 0 = \int_{-1}^1 P_2(x)P_1(x) dx = \frac{2}{3}b,$$

and  $P_2(1) = a + b + c = 1$ . These three equations have the unique solution  $a = 3/2$ ,  $b = 0$ ,  $c = -1/2$ . Thus,  $P_2(x) = \frac{1}{2}(3x^2 - 1)$ . These are the first three Legendre polynomials, which are part of a larger group of polynomials to be discussed in Chapter 7. ■

## 5.2.1 Orthogonal Polynomials and Least Squares

The method of least squares is no doubt familiar to the reader. In the simplest procedure, one tries to find a linear function that most closely fits a set of data. By definition, “most closely” means that the sum of the squares of the differences between the data points and the corresponding values of the linear function is minimum. More generally, one seeks the best polynomial fit to the data.

We shall consider a related topic, namely least-square fitting of a given function with polynomials. Suppose  $f(x)$  is a function defined on  $(a, b)$ . We want to find a polynomial that most closely approximates  $f$ . Write such a polynomial as  $p(x) = \sum_{k=0}^n a_k x^k$ , where the  $a_k$ 's are to be determined such that

$$S(a_0, a_1, \dots, a_n) \equiv \int_a^b [f(x) - a_0 - a_1 x - \dots - a_n x^n]^2 dx$$

is a minimum. Differentiating  $S$  with respect to the  $a_k$ 's and setting the result equal to zero gives

$$0 = \frac{\partial S}{\partial a_j} = \int_a^b 2(-x^j) \left[ f(x) - \sum_{k=0}^n a_k x^k \right] dx,$$

- (a)  $\|f \pm g\| = \|f\| + \|g\|$ .  
(b)  $\|f + g\|^2 + \|f - g\|^2 = 2(\|f\| + \|g\|)^2$ .  
(c) Using parts (a), (b), and Theorem 1.2.8, show that  $\mathcal{L}^1(\mathbb{R})$  is not an inner product space. This shows that not all norms arise from an inner product.

**5.6.** Use Equation (5.10) to derive Equation (5.12). Hint: To find  $a_n$ , equate the coefficients of  $x^n$  on both sides of Equation (5.10). To find  $a_{n-1}$ , multiply both sides of Equation (5.10) by  $C_{n-1}w(x)$  and integrate, using the definitions of  $k_n$ ,  $k'_n$ , and  $h_n$ .

**5.7.** Evaluate the integral  $\int_a^b x^2 C_m(x) C_n(x) w(x) dx$ .

## Additional Reading

1. Boccara, N. *Functional Analysis*, Academic Press, 1990. An application oriented book with many abstract topics related to Hilbert spaces (e.g., Lebesgue measure) explained for a physics audience.
2. DeVito, C. *Functional Analysis and Linear Operator Theory*, Addison-Wesley, 1990.
3. Reed, M., and Simon, B. *Functional Analysis*, Academic Press, 1980. Coauthored by a mathematical physicist (B.S.), this first volume of a four-volume encyclopedic treatise on functional analysis and Hilbert spaces has many examples and problems to help the reader comprehend the rather abstract presentation.
4. Zeidler, E. *Applied Functional Analysis*, Springer-Verlag, 1995. Another application-oriented book on Hilbert spaces suitable for a physics audience.



# Generalized Functions

Once we allow the number of dimensions to be infinite, we open the door for numerous possibilities that are not present in the finite case. One such possibility arises because of the variety of infinities. We have encountered two types of infinity in Chapter 0, the countable infinity and the uncountable infinity. The paradigm of the former is the “number” of integers, and that of the latter is the “number” of real numbers. The nature of dimensionality of the vector space is reflected in the components of a general vector, which has a finite number of components in a finite-dimensional vector space, a countably infinite number of components in an infinite-dimensional vector space with a *countable basis*, and an uncountably infinite number of components in an infinite-dimensional vector space with no countable basis.

## 6.1 Continuous Index

To gain an understanding of the nature of, and differences between, the three types of vector spaces mentioned above, it is convenient to think of components as functions of a “counting set.” Thus, the components  $f_i$  of a vector  $|f\rangle$  in an  $N$ -dimensional vector space can be thought of as values of a function  $f$  defined on the finite set  $\{1, 2, \dots, N\}$ , and to emphasize such functional dependence, we write  $f(i)$  instead of  $f_i$ . Similarly, the components  $f_i$  of a vector  $|f\rangle$  in a Hilbert space with the countable basis  $B = \{|e_i\rangle\}_{i=1}^{\infty}$  can be thought of as values of a function  $f : \mathbb{N} \rightarrow \mathbb{C}$ , where  $\mathbb{N}$  is the (infinite) set of natural numbers. The next step is to allow the counting set to be uncountable, i.e., a continuum such as the real numbers or an interval thereof. This leads to a “component” of the form  $f(x)$  corresponding to a function  $f : \mathbb{R} \rightarrow \mathbb{C}$ . What about the vectors themselves? What sort of a basis gives rise to such components?

Because of the isomorphism of Theorem 5.2.2, we shall concentrate on  $\mathcal{L}_w^2(a, b)$ . In keeping with our earlier notation, let  $\{|e_x\rangle\}_{x \in \mathbb{R}}$  denote the elements of an orthonormal set and interpret  $f(x)$  as  $\langle e_x | f \rangle$ . The inner product of  $\mathcal{L}_w^2(a, b)$  can now be written as

$$\begin{aligned} \langle g | f \rangle &= \int_a^b g^*(x) f(x) w(x) dx = \int_a^b \langle g | e_x \rangle \langle e_x | f \rangle w(x) dx \\ &= \langle g | \left( \int_a^b |e_x\rangle w(x) \langle e_x| dx \right) | f \rangle. \end{aligned}$$

The last line suggests writing

$$\int_a^b |e_x\rangle w(x) \langle e_x| dx = \mathbf{1}.$$

In the physics literature the “ $e$ ” is ignored, and one writes  $|x\rangle$  for  $|e_x\rangle$ . Hence, we obtain the completeness relation for a continuous index:

completeness  
relation for a  
continuous index

$$\int_a^b |x\rangle w(x) \langle x| dx = \mathbf{1}, \quad \text{or} \quad \int_a^b |x\rangle \langle x| dx = \mathbf{1}, \quad (6.1)$$

where in the second integral,  $w(x)$  is set equal to unity. We also have

$$|f\rangle = \left( \int_a^b |x\rangle w(x) \langle x| dx \right) |f\rangle = \int_a^b f(x) w(x) |x\rangle dx, \quad (6.2)$$

which shows how to expand a vector  $|f\rangle$  in terms of the  $|x\rangle$ 's.

Take the inner product of (6.2) with  $\langle x'|$ :

$$\langle x' | f \rangle = f(x') = \int_a^b f(x) w(x) \langle x' | x \rangle dx$$

where  $x'$  is assumed to lie in the interval  $(a, b)$ , otherwise  $f(x') = 0$  by definition. This equation, which holds for arbitrary  $f$ , tells us immediately that  $w(x) \langle x' | x \rangle$  is no ordinary function of  $x$  and  $x'$ . For instance, suppose  $f(x') = 0$ . Then, the result of integration is always zero, regardless of the behavior of  $f$  at other points. Clearly, there is an infinitude of functions that vanish at  $x'$ , yet all of them give the same integral! Pursuing this line of argument more quantitatively, one can show that  $w(x) \langle x' | x \rangle = 0$  if  $x \neq x'$ ,  $w(x) \langle x | x \rangle = \infty$ ,  $w(x) \langle x' | x \rangle$  is an even function of  $x - x'$ , and  $\int_a^b w(x) \langle x' | x \rangle dx = 1$ . The proof is left as a problem. The reader

Dirac delta function

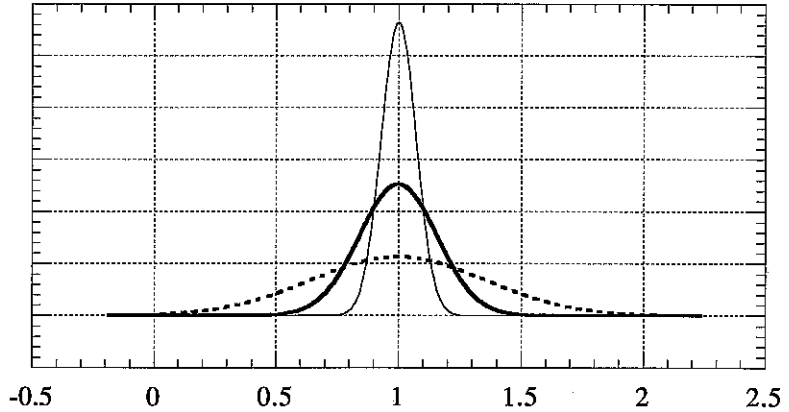
may recognize this as the **Dirac delta function**

$$\delta(x - x') = w(x) \langle x' | x \rangle, \quad (6.3)$$

which, for a function  $f$  defined on the interval  $(a, b)$ , has the following property:<sup>1</sup>

$$\int_a^b f(x) \delta(x - x') dx = \begin{cases} f(x') & \text{if } x \in (a, b), \\ 0 & \text{if } x \notin (a, b). \end{cases} \quad (6.4)$$

<sup>1</sup>For an elementary discussion of the Dirac delta function with many examples of its application, see [Hass 99].



**Figure 6.1** The Gaussian bell-shaped curve approaches the Dirac delta function as the width of the curve approaches zero. The value of  $\epsilon$  is 1 for the dashed curve, 0.25 for the heavy curve and 0.05 for the light curve.

Written in the form  $\langle x' | x \rangle = \delta(x - x')/w(x)$ , Equation (6.3) is the generalization of the orthonormality relation of vectors to the case of a continuous index.

The Dirac delta function is anything but a “function.” Nevertheless, there is a well-developed branch of mathematics, called generalized function theory or functional analysis, studying it and many other functions like it in a highly rigorous fashion. We shall only briefly explore this territory of mathematics in the next section. At this point we simply mention the fact that the Dirac delta function can be represented as the limit of certain sequences of ordinary functions. The following three examples illustrate some of these representations.

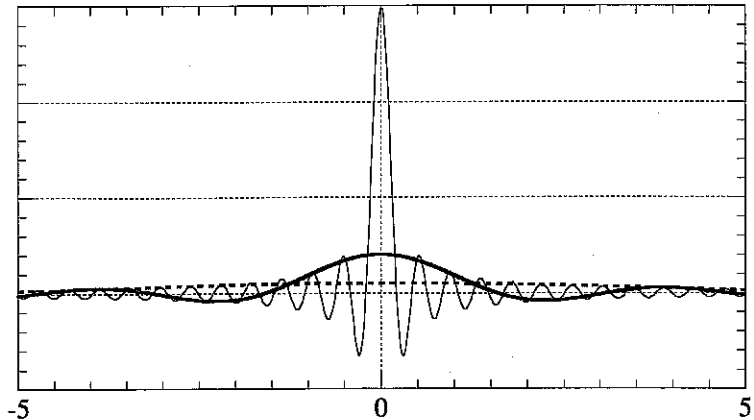
**6.1.1. Example.** Consider a Gaussian curve whose width approaches zero at the same time that its height approaches infinity in such a way that its area remains constant. In the infinite limit, we obtain the Dirac delta function. In fact, we have  $\delta(x - x') = \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{\epsilon\pi}} e^{-(x-x')^2/\epsilon}$ . In the limit of  $\epsilon \rightarrow 0$ , the height of this Gaussian goes to infinity while its width goes to zero (see Figure 6.1). Furthermore, for any nonzero value of  $\epsilon$ , we can easily verify that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{\epsilon\pi}} e^{-(x-x')^2/\epsilon} dx = 1.$$

This relation is independent of  $\epsilon$  and therefore still holds in the limit  $\epsilon \rightarrow 0$ . The limit of the Gaussian behaves like the Dirac delta function. ■

**6.1.2. Example.** Consider the function  $D_T(x - x')$  defined as

$$D_T(x - x') \equiv \frac{1}{2\pi} \int_{-T}^T e^{i(x-x')t} dt.$$



**Figure 6.2** The function  $\sin T x/x$  also approaches the Dirac delta function as the width of the curve approaches zero. The value of  $T$  is 0.5 for the dashed curve, 2 for the heavy curve, and 15 for the light curve.

The integral is easily evaluated, with the result

$$D_T(x - x') = \frac{1}{2\pi} \left. \frac{e^{i(x-x')t}}{i(x-x')} \right|_{-T}^T = \frac{1}{\pi} \frac{\sin T(x - x')}{x - x'}$$

The graph of  $D_T(x - 0)$  as a function of  $x$  for various values of  $T$  is shown in Figure 6.2. Note that the width of the curve decreases as  $T$  increases. The area under the curve can be calculated:

$$\int_{-\infty}^{\infty} D_T(x - x') dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin T(x - x')}{x - x'} dx = \frac{1}{\pi} \underbrace{\int_{-\infty}^{\infty} \frac{\sin y}{y} dy}_{=\pi} = 1.$$

Figure 6.2 shows that  $D_T(x - x')$  becomes more and more like the Dirac delta function as  $T$  gets larger and larger. In fact, we have

$$\delta(x - x') = \lim_{T \rightarrow \infty} \frac{1}{\pi} \frac{\sin T(x - x')}{x - x'} \tag{6.5}$$

To see this, we note that for any finite  $T$  we can write

$$D_T(x - x') = \frac{T}{\pi} \frac{\sin T(x - x')}{T(x - x')}$$

Furthermore, for values of  $x$  that are very close to  $x'$ ,

$$T(x - x') \rightarrow 0 \quad \text{and} \quad \frac{\sin T(x - x')}{T(x - x')} \rightarrow 1.$$

Thus, for such values of  $x$  and  $x'$ , we have  $D_T(x - x') \approx (T/\pi)$ , which is large when  $T$  is large. This is as expected of a delta function:  $\delta(0) = \infty$ . On the other hand, the width

of  $D_T(x - x')$  around  $x'$  is given, roughly, by the distance between the points at which  $D_T(x - x')$  drops to zero:  $T(x - x') = \pm\pi$ , or  $x - x' = \pm\pi/T$ . This width is roughly  $\Delta x = 2\pi/T$ , which goes to zero as  $T$  grows. Again, this is as expected of the delta function. ■

The preceding example suggests another representation of the Dirac delta function:

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(x-x')t} dt. \tag{6.6}$$

**6.1.3. Example.** A third representation of the Dirac delta function involves the **step function**  $\theta(x - x')$ , which is defined as

$$\theta(x - x') \equiv \begin{cases} 0 & \text{if } x < x', \\ 1 & \text{if } x > x' \end{cases}$$

and is discontinuous at  $x = x'$ . We can approximate this step function by many continuous functions, such as  $T_\epsilon(x - x')$  defined by

$$T_\epsilon(x - x') \equiv \begin{cases} 0 & \text{if } x \leq x' - \epsilon, \\ \frac{1}{2\epsilon}(x - x' + \epsilon) & \text{if } x' - \epsilon \leq x \leq x' + \epsilon, \\ 1 & \text{if } x \geq x' + \epsilon, \end{cases}$$

where  $\epsilon$  is a small positive number as shown in Figure 6.3. It is clear that

$$\theta(x - x') = \lim_{\epsilon \rightarrow 0} T_\epsilon(x - x').$$

Now let us consider the derivative of  $T_\epsilon(x - x')$  with respect to  $x$ :

$$\frac{dT_\epsilon}{dx}(x - x') = \begin{cases} 0 & \text{if } x < x' - \epsilon, \\ \frac{1}{2\epsilon} & \text{if } x' - \epsilon < x < x' + \epsilon, \\ 0 & \text{if } x > x' + \epsilon. \end{cases}$$

We note that the derivative is not defined at  $x = x' - \epsilon$  and  $x = x' + \epsilon$ , and that  $dT_\epsilon/dx$  is zero everywhere except when  $x$  lies in the interval  $(x' - \epsilon, x' + \epsilon)$ , where it is equal to  $1/(2\epsilon)$  and goes to infinity as  $\epsilon \rightarrow 0$ . Here again we see signs of the delta function. In fact, we also note that

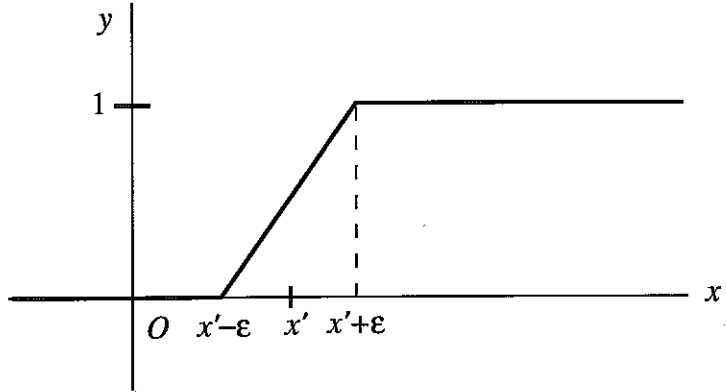
$$\int_{-\infty}^{\infty} \left(\frac{dT_\epsilon}{dx}\right) dx = \int_{x'-\epsilon}^{x'+\epsilon} \left(\frac{dT_\epsilon}{dx}\right) dx = \int_{x'-\epsilon}^{x'+\epsilon} \frac{1}{2\epsilon} dx = 1.$$

It is not surprising, then, to find that  $\lim_{\epsilon \rightarrow 0} \frac{dT_\epsilon}{dx}(x - x') = \delta(x - x')$ . Assuming that the interchange of the order of differentiation and the limiting process is justified, we obtain the important identity

$$\frac{d}{dx}\theta(x - x') = \delta(x - x'). \tag{6.7}$$

$\delta$  function as derivative of  $\theta$  function





**Figure 6.3** The step function, or  $\theta$ -function, shown in the figure has the Dirac delta function as its derivative.

Now that we have some understanding of one continuous index, we can generalize the results to several continuous indices. In the earlier discussion we looked at  $f(x)$  as the  $x$ th component of some abstract vector  $|f\rangle$ . For functions of  $n$  variables, we can think of  $f(x_1, \dots, x_n)$  as the component of an abstract vector  $|f\rangle$  along a basis vector  $|x_1, \dots, x_n\rangle$ .<sup>2</sup> This basis is a direct generalization of one continuous index to  $n$ . Then  $f(x_1, \dots, x_n)$  is defined as  $f(x_1, \dots, x_n) = \langle x_1, \dots, x_n | f \rangle$ . If the region of integration is denoted by  $\Omega$ , and we use the abbreviations

$$\mathbf{r} \equiv (x_1, x_2, \dots, x_n), \quad d^n x = dx_1 dx_2 \dots dx_n,$$

$$|x_1, x_2, \dots, x_n\rangle = |\mathbf{r}\rangle, \quad \delta(x_1 - x'_1) \dots \delta(x_n - x'_n) = \delta(\mathbf{r} - \mathbf{r}'),$$

then we can write

$$\begin{aligned} |f\rangle &= \int_{\Omega} d^n x f(\mathbf{r}) w(\mathbf{r}) |\mathbf{r}\rangle, & \int_{\Omega} d^n x |\mathbf{r}\rangle w(\mathbf{r}) \langle \mathbf{r}| &= \mathbf{1}, \\ f(\mathbf{r}') &= \int_{\Omega} d^n x f(\mathbf{r}) w(\mathbf{r}) \langle \mathbf{r}' | \mathbf{r} \rangle, & \langle \mathbf{r}' | \mathbf{r} \rangle w(\mathbf{r}) &= \delta(\mathbf{r} - \mathbf{r}'), \end{aligned} \tag{6.8}$$

where  $d^n x$  is the “volume” element and  $\Omega$  is the region of integration of interest.

For instance, if the region of definition of the functions under consideration is the surface of the unit sphere, then [with  $w(\mathbf{r}) = 1$ ], one gets

$$\int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta |\theta, \phi\rangle \langle \theta, \phi| = \mathbf{1}. \tag{6.9}$$

<sup>2</sup>Do not confuse this with an  $n$ -dimensional vector. In fact, the dimension is  $n$ -fold infinite: each  $x_i$  counts one infinite set of numbers!

This will be used in our discussion of spherical harmonics in Chapter 12.

An important identity using the three-dimensional Dirac delta function comes from potential theory. This is (see [Hass 99] for a discussion of this equation)

$$\nabla^2 \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) = -4\pi \delta(\mathbf{r} - \mathbf{r}'). \quad (6.10)$$

## 6.2 Generalized Functions

Paul Adrian Maurice Dirac discovered the delta function in the late 1920s while investigating scattering problems in quantum mechanics. This “function” seemed to violate most properties of other functions known to mathematicians at the time. Furthermore, the derivative of the delta function,  $\delta'(x - x')$  is such that for any ordinary function  $f(x)$ ,

$$\int_{-\infty}^{\infty} f(x) \delta'(x - x') dx = - \int_{-\infty}^{\infty} f'(x) \delta(x - x') dx = -f'(x').$$

We can *define*  $\delta'(x - x')$  by this relation. In addition, we can define the derivative of any function, including discontinuous functions, at any point (including points of discontinuity, where the usual definition of derivative fails) by this relation. That is, if  $\varphi(x)$  is a “bad” function whose derivative is not defined at some point(s), and  $f(x)$  is a “good” function, we can define the derivative of  $\varphi(x)$  by

$$\int_{-\infty}^{\infty} f(x) \varphi'(x) dx \equiv - \int_{-\infty}^{\infty} f'(x) \varphi(x) dx.$$

The integral on the RHS is well-defined.

Functions such as the Dirac delta function and its derivatives of all orders are not functions in the traditional sense. What is common among all of them is that in most applications they appear inside an integral, and we saw in Chapter 1 that integration can be considered as a linear functional on the space of continuous functions. It is therefore natural to describe such functions in terms of linear functionals. This idea was picked up by Laurent Schwartz in the 1950s who developed it into a new branch of mathematics called **generalized functions**, or **distributions**.

A distribution is a mathematical entity that appears inside an integral in conjunction with a well-behaved **test function**—which we assume to depend on  $n$  variables—such that the result of integration is a well-defined number. Depending on the type of test function used, different kinds of distributions can be defined. If we want to include the Dirac delta function and its derivatives of all orders, then the test functions must be infinitely differentiable, that is, they must be  $\mathcal{C}^\infty$  functions on  $\mathbb{R}^n$  (or  $\mathbb{C}^n$ ). Moreover, in order for the theory of distributions to be mathematically feasible, all the test functions must vanish outside a finite “volume” of  $\mathbb{R}^n$  (or  $\mathbb{C}^n$ ).<sup>3</sup> One common notation for such functions is  $\mathcal{C}_F^\infty(\mathbb{R}^n)$  or  $\mathcal{C}_F^\infty(\mathbb{C}^n)$

<sup>3</sup>Such functions are said to be of **compact support**.

( $F$  stands for “finite”). The definitive property of distributions concerns the way they combine with test functions to give a number. The test functions used clearly form a vector space over  $\mathbb{R}$  or  $\mathbb{C}$ . In this vector-space language, distributions are linear functionals. The linearity is a simple consequence of the properties of the integral. We therefore have the following definition of a distribution.

generalized functions  
and distributions  
defined

**6.2.1. Definition.** A *distribution, or generalized function*, is a continuous<sup>4</sup> linear functional on the space  $\mathcal{C}_F^\infty(\mathbb{R}^n)$  or  $\mathcal{C}_F^\infty(\mathbb{C}^n)$ . If  $f \in \mathcal{C}_F^\infty$  and  $\varphi$  is a distribution, then  $\varphi[f] = \int_{-\infty}^{\infty} \varphi(\mathbf{r}) f(\mathbf{r}) d^n x$ .

Another notation used in place of  $\varphi[f]$  is  $\langle \varphi, f \rangle$ . This is more appealing not only because  $\varphi$  is linear, in the sense that  $\varphi[\alpha f + \beta g] = \alpha \varphi[f] + \beta \varphi[g]$ , but also because the set of all such linear functionals forms a vector space; that is, the linear combination of the  $\varphi$ 's is also defined. Thus,  $\langle \varphi, f \rangle$  suggests a mutual “democracy” for both  $f$ 's and  $\varphi$ 's.

We now have a shorthand way of writing integrals. For instance, if  $\delta_a$  represents the Dirac delta function  $\delta(x - a)$ , with an integration over  $x$  understood, then  $\langle \delta_a, f \rangle = f(a)$ . Similarly,  $\langle \delta'_a, f \rangle = -f'(a)$ , and for linear combinations,  $\langle \alpha \delta_a + \beta \delta'_a, f \rangle = \alpha f(a) - \beta f'(a)$ .

**6.2.2. Example.** An ordinary (continuous) function  $g$  can be thought of as a special case of a distribution. The linear functional  $g : \mathcal{C}_F^\infty(\mathbb{R}) \rightarrow \mathbb{R}$  is simply defined by  $\langle g, f \rangle \equiv g[f] = \int_{-\infty}^{\infty} g(x) f(x) dx$ . ■

**6.2.3. Example.** An interesting application of distributions (generalized functions) occurs when the notion of density is generalized to include not only (smooth) volume densities, but also point-like, linear, and surface densities.

A point charge  $q$  located at  $\mathbf{r}_0$  can be thought of as having a charge density  $\rho(\mathbf{r}) = q \delta(\mathbf{r} - \mathbf{r}_0)$ . In the language of linear functionals, we interpret  $\rho$  as a distribution,  $\rho : \mathcal{C}_F^\infty(\mathbb{R}^3) \rightarrow \mathbb{R}$ , which for an arbitrary function  $f$  gives

$$\rho[f] = \langle \rho, f \rangle = q f(\mathbf{r}_0). \tag{6.11}$$

The delta function character of  $\rho$  can be detected from this equation by recalling that the LHS is

$$\int \rho(\mathbf{r}) f(\mathbf{r}) d^3 x = \lim_{\substack{N \rightarrow \infty \\ \Delta V_i \rightarrow 0}} \sum_{i=1}^N \rho(\mathbf{r}_i) f(\mathbf{r}_i) \Delta V_i.$$

On the RHS of this equation, the only volume element that contributes is the one that contains the point  $\mathbf{r}_0$ ; all the rest contribute zero. As  $\Delta V_i \rightarrow 0$ , the only way that the RHS can give a nonzero number is for  $\rho(\mathbf{r}_0) f(\mathbf{r}_0)$  to be infinite. Since  $f$  is a well-behaved function,  $\rho(\mathbf{r}_0)$  must be infinite, implying that  $\rho(\mathbf{r})$  acts as a delta function. This shows that the definition of Equation (6.11) leads to a delta-function behavior for  $\rho$ . Similarly for linear and surface densities. ■

<sup>4</sup>See [Zeidler, 95], pp. 27. 156–160, for a formal definition of the continuity of linear functionals.



The example above and Problems 6.5 and 6.6 suggest that a distribution that confines an integral to a lower-dimensional space must have a delta function in its definition.

“Physical Laws should have mathematical beauty.” This statement was Dirac’s response to the question of his philosophy of physics, posed to him in Moscow in 1955. He wrote it on a blackboard that is still preserved today.



**Paul Adrien Maurice Dirac** (1902–1984), was born in 1902 in Bristol, England, of a Swiss, French-speaking father and an English mother. His father, a taciturn man who refused to receive friends at home, enforced young Paul’s silence by requiring that only French be spoken at the dinner table. Perhaps this explains Dirac’s later disinclination toward collaboration and his general tendency to be a loner in most aspects of his life. The fundamental nature of his work made the involvement of students difficult, so perhaps Dirac’s personality was well-suited to his extraordinary accomplishments.

Dirac went to Merchant Venturer’s School, the public school where his father taught French, and while there displayed great mathematical abilities. Upon graduation, he followed in his older brother’s footsteps and went to Bristol University to study electrical engineering. He was 19 when he graduated Bristol University in 1921. Unable to find a suitable engineering position due to the economic recession that gripped post-World War I England, Dirac accepted a fellowship to study mathematics at Bristol University. This fellowship, together with a grant from the Department of Scientific and Industrial Research, made it possible for Dirac to go to Cambridge as a research student in 1923. At Cambridge Dirac was exposed to the experimental activities of the Cavendish Laboratory, and he became a member of the intellectual circle over which Rutherford and Fowler presided. He took his Ph.D. in 1926 and was elected in 1927 as a fellow. His appointment as university lecturer came in 1929. He assumed the Lucasian professorship following Joseph Larmor in 1932 and retired from it in 1969. Two years later he accepted a position at Florida State University where he lived out his remaining years. The FSU library now carries his name.

In the late 1920s the relentless march of ideas and discoveries had carried physics to a generally accepted relativistic theory of the electron. Dirac, however, was dissatisfied with the prevailing ideas and, somewhat in isolation, sought for a better formulation. By 1928 he succeeded in finding an equation, the *Dirac equation*, that accorded with his own ideas and also fit most of the established principles of the time. Ultimately, this equation, and the physical theory behind it, proved to be one of the great intellectual achievements of the period. It was particularly remarkable for the internal beauty of its mathematical structure, which not only clarified previously mysterious phenomena such as **spin** and the **Fermi-Dirac** statistics associated with it, but also predicted the existence of an electron-like particle of negative energy, the antielectron, or *positron*, and, more recently, it has come to play a role of great importance in modern mathematics, particularly in the interrelations between topology, geometry, and analysis. Heisenberg characterized the discovery of antimatter by Dirac as “the most decisive discovery in connection with the properties or the nature of elementary particles . . . . This discovery of particles and antiparticles by Dirac . . . changed our whole outlook on atomic physics completely.” One of the interesting implications of

“The amount of theoretical ground one has to cover before being able to solve problems of real practical value is rather large, but this circumstance is an inevitable consequence of the fundamental part played by transformation theory and is likely to become more pronounced in the theoretical physics of the future.”  
P.A.M. Dirac (1930)

his work that predicted the positron was the prediction of a *magnetic monopole*. Dirac won the Nobel Prize in 1933 for this work.

Dirac is not only one of the chief authors of quantum mechanics, but he is also the creator of *quantum electrodynamics* and one of the principal architects of *quantum field theory*. While studying the scattering theory of quantum particles, he invented the (Dirac) *delta function*; in his attempt at quantizing the general theory of relativity, he founded *constrained Hamiltonian dynamics*, which is one of the most active areas of theoretical physics research today. One of his greatest contributions is the invention of *bra*  $\langle |$  and *ket*  $| \rangle$ .

While at Cambridge, Dirac did not accept many research students. Those who worked with him generally thought that he was a good supervisor, but one who did not spend much time with his students. A student needed to be extremely independent to work under Dirac. One such student was Dennis Sciama, who later became the supervisor of Stephen Hawking, the current holder of the Lucasian chair. Salam and Wigner, in their preface to the Festschrift that honors Dirac on his seventieth birthday and commemorates his contributions to quantum mechanics succinctly assessed the man:

Dirac is one of the chief creators of quantum mechanics . . . . Posterity will rate Dirac as one of the greatest physicists of all time. The present generation values him as one of its greatest teachers . . . . On those privileged to know him, Dirac has left his mark . . . by his human greatness. He is modest, affectionate, and sets the highest possible standards of personal and scientific integrity. He is a legend in his own lifetime and rightly so.

(Taken from Schweber, S. S. "Some chapters for a history of quantum field theory: 1938-1952", in *Relativity, Groups, and Topology II* vol. 2, B. S. DeWitt and R. Stora, eds., North-Holland, Amsterdam, 1984.)

We have seen that the delta function can be thought of as the limit of an ordinary function. This idea can be generalized.

**6.2.4. Definition.** Let  $\{\varphi_n(x)\}$  be a sequence of functions such that

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \varphi_n(x) f(x) dx$$

exists for all  $f \in \mathcal{C}_F^\infty(\mathbb{R})$ . Then the sequence is said to converge to the distribution  $\varphi$ , defined by

$$\langle \varphi, f \rangle = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \varphi_n(x) f(x) dx \quad \forall f.$$

This convergence is denoted by  $\varphi_n \rightarrow \varphi$ .

For example, it can be verified that

$$\frac{n}{\sqrt{\pi}} e^{-n^2 x^2} \rightarrow \delta(x) \quad \text{and} \quad \frac{1 - \cos nx}{n\pi x^2} \rightarrow \delta(x)$$

and so on. The proofs are left as exercises.

derivative of a  
distribution

**6.2.5. Definition.** The *derivative* of a distribution  $\varphi$  is another distribution  $\varphi'$  defined by  $\langle \varphi', f \rangle = -\langle \varphi, f' \rangle \quad \forall f \in \mathcal{C}_F^\infty$ .

**6.2.6. Example.** We can combine the last two definitions to show that if the functions  $\theta_n$  are defined as

$$\theta_n(x) \equiv \begin{cases} 0 & \text{if } x \leq -\frac{1}{n}, \\ (nx+1)/2 & \text{if } -\frac{1}{n} \leq x \leq \frac{1}{n}, \\ 1 & \text{if } x \geq \frac{1}{n}, \end{cases}$$

then  $\theta'_n(x) \rightarrow \delta(x)$ .

We write the definition of the derivative,  $\langle \theta'_n, f \rangle = -\langle \theta_n, f' \rangle$ , in terms of integrals:

$$\begin{aligned} \int_{-\infty}^{\infty} \theta'_n(x) f(x) dx &= - \int_{-\infty}^{\infty} \theta_n(x) \frac{df}{dx} dx = - \int_{-\infty}^{\infty} \theta_n(x) df \\ &= - \left( \int_{-\infty}^{-1/n} \theta_n(x) df + \int_{-1/n}^{1/n} \theta_n(x) df + \int_{1/n}^{\infty} \theta_n(x) df \right) \\ &= - \left( 0 + \int_{-1/n}^{1/n} \frac{nx+1}{2} df + \int_{1/n}^{\infty} df \right) \\ &= -\frac{n}{2} \int_{-1/n}^{1/n} x df - \frac{1}{2} \int_{-1/n}^{1/n} df - \int_{1/n}^{\infty} df \\ &= -\frac{n}{2} \left( xf(x) \Big|_{-1/n}^{1/n} - \int_{-1/n}^{1/n} f(x) dx \right) \\ &\quad - \frac{1}{2} (f(1/n) - f(-1/n)) - f(\infty) + f(1/n). \end{aligned}$$

For large  $n$ , we have  $1/n \approx 0$  and  $f(\pm 1/n) \approx f(0)$ . Thus,

$$\int_{-\infty}^{\infty} \theta'_n(x) f(x) dx \approx -\frac{n}{2} \left( \frac{1}{n} f\left(\frac{1}{n}\right) + \frac{1}{n} f\left(-\frac{1}{n}\right) - \frac{2}{n} f(0) \right) + f(0) \approx f(0).$$

The approximation becomes equality in the limit  $n \rightarrow \infty$ . Thus,

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \theta'_n(x) f(x) dx = f(0) = \langle \delta_0, f \rangle \Rightarrow \theta'_n \rightarrow \delta.$$

Note that  $f(\infty) = 0$  because of the assumption that all functions must vanish outside a finite volume. ■

## 6.3 Problems

**6.1.** Write a density function for two point charges  $q_1$  and  $q_2$  located at  $\mathbf{r} = \mathbf{r}_1$  and  $\mathbf{r} = \mathbf{r}_2$ , respectively.

**6.2.** Write a density function for four point charges  $q_1 = q$ ,  $q_2 = -q$ ,  $q_3 = q$  and  $q_4 = -q$ , located at the corners of a square of side  $2a$ , lying in the  $xy$ -plane, whose center is at the origin and whose first corner is at  $(a, a)$ .

**6.3.** Show that  $\delta(f(x)) = \frac{1}{|f'(x_0)|} \delta(x - x_0)$ , where  $x_0$  is a root of  $f$  and  $x$  is confined to values close to  $x_0$ . Hint: Make a change of variable to  $y = f(x)$ .

**6.4.** Show that

$$\delta(f(x)) = \sum_{k=1}^m \frac{1}{|f'(x_k)|} \delta(x - x_k),$$

where the  $x_k$ 's are all the roots of  $f$  in the interval on which  $f$  is defined.

**6.5.** Define the distribution  $\rho : \mathcal{C}^\infty(\mathbb{R}^3) \rightarrow \mathbb{R}$  by

$$\langle \rho, f \rangle = \int_S \sigma(\mathbf{r}) f(\mathbf{r}) da(\mathbf{r}),$$

where  $\sigma(\mathbf{r})$  is a smooth function on a smooth surface  $S$  in  $\mathbb{R}^3$ . Show that  $\rho(\mathbf{r})$  is zero if  $\mathbf{r}$  is not on  $S$  and infinite if  $\mathbf{r}$  is on  $S$ .

**6.6.** Define the distribution  $\rho : \mathcal{C}^\infty(\mathbb{R}^3) \rightarrow \mathbb{R}$  by

$$\langle \rho, f \rangle = \int_C \lambda(\mathbf{r}) f(\mathbf{r}) dl(\mathbf{r}),$$

where  $\lambda(\mathbf{r})$  is a smooth function on a smooth curve  $C$  in  $\mathbb{R}^3$ . Show that  $\rho(\mathbf{r})$  is zero if  $\mathbf{r}$  is not on  $C$  and infinite if  $\mathbf{r}$  is on  $C$ .

**6.7.** Express the three-dimensional Dirac delta function as a product of three one-dimensional delta functions involving the coordinates in

- (a) cylindrical coordinates,
- (b) spherical coordinates,
- (c) general curvilinear coordinates.

Hint: The Dirac delta function in  $\mathbb{R}^3$  satisfies  $\iiint \delta(\mathbf{r}) d^3x = 1$ .

**6.8.** Show that  $\int_{-\infty}^{\infty} \delta'(x) f(x) dx = -f'(0)$  where  $\delta'(x) \equiv \frac{d}{dx} \delta(x)$ .

**6.9.** Evaluate the following integrals:

$$\begin{aligned} (a) \int_{-\infty}^{\infty} \delta(x^2 - 5x + 6)(3x^2 - 7x + 2) dx. & \quad (b) \int_{-\infty}^{\infty} \delta(x^2 - \pi^2) \cos x dx. \\ (c) \int_{0.5}^{\infty} \delta(\sin \pi x) \left(\frac{2}{3}\right)^x dx. & \quad (d) \int_{-\infty}^{\infty} \delta(e^{-x^2}) \ln x dx. \end{aligned}$$

Hint: Use the result of Problem 6.4.

**6.10.** Consider  $|x|$  as a generalized function and find its derivative.

**6.11.** Let  $\eta \in C^\infty(\mathbb{R}^n)$  be a smooth function on  $\mathbb{R}^n$ , and let  $\varphi$  be a distribution. Show that  $\eta\varphi$  is also a distribution. What is the natural definition for  $\eta\varphi$ ? What is  $(\eta\varphi)'$ , the derivative of  $\eta\varphi$ ?

**6.12.** Show that each of the following sequences of functions approaches  $\delta(x)$  in the sense of Definition 6.2.4.

$$(a) \frac{n}{\sqrt{\pi}} e^{-n^2 x^2}. \quad (b) \frac{1 - \cos nx}{\pi n x^2}. \quad (c) \frac{n}{\pi} \frac{1}{1 + n^2 x^2}. \quad (d) \frac{\sin nx}{\pi x}.$$

Hint: Approximate  $\varphi_n(x)$  for large  $n$  and  $x \approx 0$ , and then evaluate the appropriate integral.

**6.13.** Show that  $\frac{1}{2}(1 + \tanh nx) \rightarrow \theta(x)$  as  $n \rightarrow \infty$ .

**6.14.** Show that  $x\delta'(x) = -\delta(x)$ .

## Additional Reading

1. Hassani, S. *Mathematical Methods*, Springer-Verlag, 2000. An elementary treatment of the Dirac delta function with many examples drawn from mechanics and electromagnetism.
2. Rudin, W. *Functional Analysis*, McGraw-Hill, 1991. Part II of this mathematical but (for those with a strong undergraduate mathematics background) very readable book is devoted to the theory of distributions.
3. Reed, M. and Simon, B. *Functional Analysis*, Academic Press, 1980.

# Classical Orthogonal Polynomials

The last example of Chapter 5 discussed only one of the many types of the so-called classical orthogonal polynomials. Historically, these polynomials were discovered as solutions to differential equations arising in various physical problems.

Such polynomials can be produced by starting with  $1, x, x^2, \dots$  and employing the Gram-Schmidt process. However, there is a more elegant, albeit less general, approach that simultaneously studies most polynomials of interest to physicists. We will employ this approach.<sup>1</sup>

## 7.1 General Properties

Most relevant properties of the polynomials of interest are contained in

**7.1.1. Theorem.** *Consider the functions*

$$F_n(x) = \frac{1}{w(x)} \frac{d^n}{dx^n}(ws^n) \quad \text{for } n = 0, 1, 2, \dots, \quad (7.1)$$

where

1.  $F_1(x)$  is a first-degree polynomial in  $x$ ,
2.  $s(x)$  is a polynomial in  $x$  of degree less than or equal to 2 with only real roots,
3.  $w(x)$  is a strictly positive function, integrable in the interval  $(a, b)$ , that satisfies the boundary conditions  $w(a)s(a) = 0 = w(b)s(b)$ .

---

<sup>1</sup>This approach is due to F. G. Tricomi [Tric 55]. See also [Denn 67].

Then  $F_n(x)$  is a polynomial of degree  $n$  in  $x$  and is orthogonal—on the interval  $(a, b)$ , with weight  $w(x)$ —to any polynomial  $p_k(x)$  of degree  $k < n$ , i.e.,  $\int_a^b p_k(x)F_n(x)w(x) dx = 0$  for  $k < n$ . These polynomials are collectively called **classical orthogonal polynomials**.

Before proving the theorem, we need two lemmas:<sup>2</sup>

**7.1.2. Lemma.** *The following identity holds:*

$$\frac{d^m}{dx^m}(ws^n p_{\leq k}) = ws^{n-m} p_{\leq k+m}, \quad m \leq n.$$

*Proof.* See Problem 7.1. □

**7.1.3. Lemma.** *All the derivatives  $d^m/dx^m(ws^n)$  vanish at  $x = a$  and  $x = b$ , for all values of  $m < n$ .*

*Proof.* Set  $k = 0$  in the identity of the previous lemma and let  $p_{\leq 0} = 1$ . Then we have  $\frac{d^m}{dx^m}(ws^n) = ws^{n-m} p_{\leq m}$ . The RHS vanishes at  $x = a$  and  $x = b$  due to the third condition stated in the theorem. □

*Proof of the theorem.* We prove the orthogonality first. The proof involves multiple use of integration by parts:

$$\begin{aligned} \int_a^b p_k(x)F_n(x)w(x) dx &= \int_a^b p_k(x) \frac{1}{w} \left[ \frac{d^n}{dx^n}(ws^n) \right] w dx \\ &= \int_a^b p_k(x) \frac{d}{dx} \left[ \frac{d^{n-1}}{dx^{n-1}}(ws^n) \right] dx \\ &= p_k(x) \underbrace{\frac{d^{n-1}}{dx^{n-1}}(ws^n)} \Big|_a^b - \int_a^b \frac{dp_k}{dx} \frac{d^{n-1}}{dx^{n-1}}(ws^n) dx. \\ & \qquad \qquad \qquad = 0 \text{ by Lemma 7.1.3} \end{aligned}$$

This shows that each integration by parts transfers one differentiation from  $ws^n$  to  $p_k$  and introduces a minus sign. Thus, after  $k$  integrations by parts, we get

$$\begin{aligned} \int_a^b p_k(x)F_n(x)w(x) dx &= (-1)^k \int_a^b \frac{d^k p_k}{dx^k} \frac{d^{n-k}}{dx^{n-k}}(ws^n) dx \\ &= C \int_a^b \frac{d}{dx} \left[ \frac{d^{n-k-1}}{dx^{n-k-1}}(ws^n) \right] dx = C \frac{d^{n-k-1}}{dx^{n-k-1}}(ws^n) \Big|_a^b = 0, \end{aligned}$$

<sup>2</sup>Recall that  $p_{\leq k}$  is a generic polynomial with degree less than or equal to  $k$ .

where we have used the fact that the  $k$ th derivative of a polynomial of degree  $k$  is a constant. Note that  $n - k - 1 \geq 0$  because  $k < n$ , so that the last line of the equation is well-defined.

To prove the first part of the theorem, we use Lemma 7.1.2 with  $k = 0$  and  $m = n$  to get  $\frac{d^n}{dx^n}(ws^n) = wp_{\leq n}$ , or  $F_n(x) = \frac{1}{w} \frac{d^n}{dx^n}(ws^n) = p_{\leq n}$ . To prove that  $F_n(x)$  is a polynomial of degree precisely equal to  $n$ , we write  $F_n(x) = p_{\leq n-1}(x) + k_n x^n$ , multiply both sides by  $w(x)F_n(x)$ , and integrate over  $(a, b)$ :

$$\int_a^b [F_n(x)]^2 w(x) dx = \int_a^b p_{\leq n-1} F_n(x) w(x) dx + k_n \int_a^b x^n F_n(x) w(x) dx.$$

The LHS is a positive quantity because both  $w(x)$  and  $[F_n(x)]^2$  are positive, and the first integral on the RHS vanishes by the first part of the proof. Therefore, the second term on the RHS cannot be zero. In particular,  $k_n \neq 0$ , and  $F_n(x)$  is of degree  $n$ . □

It is customary to introduce a normalization constant in the definition of  $F_n(x)$ , and write

$$F_n(x) = \frac{1}{K_n w} \frac{d^n}{dx^n}(ws^n). \tag{7.2}$$

generalized  
Rodriguez formula

This equation is called the **generalized Rodriguez formula**. For historical reasons, different polynomial functions are normalized differently, which is why  $K_n$  is introduced here.

From Theorem 7.1.1 it is clear that the sequence  $F_0(x), F_1(x), F_2(x), \dots$  of polynomials forms an orthogonal set of polynomials on  $[a, b]$  with weight function  $w(x)$ .

All the varieties of classical orthogonal polynomials were discovered as solutions of differential equations. Here, we give a single generic differential equation satisfied by all the  $F_n$ 's. The proof is outlined in Problem 7.4.

differential equation  
for classical  
orthogonal  
polynomials

**7.1.4. Proposition.** *Let  $k_1$  be the coefficient of  $x$  in  $F_1(x)$  and  $\sigma_2$  the coefficient of  $x^2$  in  $s(x)$ . Then the orthogonal polynomials  $F_n$  satisfy the differential equation<sup>3</sup>*

$$(wsF_n')' = w\lambda_n F_n(x) \quad \text{where} \quad \lambda_n = K_1 k_1 n + \sigma_2 n(n-1).$$

We shall study the differential equation above in the context of the Sturm-Liouville problem (see Chapters 18 and 19), which is an eigenvalue problem involving differential operators.

<sup>3</sup>A prime is a symbol for derivative with respect to  $x$ .



## 7.2 Classification

Let us now investigate the consequences of various choices of  $s(x)$ . We start with  $F_1(x)$ , and note that it satisfies Equation (7.2) with  $n = 1$ :

$$F_1(x) = \frac{1}{K_1 w} \frac{d}{dx}(ws), \quad \text{or} \quad \frac{1}{ws} \frac{d}{dx}(ws) = \frac{K_1 F_1(x)}{s},$$

which can be integrated to yield  $ws = A \exp(\int K_1 F_1(x) dx/s)$  where  $A$  is a constant. On the other hand, being a polynomial of degree 1,  $F_1(x)$  can be written as  $F_1(x) = k_1 x + k'_1$ . It follows that

$$\begin{aligned} w(x)s(x) &= A \exp\left(\int \frac{K_1(k_1 x + k'_1)}{s} dx\right), \\ w(a)s(a) = 0 &= w(b)s(b). \end{aligned} \tag{7.3}$$

Next we look at the three choices for  $s(x)$ : a constant, a polynomial of degree 1, and a polynomial of degree 2. For a constant  $s(x)$ , Equation (7.3) can be easily integrated:

$$\begin{aligned} w(x)s(x) &= A \exp\left(\int \frac{K_1(k_1 x + k'_1)}{s} dx\right) \equiv A \exp\left(\int (2\alpha x + \beta) dx\right) \\ &= A e^{\alpha x^2 + \beta x + C} = B e^{\alpha x^2 + \beta x}. \end{aligned}$$

The interval  $(a, b)$  is determined by  $w(a)s(a) = 0 = w(b)s(b)$ , which yields  $B e^{\alpha a^2 + \beta a} = 0 = B e^{\alpha b^2 + \beta b}$ . The only way that this equality can hold is for  $a$  and  $b$  to be infinite. Since  $a < b$ , we must take  $a = -\infty$  and  $b = +\infty$ , in which case  $\alpha < 0$ . With  $y = \sqrt{|\alpha|}(x + \beta/(2\alpha))$  and choosing  $B = s \exp(\beta^2/(4\alpha))$ , we obtain  $w(y) = \exp(-y^2)$ . We also take the constant  $s$  to be 1. This is always possible by a proper choice of constants such as  $B$ .

If the degree of  $s$  is 1, then  $s(x) = \sigma_1 x + \sigma_0$  and

$$\begin{aligned} w(x)(\sigma_1 x + \sigma_0) &= A \exp\left(\int \frac{K_1(k_1 x + k'_1)}{\sigma_1 x + \sigma_0} dx\right) \\ &= A \exp\left[\int \left(\frac{K_1 k_1}{\sigma_1} + \frac{K_1 k'_1 - K_1 k_1 \sigma_0 / \sigma_1}{\sigma_1 x + \sigma_0}\right) dx\right] \\ &\equiv B(\sigma_1 x + \sigma_0)^\rho e^{\gamma x}, \end{aligned}$$

where  $\gamma = K_1 k_1 / \sigma_1$ ,  $\rho = K_1 k'_1 / \sigma_1 - K_1 k_1 \sigma_0 / \sigma_1^2$ , and  $B$  is  $A$  modified by the constant of integration. The last equation above must satisfy the boundary conditions at  $a$  and  $b$ :  $B(\sigma_1 a + \sigma_0)^\rho e^{\gamma a} = 0 = B(\sigma_1 b + \sigma_0)^\rho e^{\gamma b}$ , which give  $a = -\sigma_0 / \sigma_1$ ,  $\rho > 0$ ,  $\gamma < 0$ , and  $b = +\infty$ . With appropriate redefinition of variables and parameters, we can write  $w(y) = y^\nu e^{-y}$ ,  $\nu > -1$ , and  $s(x) = x$ ,  $a = 0$ ,  $b = +\infty$ .

$\mu$	$\nu$	$w(x)$	Polynomial
0	0	1	Legendre, $P_n(x)$
$\lambda - \frac{1}{2}$	$\lambda - \frac{1}{2}$	$(1 - x^2)^{\lambda-1/2}$	Gegenbauer, $C_n^\lambda(x)$ , $\lambda > -\frac{1}{2}$
$-\frac{1}{2}$	$-\frac{1}{2}$	$(1 - x^2)^{-1/2}$	Chebyshev of the first kind, $T_n(x)$
$\frac{1}{2}$	$\frac{1}{2}$	$(1 - x^2)^{1/2}$	Chebyshev of the second kind, $U_n(x)$

**Table 7.1** Special cases of Jacobi polynomials

Similarly, we can obtain the weight function and the interval of integration for the case when  $s(x)$  is of degree 2. This result, as well as the results obtained above, are collected in the following proposition.

**7.2.1. Proposition.** *If the conditions of Theorem 7.1.1 prevail, then*

- (a) *For  $s(x)$  of degree zero we get  $w(x) = e^{-x^2}$  with  $s(x) = 1$ ,  $a = -\infty$ , and  $b = +\infty$ . The resulting polynomials are called **Hermite polynomials** and are denoted by  $H_n(x)$ .*
- (b) *For  $s(x)$  of degree 1, we obtain  $w(x) = x^\nu e^{-x}$  with  $\nu > -1$ ,  $s(x) = x$ ,  $a = 0$ , and  $b = +\infty$ . The resulting polynomials are called **Laguerre polynomials** and are denoted by  $L_n^\nu(x)$ .*
- (c) *For  $s(x)$  of degree 2, we get  $w(x) = (1 + x)^\mu(1 - x)^\nu$  with  $\mu, \nu > -1$ ,  $s(x) = 1 - x^2$ ,  $a = -1$ , and  $b = +1$ . The resulting polynomials are called **Jacobi polynomials** and are denoted by  $P_n^{\mu,\nu}(x)$ .*

Jacobi polynomials are themselves divided into other subcategories depending on the values of  $\mu$  and  $\nu$ . The most common and widely used of these are collected in Table 7.1. Note that the definition of each of the preceding polynomials involves a "standardization," which boils down to a particular choice of  $K_n$  in the generalized Rodriguez formula.

### 7.3 Recurrence Relations

Besides the recurrence relations obtained in Section 5.2, we can use the differential equation of Proposition 7.1.4 to construct new recurrence relations involving derivatives. These relations apply only to *classical* orthogonal polynomials, and not to general ones. We start with Equation (5.12)

$$F_{n+1}(x) = (\alpha_n x + \beta_n)F_n(x) + \gamma_n F_{n-1}(x), \quad (7.4)$$

differentiate both sides twice, and substitute for the second derivative from the differential equation of Proposition 7.1.4. This will yield

$$2ws\alpha_n F'_n + \left[ \alpha_n \frac{d}{dx}(ws) + w\lambda_n(\alpha_n x + \beta_n) \right] F_n - w\lambda_{n+1}F_{n+1} + w\gamma_n\lambda_{n-1}F_{n-1} = 0. \quad (7.5)$$

**Karl Gustav Jacob Jacobi** (1804–1851) was the second son born to a well-to-do Jewish banking family in Potsdam. An obviously bright young man, Jacobi was soon moved to the highest class in spite of his youth and remained at the gymnasium for four years only because he could not enter the university until he was sixteen. He excelled at the University of Berlin in all the classical subjects as well as mathematical studies, the topic he soon chose as his career. He passed the examination to become a secondary school teacher, then later the examination that allowed university teaching, and joined the faculty at Berlin at the age of twenty. Since promotion there appeared unlikely, he moved in 1826 to the University of Königsberg in search of a more permanent position. He was known as a lively and creative lecturer who often injected his latest research topics into the lectures. He began what is now a common practice at most universities—the research seminar—for the most advanced students and his faculty collaborators. The Jacobi “school,” together with the influence of Bessel and Neumann (also at Königsberg), sparked a renewal of mathematical excellence in Germany.



In 1843 Jacobi fell gravely ill with diabetes. After seeing his condition, Dirichlet, with the help of von Humboldt, secured a donation to enable Jacobi to spend several months in Italy, a therapy recommended by his doctor. The friendly atmosphere and healthful climate there soon improved his condition. Jacobi was later given royal permission to move from Königsberg to Berlin so that his health would not be affected by the harsh winters in the former location. A salary bonus given to Jacobi to offset the higher cost of living in the capital was revoked after he made some politically sensitive remarks in an impromptu speech. A permanent position at Berlin was also refused, and the reduced salary and lack of security caused considerable hardship for Jacobi and his family. Only after he accepted a position in Vienna did the Prussian government recognize the desirability of keeping the distinguished mathematician within its borders, offering him special concessions that together with his love for his homeland convinced Jacobi to stay. In 1851 Jacobi died after contracting both influenza and smallpox.

Jacobi’s mathematical reputation began largely with his heated competition with Abel in the study of elliptic functions. Legendre, formerly the star of such studies, wrote Jacobi of his happiness at having “lived long enough to witness these magnanimous contests between two young athletes equally strong.” Although Jacobi and Abel could reasonably be considered contemporary researchers who arrived at many of the same results independently, Jacobi suggested the names “Abelian functions” and “Abelian theorem” in a review he wrote for Crelle’s Journal. Jacobi also extended his discoveries in elliptic functions to number theory and the theory of integration. He also worked in other areas of number theory, such as the theory of quadratic forms and the representation of integers as sums of squares and cubes. He

presented the well-known *Jacobian*, or functional determinant, in 1841. To physicists, Jacobi is probably best known for his work in dynamics with the form introduced by Hamilton. Although elegant and quite general, Hamiltonian dynamics did not lend itself to easy solution of many practical problems in mechanics. In the spirit of Lagrange, Poisson, and others, Jacobi investigated transformations of Hamilton's equations that preserved their canonical nature (loosely speaking, that preserved the Poisson brackets in each representation). After much work and a little simplification, the resulting equations of motion, now known as *Hamilton-Jacobi equations*, allowed Jacobi to solve several important problems in ordinary and celestial mechanics. Clebsch and later Helmholtz amplified their use in other areas of physics.

We can get another recurrence relation involving derivatives by substituting (7.4) in (7.5) and simplifying:

$$2ws\alpha_n F'_n + \left[ \alpha_n \frac{d}{dx}(ws) + w(\lambda_n - \lambda_{n+1})(\alpha_n x + \beta_n) \right] F_n + w\gamma_n(\lambda_{n-1} - \lambda_{n+1})F_{n-1} = 0. \quad (7.6)$$

Two other recurrence relations can be obtained by differentiating Equations (7.6) and (7.5), respectively, and using the differential equation for  $F_n$ . Now solve the first equation so obtained for  $\gamma_n(d/dx)(wF_{n-1})$  and substitute the result in the second equation. After simplification, the result will be

$$2w\alpha_n\lambda_n F_n + \frac{d}{dx} \left\{ \left[ \alpha_n \frac{d}{dx}(ws) + w(\lambda_n - \lambda_{n-1})(\alpha_n x + \beta_n) \right] F_n \right\} + (\lambda_{n-1} - \lambda_{n+1}) \frac{d}{dx}(wF_{n+1}) = 0. \quad (7.7)$$

Finally, we record one more useful recurrence relation:

$$A_n(x)F_n - \lambda_{n+1}(\alpha_n x + \beta_n) \frac{dw}{dx} F_{n+1} + \gamma_n \lambda_{n-1}(\alpha_n x + \beta_n) \frac{dw}{dx} F_{n-1} + B_n(x)F'_{n+1} + \gamma_n D_n(x)F'_{n-1} = 0, \quad (7.8)$$

where

$$\begin{aligned} A_n(x) &= (\alpha_n x + \beta_n) \left[ 2w\alpha_n\lambda_n + \alpha_n \frac{d^2}{dx^2}(ws) + \lambda_n(\alpha_n x + \beta_n) \frac{dw}{dx} \right] \\ &\quad - \alpha_n^2 \frac{d}{dx}(ws), \\ B_n(x) &= \alpha_n \frac{d}{dx}(ws) - w(\alpha_n x + \beta_n)(\lambda_{n+1} - \lambda_n), \\ D_n(x) &= w(\alpha_n x + \beta_n)(\lambda_{n-1} - \lambda_n) - \alpha_n \frac{d}{dx}(ws). \end{aligned}$$

Details of the derivation of this relation are left for the reader. All these recurrence relations seem to be very complicated. However, complexity is the price we pay

useful recurrence  
relations for Hermite  
and Legendre  
polynomials

for generality. When we work with specific orthogonal polynomials, the equations simplify considerably. For instance, for Hermite and Legendre polynomials Equation (7.6) yields, respectively,

$$H'_n = 2nH_{n-1}, \quad \text{and} \quad (1-x^2)P'_n + nxP_n - nP_{n-1} = 0. \quad (7.9)$$

Also, applying Equation (7.7) to Legendre polynomials gives

$$P'_{n+1} - xP'_n - (n+1)P_n = 0, \quad (7.10)$$

and Equation (7.8) yields

$$P'_{n+1} - P'_{n-1} - (2n+1)P_n = 0. \quad (7.11)$$

It is possible to find many more recurrence relations by manipulating the existing recurrence relations.

Before studying specific orthogonal polynomials, let us pause for a moment to appreciate the generality and elegance of the foregoing discussion. With a few assumptions and a single defining equation we have severely restricted the choice of the weight function and with it the choice of the interval  $(a, b)$ . We have nevertheless exhausted the list of the so-called classical orthogonal polynomials.

## 7.4 Examples of Classical Orthogonal Polynomials

We now construct the specific polynomials used frequently in physics. We have seen that the four parameters  $K_n$ ,  $k_n$ ,  $k'_n$ , and  $h_n$  determine all the properties of the polynomials. Once  $K_n$  is fixed by some standardization, we can determine all the other parameters:  $k_n$  and  $k'_n$  will be given by the generalized Rodriguez formula, and  $h_n$  can be calculated as

$$\begin{aligned} h_n &= \int_a^b F_n^2(x)w(x) dx = \int_a^b (k_n x^n + \dots) F_n(x)w(x) dx \\ &= k_n \int_a^b wx^n \frac{1}{K_n w} \frac{d^n}{dx^n} (ws^n) dx = \frac{k_n}{K_n} \int_a^b x^n \frac{d}{dx} \left[ \frac{d^{n-1}}{dx^{n-1}} (ws^n) \right] dx \\ &= \frac{k_n}{K_n} x^n \frac{d^{n-1}}{dx^{n-1}} (ws^n) \Big|_a^b - \frac{k_n}{K_n} \int_a^b \frac{d}{dx} (x^n) \frac{d^{n-1}}{dx^{n-1}} (ws^n) dx \end{aligned}$$

The first term of the last line is zero by Lemma 7.1.3. It is clear that each integration by parts introduces a minus sign and shifts one differentiation from  $ws^n$  to  $x^n$ . Thus, after  $n$  integrations by parts and noting that  $d^0/dx^0(ws^n) = ws^n$  and  $d^n/dx^n(x^n) = n!$ , we obtain

$$h_n = \frac{(-1)^n k_n n!}{K_n} \int_a^b ws^n dx. \quad (7.12)$$

### 7.4.1 Hermite Polynomials

summary of  
properties of Hermite  
polynomials

The Hermite polynomials are standardized such that  $K_n = (-1)^n$ . Thus, the generalized Rodriguez formula (7.2) and Proposition 7.2.1 give

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}). \quad (7.13)$$

It is clear that each time  $e^{-x^2}$  is differentiated, a factor of  $-2x$  is introduced. The highest power of  $x$  is obtained when we differentiate  $e^{-x^2}$   $n$  times. This yields  $(-1)^n e^{x^2} (-2x)^n e^{-x^2} = 2^n x^n \Rightarrow k_n = 2^n$ .

To obtain  $k'_n$ , we find it helpful to see whether the polynomial is even or odd. We substitute  $-x$  for  $x$  in Equation (7.13) and get  $H_n(-x) = (-1)^n H_n(x)$ , which shows that if  $n$  is even (odd),  $H_n$  is an even (odd) polynomial, i.e., it can have only even (odd) powers of  $x$ . In either case, the next-highest power of  $x$  in  $H_n(x)$  is not  $n-1$  but  $n-2$ . Thus, the coefficient of  $x^{n-1}$  is zero for  $H_n(x)$ , and we have  $k'_n = 0$ . For  $h_n$ , we use (7.12) to obtain  $h_n = \sqrt{\pi} 2^n n!$ .

Next we calculate the recurrence relation of Equation (5.12). We can readily calculate the constants needed:  $\alpha_n = 2$ ,  $\beta_n = 0$ ,  $\gamma_n = -2n$ . Then substitute these in Equation (5.12) to obtain

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x). \quad (7.14)$$

Other recurrence relations can be obtained similarly.

Finally, the differential equation of  $H_n(x)$  is obtained by first noting that  $K_1 = -1$ ,  $\sigma_2 = 0$ ,  $F_1(x) = 2x \Rightarrow k_1 = 2$ . All of this gives  $\lambda_n = -2n$ , which can be used in the equation of Proposition 7.1.4 to get

$$\frac{d^2 H_n}{dx^2} - 2x \frac{dH_n}{dx} + 2nH_n = 0. \quad (7.15)$$

### 7.4.2 Laguerre Polynomials

summary of  
properties of  
Laguerre  
polynomials

For Laguerre polynomials, the standardization is  $K_n = n!$ . Thus, the generalized Rodriguez formula (7.2) and Proposition 7.2.1 give

$$L_n^v(x) = \frac{1}{n! x^v e^{-x}} \frac{d^n}{dx^n} (x^v e^{-x} x^n) = \frac{1}{n!} x^{-v} e^x \frac{d^n}{dx^n} (x^{n+v} e^{-x}). \quad (7.16)$$

To find  $k_n$  we note that differentiating  $e^{-x}$  does not introduce any new powers of  $x$  but only a factor of  $-1$ . Thus, the highest power of  $x$  is obtained by leaving  $x^{n+v}$  alone and differentiating  $e^{-x}$   $n$  times. This gives

$$\frac{1}{n!} x^{-v} e^x x^{n+v} (-1)^n e^{-x} = \frac{(-1)^n}{n!} x^n \Rightarrow k_n = \frac{(-1)^n}{n!}.$$

We may try to check the evenness or oddness of  $L_n^\nu(x)$ ; however, this will not be helpful because changing  $x$  to  $-x$  distorts the RHS of Equation (7.16). In fact,  $k'_n \neq 0$  in this case, and it can be calculated by noticing that the next-highest power of  $x$  is obtained by adding the first derivative of  $x^{n+\nu}$   $n$  times and multiplying the result by  $(-1)^{n-1}$ , which comes from differentiating  $e^{-x}$ . We obtain

$$\frac{1}{n!} x^{-\nu} e^x [(-1)^{n-1} n(n+\nu) x^{n+\nu-1} e^{-x}] = \frac{(-1)^{n-1} (n+\nu)}{(n-1)!} x^{n-1},$$

and therefore  $k'_n = (-1)^{n-1} (n+\nu)/(n-1)!$ .

Finally, for  $h_n$  we get

$$h_n = \frac{(-1)^n [(-1)^n / n!] n!}{n!} \int_0^\infty x^\nu e^{-x} x^n dx = \frac{1}{n!} \int_0^\infty x^{n+\nu} e^{-x} dx.$$

If  $\nu$  is not an integer (and it need not be), the integral on the RHS cannot be evaluated by elementary methods. In fact, this integral occurs so frequently in mathematical applications that it is given a special name, the **gamma function**. A detailed discussion of this function can be found in Chapter 11. At this point, we simply note that

$$\Gamma(z+1) \equiv \int_0^\infty x^z e^{-x} dx, \quad \Gamma(n+1) = n! \quad \text{for } n \in \mathbb{N}, \quad (7.17)$$

and write  $h_n$  as

$$h_n = \frac{\Gamma(n+\nu+1)}{n!} = \frac{\Gamma(n+\nu+1)}{\Gamma(n+1)}.$$

The relevant parameters for the recurrence relation can be easily calculated:

$$\alpha_n = -\frac{1}{n+1}, \quad \beta_n = \frac{2n+\nu+1}{n+1}, \quad \gamma_n = -\frac{n+\nu}{n+1}.$$

Substituting these in Equation (5.12) and simplifying yields

$$(n+1)L_{n+1}^\nu = (2n+\nu+1-x)L_n^\nu - (n+\nu)L_{n-1}^\nu.$$

With  $k_1 = -1$  and  $\sigma_2 = 0$ , we get  $\lambda_n = -n$ , and the differential equation of Proposition 7.1.4 becomes

$$x \frac{d^2 L_n^\nu}{dx^2} + (\nu+1-x) \frac{dL_n^\nu}{dx} + nL_n^\nu = 0. \quad (7.18)$$

### 7.4.3 Legendre Polynomials

summary of  
properties of  
Legendre  
polynomials

Instead of discussing the Jacobi polynomials as a whole, we will discuss a special case of them, the Legendre polynomials  $P_n(x)$ , which are more widely used in physics.

With  $\mu = 0 = \nu$ , corresponding to the Legendre polynomials, the weight function for the Jacobi polynomials reduces to  $w(x) = 1$ . The standardization is  $K_n = (-1)^n 2^n n!$ . Thus, the generalized Rodriguez formula reads

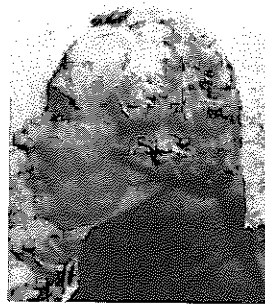
$$P_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} [(1-x^2)^n]. \quad (7.19)$$

To find  $k_n$ , we expand the expression in square brackets using the binomial theorem and take the  $n$ th derivative of the highest power of  $x$ . This yields

$$\begin{aligned} k_n x^n &= \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} [(-x^2)^n] = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^{2n}) \\ &= \frac{1}{2^n n!} 2n(2n-1)(2n-2) \cdots (n+1)x^n. \end{aligned}$$

After some algebra (see Problem 7.7), we get  $k_n = \frac{2^n \Gamma(n + \frac{1}{2})}{n! \Gamma(\frac{1}{2})}$ .

**Adrien-Marie Legendre** (1752–1833) came from a well-to-do Parisian family and received an excellent education in science and mathematics. His university work was advanced enough that his mentor used many of Legendre's essays in a treatise on mechanics. A man of modest fortune until the revolution, Legendre was able to devote himself to study and research without recourse to an academic position. In 1782 he won the prize of the Berlin Academy for calculating the trajectories of cannonballs taking air resistance into account. This essay brought him to the attention of Lagrange and helped pave the way to acceptance in French scientific circles, notably the Academy of Sciences, to which Legendre submitted numerous papers. In July 1784 he submitted a paper on planetary orbits that contained the now-famous *Legendre polynomials*, mentioning that Lagrange had been able to "present a more complete theory" in a recent paper by using Legendre's results. In the years that followed, Legendre concentrated his efforts in number theory, celestial mechanics, and the theory of elliptic functions. In addition, he was a prolific calculator, producing large tables of the values of special functions, and he also authored an elementary textbook that remained in use for many decades. In 1824 Legendre refused to vote for the government's candidate for *Institut National*. Because of this, his pension was stopped and he died in poverty and in pain at the age of 80 after several years of failing health.



Legendre produced a large number of useful ideas but did not always develop them in the most rigorous manner, claiming to hold the priority for an idea if he had presented



merely a reasonable argument for it. Gauss, with whom he had several quarrels over priority, considered rigorous proof the standard of ownership. To Legendre's credit, however, he was an enthusiastic supporter of his young rivals Abel and Jacobi and gave their work considerable attention in his writings. Especially in the theory of elliptic functions, the area of competition with Abel and Jacobi, Legendre is considered more of a trailblazer than a great builder. Hermite wrote that Legendre "is considered the founder of the theory of elliptic functions" and "greatly smoothed the way for his successors," but notes that the recognition of the double periodicity of the inverse function, which allowed the great progress of others, was missing from Legendre's work.

Legendre also contributed to practical efforts in science and mathematics. He and two of his contemporaries were assigned in 1787 to a panel conducting geodetic work in cooperation with the observatories at Paris and Greenwich. Four years later the same panel members were appointed as the Academy's commissioners to undertake the measurements and calculations necessary to determine the length of the standard meter. Legendre's seemingly tireless skill at calculating produced large tables of the values of trigonometric and elliptic functions, logarithms, and solutions to various special equations.

In his famous textbook *Eléments de géométrie* (1794) he gave a simple proof that  $\pi$  is irrational and conjectured that it is not the root of any algebraic equation of finite degree with rational coefficients. The textbook was somewhat dogmatic in its presentation of ordinary Euclidean thought and includes none of the non-Euclidean ideas beginning to be formed around that time. It was Legendre who first gave a rigorous proof of the theorem (assuming all of Euclid's postulates, of course) that the sum of the angles of a triangle is "equal to two right angles." Very little of his research in this area was of memorable quality. The same could possibly be argued for the balance of his writing, but one must acknowledge the very fruitful ideas he left behind in number theory and elliptic functions and, of course, the introduction of Legendre polynomials and the important *Legendre transformation* used both in thermodynamics and Hamiltonian mechanics.

To find  $k'_n$ , we look at the evenness or oddness of the polynomials. By an investigation of the Rodriguez formula—as in our study of Hermite polynomials—we note that  $P_n(-x) = (-1)^n P_n(x)$ , which tells us that  $P_n(x)$  is either even or odd. In either case,  $x$  will not have an  $(n-1)$ st power. Therefore,  $k'_n = 0$ .

We now calculate  $h_n$  as given by (7.12):

$$h_n = \frac{(-1)^n k_n n!}{K_n} \int_{-1}^1 (1-x^2)^n dx = \frac{2^n \Gamma(n + \frac{1}{2}) / \Gamma(\frac{1}{2})}{2^n n!} \int_{-1}^1 (1-x^2)^n dx.$$

The integral can be evaluated by repeated integration by parts (see Problem 7.8). Substituting the result in the expression above yields  $h_n = 2/(2n+1)$ .

We need  $\alpha_n$ ,  $\beta_n$  and  $\gamma_n$  for the recurrence relation:

$$\alpha_n = \frac{k_{n+1}}{k_n} = \frac{2^{n+1} \Gamma(n+1 + \frac{1}{2})}{(n+1)! \Gamma(\frac{1}{2})} \frac{n! \Gamma(\frac{1}{2})}{2^n \Gamma(n + \frac{1}{2})} = \frac{2n+1}{n+1},$$

where we used the relation  $\Gamma(n+1 + \frac{1}{2}) = (n + \frac{1}{2}) \Gamma(n + \frac{1}{2})$ . We also have  $\beta_n = 0$  (because  $k'_n = 0 = k'_{n+1}$ ) and  $\gamma_n = -n/(n+1)$ . Therefore, the recurrence relation

is

$$(n + 1)P_{n+1}(x) = (2n + 1)xP_n(x) - nP_{n-1}(x). \quad (7.20)$$

Now we use  $K_1 = -2$ ,  $P_1(x) = x \Rightarrow k_1 = 1$ , and  $\sigma_2 = -1$  to obtain  $\lambda_n = -n(n + 1)$ , which yields the following differential equation:

$$\frac{d}{dx} \left[ (1 - x^2) \frac{dP_n}{dx} \right] = -n(n + 1)P_n. \quad (7.21)$$

This can also be expressed as

$$(1 - x^2) \frac{d^2 P_n}{dx^2} - 2x \frac{dP_n}{dx} + n(n + 1)P_n = 0. \quad (7.22)$$

### 7.4.4 Other Classical Orthogonal Polynomials

The rest of the classical orthogonal polynomials can be constructed similarly. For the sake of completeness, we merely quote the results.

#### Jacobi Polynomials, $P_n^{\mu, \nu}(x)$

**Standardization:**  $K_n = (-2)^n n!$

**Constants:**  $k_n = 2^{-n} \frac{\Gamma(2n + \mu + \nu + 1)}{n! \Gamma(n + \mu + \nu + 1)}$ ,  $k'_n = \frac{n(\nu - \mu)}{2n + \mu + \nu} k_n$ ,

$$h_n = \frac{2^{\mu+\nu+1} \Gamma(n + \mu + 1) \Gamma(n + \nu + 1)}{n! (2n + \mu + \nu + 1) \Gamma(n + \mu + \nu + 1)}$$

**Rodriguez formula:**

$$P_n^{\mu, \nu}(x) = \frac{(-1)^n}{2^n n!} (1 + x)^{-\mu} (1 - x)^{-\nu} \frac{d^n}{dx^n} [(1 + x)^{\mu+n} (1 - x)^{\nu+n}]$$

**Differential Equation:**

$$(1 - x^2) \frac{d^2 P_n^{\mu, \nu}}{dx^2} + [\mu - \nu - (\mu + \nu + 2)x] \frac{dP_n^{\mu, \nu}}{dx} + n(n + \mu + \nu + 1)P_n^{\mu, \nu} = 0$$

**A Recurrence Relation:**

$$\begin{aligned} & 2(n + 1)(n + \mu + \nu + 1)(2n + \mu + \nu)P_{n+1}^{\mu, \nu} \\ &= (2n + \mu + \nu + 1)[(2n + \mu + \nu)(2n + \mu + \nu + 2)x + \nu^2 - \mu^2]P_n^{\mu, \nu} \\ & \quad - 2(n + \mu)(n + \nu)(2n + \mu + \nu + 2)P_{n-1}^{\mu, \nu} \end{aligned}$$

### Gegenbauer Polynomials, $C_n^\lambda(x)$

**Standardization:**  $K_n = (-2)^n n! \frac{\Gamma(n + \lambda + \frac{1}{2})\Gamma(2\lambda)}{\Gamma(n + 2\lambda)\Gamma(\lambda + \frac{1}{2})}$

**Constants:**  $k_n = \frac{2^n \Gamma(n + \lambda)}{n! \Gamma(\lambda)}$ ,  $k'_n = 0$ ,  $h_n = \frac{\sqrt{\pi}\Gamma(n + 2\lambda)\Gamma(\lambda + \frac{1}{2})}{n!(n + \lambda)\Gamma(2\lambda)\Gamma(\lambda)}$

**Rodriguez Formula:**

$$C_n^\lambda(x) = \frac{(-1)^n \Gamma(n + 2\lambda)\Gamma(\lambda + \frac{1}{2})}{2^n n! \Gamma(n + \lambda + \frac{1}{2})\Gamma(2\lambda)} (1 - x^2)^{-\lambda + 1/2} \frac{d^n}{dx^n} [(1 - x^2)^{n + \lambda - 1/2}]$$

**Differential Equation:**

$$(1 - x^2) \frac{d^2 C_n^\lambda}{dx^2} - (2\lambda + 1)x \frac{dC_n^\lambda}{dx} + n(n + 2\lambda)C_n^\lambda = 0$$

**A Recurrence Relation:**

$$(n + 1)C_{n+1}^\lambda = 2(n + \lambda)x C_n^\lambda - (n + 2\lambda - 1)C_{n-1}^\lambda$$

### Chebyshev Polynomials of the First Kind, $T_n(x)$

**Standardization:**  $K_n = (-1)^n \frac{(2n)!}{2^n n!}$

**Constants:**  $k_n = 2^{n-1}$ ,  $k'_n = 0$ ,  $h_n = \frac{\pi}{2}$

**Rodriguez Formula:**  $T_n(x) = \frac{(-1)^n 2^n n!}{(2n)!} (1 - x^2)^{1/2} \frac{d^n}{dx^n} [(1 - x^2)^{n-1/2}]$

**Differential Equation:**  $(1 - x^2) \frac{d^2 T_n}{dx^2} - x \frac{dT_n}{dx} + n^2 T_n = 0$

**A Recurrence Relation:**  $T_{n+1} = 2xT_n - T_{n-1}$

### Chebyshev Polynomials of the Second Kind, $U_n(x)$

**Standardization:**  $K_n = (-1)^n \frac{(2n + 1)!}{2^n (n + 1)!}$

**Constants:**  $k_n = 2^n$ ,  $k'_n = 0$ ,  $h_n = \frac{\pi}{2}$

**Rodriguez Formula:**

$$U_n(x) = \frac{(-1)^n 2^n (n+1)!}{(2n+1)!} (1-x^2)^{-1/2} \frac{d^n}{dx^n} [(1-x^2)^{n+1/2}]$$

**Differential Equation:**  $(1-x^2) \frac{d^2 U_n}{dx^2} - 3x \frac{dU_n}{dx} + n(n+2)U_n = 0$

**A Recurrence Relation:**  $U_{n+1} = 2xU_n - U_{n-1}$

## 7.5 Expansion in Terms of Orthogonal Polynomials

Having studied the different classical orthogonal polynomials, we can now use them to write an arbitrary function  $f \in \mathcal{L}_w^2(a, b)$  as a series of these polynomials. If we denote a complete set of orthogonal (not necessarily classical) polynomials by  $|C_k\rangle$  and the given function by  $|f\rangle$ , we may write

$$|f\rangle = \sum_{k=0}^{\infty} a_k |C_k\rangle, \quad (7.23)$$

where  $a_k$  is found by multiplying both sides of the equation by  $\langle C_i|$  and using the orthogonality of the  $|C_k\rangle$ 's:

$$\langle C_i | f \rangle = \sum_{k=0}^{\infty} a_k \langle C_i | C_k \rangle = a_i \langle C_i | C_i \rangle \Rightarrow a_i = \frac{\langle C_i | f \rangle}{\langle C_i | C_i \rangle}. \quad (7.24)$$

This is written in function form as

$$a_i = \frac{\int_a^b C_i^*(x) f(x) w(x) dx}{\int_a^b |C_i(x)|^2 w(x) dx}. \quad (7.25)$$

We can also “derive” the functional form of Equation (7.23) by multiplying both of its sides by  $\langle x|$  and using the fact that  $\langle x| f \rangle = f(x)$  and  $\langle x| C_k \rangle = C_k(x)$ . The result will be

$$f(x) = \sum_{k=0}^{\infty} a_k C_k(x). \quad (7.26)$$

**7.5.1. Example.** The solution of Laplace’s equation in spherically symmetric electrostatic problems that are independent of the azimuthal angle is given by

$$\Phi(r, \theta) = \sum_{k=0}^{\infty} \left( \frac{b_k}{r^{k+1}} + c_k r^k \right) P_k(\cos \theta). \quad (7.27)$$

Consider two conducting hemispheres of radius  $a$  separated by a small insulating gap at the equator. The upper hemisphere is held at potential  $V_0$  and the lower one at  $-V_0$ , as

shown in Figure 7.1. We want to find the potential at points outside the resulting sphere. Since the potential must vanish at infinity, we expect the second term in Equation (7.27) to be absent, i.e.,  $c_k = 0 \forall k$ . To find  $b_k$ , substitute  $a$  for  $r$  in (7.27) and let  $\cos \theta \equiv x$ . Then,

$$\Phi(a, x) = \sum_{k=0}^{\infty} \frac{b_k}{a^{k+1}} P_k(x),$$

where

$$\Phi(a, x) = \begin{cases} -V_0 & \text{if } -1 < x < 0, \\ +V_0 & \text{if } 0 < x < 1. \end{cases}$$

From Equation (7.25), we have

$$\begin{aligned} \frac{b_k}{a^{k+1}} &= \frac{\int_{-1}^1 P_k(x) \Phi(a, x) dx}{\underbrace{\int_{-1}^1 |P_k(x)|^2 dx}_{=h_k}} = \frac{2k+1}{2} \int_{-1}^1 P_k(x) \Phi(a, x) dx \\ &= \frac{2k+1}{2} V_0 \left[ -\int_{-1}^0 P_k(x) dx + \int_0^1 P_k(x) dx \right]. \end{aligned}$$

To proceed, we rewrite the first integral:

$$\int_{-1}^0 P_k(x) dx = -\int_{+1}^0 P_k(-y) dy = \int_0^1 P_k(-y) dy = (-1)^k \int_0^1 P_k(x) dx,$$

where we made use of the parity property of  $P_k(x)$ . Therefore,

$$\frac{b_k}{a^{k+1}} = \frac{2k+1}{2} V_0 [1 - (-1)^k] \int_0^1 P_k(x) dx.$$

It is now clear that only odd polynomials contribute to the expansion. Using the result of Problem 7.26, we get

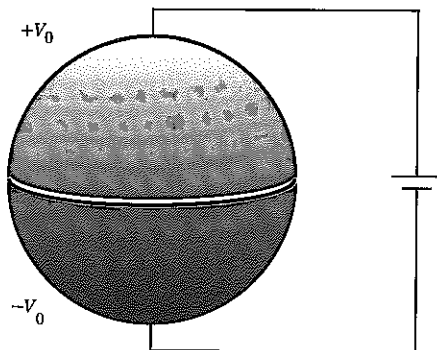
$$\frac{b_k}{a^{k+1}} = (-1)^{(k-1)/2} \frac{(2k+1)(k-1)!}{2^k \left(\frac{k+1}{2}\right)! \left(\frac{k-1}{2}\right)!} V_0, \quad k \text{ odd},$$

or

$$b_{2m+1} = (4m+3)a^{2m+2} V_0 (-1)^m \frac{(2m)!}{2^{2m+1} m!(m+1)!}.$$

Note that  $\Phi(a, x)$  is an odd function; that is,  $\Phi(a, -x) = -\Phi(a, x)$  as is evident from its definition. Thus, only odd polynomials appear in the expansion of  $\Phi(a, x)$  to preserve this property. Having found the coefficients, we can write the potential:

$$\Phi(r, \theta) = V_0 \sum_{k=0}^{\infty} (-1)^m \frac{(4m+3)(2m)!}{2^{2m+1} m!(m+1)!} \left(\frac{a}{r}\right)^{2m+2} P_{2m+1}(\cos \theta). \quad \blacksquare$$



**Figure 7.1** The voltage is  $+V_0$  for the upper hemisphere, where  $0 \leq \theta < \pi/2$ , or where  $0 < \cos \theta \leq 1$ . It is  $-V_0$  for the lower hemisphere, where  $\pi/2 < \theta \leq \pi$ , or where  $-1 \leq \cos \theta < 0$ .

The place where Legendre polynomials appear most naturally is, as mentioned above, in the solution of Laplace's equation in spherical coordinates. After the partial differential equation is transformed into three ordinary differential equations using the method of the separation of variables, the differential equation corresponding to the polar angle  $\theta$  gives rise to solutions of which Legendre polynomials are special cases. This differential equation simplifies to Legendre differential equation if the substitution  $x = \cos \theta$  is made; in that case, the solutions will be Legendre polynomials in  $x$ , or in  $\cos \theta$ . That is why the argument of  $P_k(x)$  is restricted to the interval  $[-1, +1]$ .

**7.5.2. Example.** We can expand the Dirac delta function in terms of Legendre polynomials. We write

$$\delta(x) = \sum_{n=0}^{\infty} a_n P_n(x), \quad \text{where} \quad a_n = \frac{2n+1}{2} \int_{-1}^1 P_n(x) \delta(x) dx = \frac{2n+1}{2} P_n(0).$$

For odd  $n$  this will give zero, because  $P_n(x)$  is an odd polynomial. To evaluate  $P_n(0)$  for even  $n$ , we use the recurrence relation (7.20) for  $x = 0$ :

$$(n+1)P_{n+1}(0) = -nP_{n-1}(0),$$

or  $nP_n(0) = -(n-1)P_{n-2}(0)$ , or  $P_n(0) = -\frac{n-1}{n}P_{n-2}(0)$ . Iterating this  $m$  times, we obtain

$$P_n(0) = (-1)^m \frac{(n-1)(n-3) \cdots (n-2m+1)}{n(n-2)(n-4) \cdots (n-2m+2)} P_{n-2m}(0).$$

For  $n = 2m$ , this yields  $P_{2m}(0) = (-1)^m \frac{(2m-1)(2m-3) \cdots 3 \cdot 1}{2m(2m-2) \cdots 4 \cdot 2} P_0(0)$ . Now we "fill the gaps" in the numerator by multiplying it—and the denominator, of course—by the

denominator. This yields

$$\begin{aligned}
 P_{2m}(0) &= (-1)^m \frac{2m(2m-1)(2m-2) \cdots 3 \cdot 2 \cdot 1}{[2m(2m-2) \cdots 4 \cdot 2]^2} \\
 &= (-1)^m \frac{(2m)!}{[2^m m!]^2} = (-1)^m \frac{(2m)!}{2^{2m} (m!)^2},
 \end{aligned}$$

because  $P_0(x) = 1$ . Thus, we can write

$$\delta(x) = \sum_{m=0}^{\infty} \frac{4m+1}{2} (-1)^m \frac{(2m)!}{2^{2m} (m!)^2} P_{2m}(x).$$

We can also derive this expansion as follows. For any complete set of orthonormal vectors  $\{|f_k\rangle\}_{k=1}^{\infty}$ , we have

$$\begin{aligned}
 \delta(x-x') &= w(x) \langle x|x' \rangle = w(x) \langle x|1|x' \rangle \\
 &= w(x) \langle x| \left( \sum_k |f_k\rangle \langle f_k| \right) |x' \rangle = w(x) \sum_k f_k^*(x') f_k(x).
 \end{aligned}$$

Legendre polynomials are not orthonormal; but we can make them so by dividing  $P_k(x)$  by  $h_k^{1/2} = \sqrt{2/(2k+1)}$ . Then, noting that  $w(x) = 1$ , we obtain

$$\delta(x-x') = \sum_{k=0}^{\infty} \frac{P_k(x')}{\sqrt{2/(2k+1)}} \frac{P_k(x)}{\sqrt{2/(2k+1)}} = \sum_{k=0}^{\infty} \frac{2k+1}{2} P_k(x') P_k(x).$$

For  $x' = 0$  we get  $\delta(x) = \sum_{k=0}^{\infty} \frac{2k+1}{2} P_k(0) P_k(x)$ , which agrees with the previous result. ■

## 7.6 Generating Functions

generating function

It is possible to generate all orthogonal polynomials of a certain kind from a single function of two variables  $g(x, t)$  by repeated differentiation of that function. Such a function is called a **generating function**. This generating function is assumed to be expandable in the form

$$g(x, t) = \sum_{n=0}^{\infty} a_n t^n F_n(x), \tag{7.28}$$

so that the  $n$ th derivative of  $g(x, t)$  with respect to  $t$  evaluated at  $t = 0$  gives  $F_n(x)$  to within a multiplicative constant. The constant  $a_n$  is introduced for convenience. Clearly, for  $g(x, t)$  to be useful, it must be in closed form. The derivation of such a function for general  $F_n(x)$  is nontrivial, and we shall not attempt to derive such a general generating function—as we did, for instance, for the general Rodriguez formula. Instead, we simply quote these functions in Table 7.2, and leave the derivation of the generating functions of Hermite and Legendre polynomials as Problems 7.14 and 7.20. For the derivation of Laguerre generating function, see [Hassani, 2000] pp. 606–607.

Polynomial	Generating function	$a_n$
Hermite, $H_n(x)$	$\exp(-t^2 + 2xt)$	$1/n!$
Laguerre, $L_n^\nu(x)$	$\exp[-xt/(1-t)]/(1-t)^{\nu+1}$	1
Legendre, $P_n(x)$	$(t^2 - 2xt + 1)^{-1/2}$	1
Chebyshev (1st kind), $T_n(x)$	$(1-t^2)(t^2 - 2xt + 1)^{-1}$	$2, n \neq 0$ $a_0 = 1$
Chebyshev (2nd kind), $U_n(x)$	$(t^2 - 2xt + 1)^{-1}$	1

Table 7.2 Generating functions for selected polynomials

## 7.7 Problems

7.1. Let  $n = 1$  in Equation (7.1) and solve for  $s \frac{dw}{dx}$ . Now substitute this in the derivative of  $ws^n p_{\leq k}$  and show that the derivative is equal to  $ws^{n-1} p_{\leq k+1}$ . Repeat this process  $m$  times to prove Lemma 7.1.2.

7.2. Find  $w(x)$ ,  $a$ , and  $b$  for the case of the classical orthogonal polynomials in which  $s(x)$  is of second degree.

7.3. Integrate by parts twice and use Lemma 7.1.2 to show that

$$\int_a^b F_m (ws F_n')' dx = 0 \quad \text{for } m < n.$$

7.4. (a) Using Lemma 7.1.2 conclude that  $(ws F_n')'/w$  is a polynomial of degree less than or equal to  $n$ .

(b) Write  $(ws F_n')'/w$  as a linear combination of  $F_i(x)$ , and use their orthogonality and Problem 7.3 to show that the linear combination collapses to a single term.

(c) Multiply both sides of the differential equation so obtained by  $F_n$  and integrate. The RHS becomes  $h_n \lambda_n$ . For the LHS, carry out the differentiation and note that  $(ws)'/w = K_1 F_1$ . Now show that  $K_1 F_1 F_n' + s F_n''$  is a polynomial of degree  $n$ , and that the LHS of the differential equation yields  $\{(K_1 k_1 n + \sigma_2 n(n-1))h_n\}$ . Now find  $\lambda_n$ .

7.5. Derive the recurrence relation of Equation (7.8). Hint: Differentiate Equation (7.5) and substitute for  $F_n''$  from the differential equation. Now multiply the resulting equation by  $\alpha_n x + \beta_n$  and substitute for  $(\alpha_n x + \beta_n) F_n'$  from one of the earlier recurrence relations.



7.6. Using only the orthogonality of Hermite polynomials

$$\int_{-\infty}^{\infty} e^{-x^2} H_m(x) H_n(x) dx = \sqrt{\pi} 2^n n! \delta_{mn}$$

(and the fact that they are polynomials) generate the first three of them.

7.7. Show that for Legendre polynomials,  $k_n = 2^n \Gamma(n + \frac{1}{2}) / [n! \Gamma(\frac{1}{2})]$ . Hint: Multiply and divide the expression given in the book by  $n!$ ; take a factor of 2 out of all terms in the numerator; the even terms yield a factor of  $n!$ , and the odd terms give a gamma function.

7.8. Using integration by parts several times, show that

$$\int_{-1}^1 (1-x^2)^n dx = \frac{2^m n(n-1) \cdots (n-m+1)}{3 \cdot 5 \cdot 7 \cdots (2m-1)} \int_{-1}^1 x^{2m} (1-x^2)^{n-m} dx.$$

Now show that  $\int_{-1}^1 (1-x^2)^n dx = 2\Gamma(\frac{1}{2})n! / [(2n+1)\Gamma(n+\frac{1}{2})]$ .

7.9. Use the generalized Rodriguez formula for Hermite polynomials and integration by parts to expand  $x^{2k}$  and  $x^{2k+1}$  in terms of Hermite polynomials.

7.10. Use the recurrence relation for Hermite polynomials to show that

$$\int_{-\infty}^{\infty} x e^{-x^2} H_m(x) H_n(x) dx = \sqrt{\pi} 2^{n-1} n! [\delta_{m,n-1} + 2(n+1)\delta_{m,n+1}].$$

What happens when  $m = n$ ?

7.11. Apply the general formalism of the recurrence relations given in the book to Hermite polynomials to find the following:

$$H_n + H'_{n-1} - 2xH_{n-1} = 0.$$

7.12. Show that  $\int_{-\infty}^{\infty} x^2 e^{-x^2} H_n^2(x) dx = \sqrt{\pi} 2^n (n + \frac{1}{2})n!$

7.13. Use a recurrence relations for Hermite polynomials to show that

$$H_n(0) = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ (-1)^m \frac{(2m)!}{m!} & \text{if } n = 2m. \end{cases}$$

7.14. Differentiate the expansion of  $g(x, t)$  for Hermite polynomials with respect to  $x$  (treating  $t$  as a constant) and choose  $a_n$  such that  $na_n = a_{n-1}$  to obtain a differential equation for  $g$ . Solve this differential equation. To determine the "constant" of integration use the result of Problem 7.13 to show that  $g(0, t) = e^{-t^2}$ .

**7.15.** Use the expansion of the generating function for Hermite polynomials to obtain

$$\sum_{m,n=0}^{\infty} e^{-x^2} H_m(x) H_n(x) \frac{s^m t^n}{m!n!} = e^{-x^2+2x(s+t)-(s^2+t^2)}.$$

Then integrate both sides over  $x$  and use the orthogonality of the Hermite polynomials to get

$$\sum_{n=0}^{\infty} \frac{(st)^n}{(n!)^2} \int_{-\infty}^{\infty} e^{-x^2} H_n^2(x) dx = \sqrt{\pi} e^{2st}.$$

Deduce from this the normalization constant  $h_n$  of  $H_n(x)$ .

**7.16.** Using the recurrence relation of Equation (7.14) repeatedly, show that

$$\int_{-\infty}^{\infty} x^k e^{-x^2} H_m(x) H_{m+n}(x) dx = \begin{cases} 0 & \text{if } n > k, \\ \sqrt{\pi} 2^m (m+k)! & \text{if } n = k. \end{cases}$$

**7.17.** Given that  $P_0(x) = 1$  and  $P_1(x) = x$ ,

(a) use (7.20) repeatedly to show that  $P_n(1) = 1$ .

(b) Using the same equation, find  $P_2(x)$ ,  $P_3(x)$ , and  $P_4(x)$ .

**7.18.** Apply the general formalism of the recurrence relations given in the book to find the following two relations for Legendre polynomials:

(a)  $nP_n - xP'_n + P'_{n-1} = 0$ .

(b)  $(1-x^2)P'_n - nP_{n-1} + nxP_n = 0$ .

**7.19.** Show that  $\int_{-1}^1 x^n P_n(x) dx = 2^{n+1} (n!)^2 / (2n+1)!$ . Hint: Use the definition of  $h_n$  and  $k_n$  and the fact that  $P_n$  is orthogonal to any polynomial of degree lower than  $n$ .

**7.20.** Differentiate the expansion of  $g(x, t)$  for Legendre polynomials, and choose  $a_n = 1$ . For  $P'_n$ , you will substitute two different expressions to get two equations. First use Equation (7.11) with  $n+1$  replaced by  $n$ , to obtain

$$(1-t^2) \frac{dg}{dx} + tg = 2 \sum_{n=2}^{\infty} nt^n P_{n-1} + 2t.$$

As an alternative, use Equation (7.10) to substitute for  $P'_n$  and get

$$(1-xt) \frac{dg}{dx} = \sum_{n=2}^{\infty} nt^n P_{n-1} + t.$$

Combine the last two equations to get  $(t^2 - 2xt + 1)g' = tg$ . Solve this differential equation and determine the constant of integration by evaluating  $g(x, 0)$ .

**7.21.** Use the generating function for Legendre polynomials to show that  $P_n(1) = 1$ ,  $P_n(-1) = (-1)^n$ ,  $P_n(0) = 0$  for odd  $n$ , and  $P'_n(1) = n(n+1)/2$ .

**7.22.** Both electrostatic and gravitational potential energies depend on the quantity  $1/|\mathbf{r} - \mathbf{r}'|$ , where  $\mathbf{r}'$  is the position of the source (charge or mass) and  $\mathbf{r}$  is the observation point.

(a) Let  $\mathbf{r}$  lie along the  $z$ -axis, and use spherical coordinates and the definition of generating functions to show that

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{r_{>}} \sum_{n=0}^{\infty} \left(\frac{r_{<}}{r_{>}}\right)^n P_n(\cos \theta),$$

where  $r_{<}(r_{>})$  is the smaller (larger) of  $r$  and  $r'$ , and  $\theta$  is the polar angle.

(b) The electrostatic or gravitational potential energy  $\Phi(\mathbf{r})$  is given by

$$\Phi(\mathbf{r}) = k \iiint \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3x',$$

where  $k$  is a constant and  $\rho(\mathbf{r}')$  is the (charge or mass) density function. Use the result of part (a) to show that if the density depends only on  $r'$ , and not on any angle (i.e.,  $\rho$  is **spherically symmetric**), then  $\Phi(\mathbf{r})$  reduces to the potential energy of a point charge at the origin for  $r > r'$ .

(c) What is  $\Phi(\mathbf{r})$ —in the form of an integral—for  $r < a$  for a spherically symmetric density that extends from origin to  $a$ ?

(d) Show that  $\mathbf{E}$  (or  $\mathbf{g}$ ) is given by  $[kQ(r)/r^2]\hat{\mathbf{e}}_r$ , where  $Q(r)$  is the charge (or mass) enclosed in a sphere of radius  $r$ .

**7.23.** Use the generating function for Legendre polynomials and their orthogonality to derive the relation

$$\int_{-1}^1 \frac{dx}{1 - 2xt + t^2} = \sum_{n=0}^{\infty} t^{2n} \int_{-1}^1 P_n^2(x) dx.$$

Integrate the LHS, expand the result in powers of  $t$ , and compare these powers on both sides to obtain the normalization constant  $h_n$ .

**7.24.** Evaluate the following integrals using the expansion of the generating function for Legendre polynomials.

$$(a) \int_0^\pi \frac{(a \cos \theta + b) \sin \theta d\theta}{\sqrt{a^2 + 2ab \cos \theta + b^2}} \quad (b) \int_0^\pi \frac{(a \cos^2 \theta + b \sin^2 \theta) \sin \theta d\theta}{\sqrt{a^2 + 2ab \cos \theta + b^2}}$$

**7.25.** Differentiate the expansion of the Legendre polynomial generating function with respect to  $x$  and manipulate the resulting expression to obtain

$$(1 - 2xt + t^2) \sum_{n=0}^{\infty} t^n P'_n(x) = t \sum_{n=0}^{\infty} t^n P_n(x).$$

Equate equal powers of  $t$  on both sides to derive the recurrence relation

$$P'_{n+1} + P'_{n-1} - 2xP'_n - P_n = 0.$$

7.26. Show that

$$\int_0^1 P_k(x) dx = \begin{cases} \delta_{k0} & \text{if } k \text{ is even,} \\ \frac{(-1)^{(k-1)/2}(k-1)!}{2^k(\frac{k-1}{2})!(\frac{k+1}{2})!} & \text{if } k \text{ is odd.} \end{cases}$$

Hint: For even  $k$ , extend the region of integration to  $(-1, 1)$  and use the orthogonality property. For odd  $k$ , note that

$$\frac{d^{k-1}}{dx^{k-1}} (1-x^2)^k \Big|_0^1$$

gives zero for the upper limit (by Lemma 7.1.3). For the lower limit, expand the expression using the binomial theorem, and carry out the differentiation, keeping in mind that only one term of the expansion contributes.

7.27. Show that  $g(x, t) = g(-x, -t)$  for both Hermite and Legendre polynomials.

Now expand  $g(x, t)$  and  $g(-x, -t)$  and compare the coefficients of  $t^n$  to obtain the **parity relations** for these polynomials:

$$H_n(-x) = (-1)^n H_n(x) \quad \text{and} \quad P_n(-x) = (-1)^n P_n(x).$$

7.28. Derive the orthogonality of Legendre polynomials directly from the differential equation they satisfy.

7.29. Expand  $|x|$  in the interval  $(-1, +1)$  in terms of Legendre polynomials. Hint: Use the result of Problem 7.26.

7.30. Apply the general formalism of the recurrence relations given in the book to find the following two relations for Laguerre polynomials:

$$(a) \quad nL_n^\nu - (n + \nu)L_{n-1}^\nu - x \frac{dL_n^\nu}{dx} = 0.$$

$$(b) \quad (n + 1)L_{n+1}^\nu - (2n + \nu + 1 - x)L_n^\nu + (n + \nu)L_{n-1}^\nu = 0.$$

7.31. From the generating function for Laguerre polynomials given in Table 7.2 deduce that  $L_n^\nu(0) = \Gamma(n + \nu + 1)/[n!\Gamma(\nu + 1)]$ .

7.32. Let  $L_n \equiv L_n^0$ . Now differentiate both sides of

$$g(x, t) = \frac{e^{-xt/(1-t)}}{1-t} = \sum_0^\infty t^n L_n(x)$$

with respect to  $x$  and compare powers of  $t$  to obtain  $L'_n(0) = -n$  and  $L''_n(0) = \frac{1}{2}n(n-1)$ . Hint: Differentiate  $1/(1-t) = \sum_{n=0}^\infty t^n$  to get an expression for  $(1-t)^{-2}$ .

7.33. Expand  $e^{-kx}$  as a series of Laguerre polynomials  $L_n^\nu(x)$ . Find the coefficients by using (a) the orthogonality of  $L_n^\nu(x)$  and (b) the generating function.

7.34. Derive the recurrence relations given in the book for Jacobi, Gegenbauer, and Chebyshev polynomials.

7.35. Show that  $T_n(-x) = (-1)^n T_n(x)$  and  $U_n(-x) = (-1)^n U_n(x)$ . Hint: Use  $g(x, t) = g(-x, -t)$ .

7.36. Show that  $T_n(1) = 1$ ,  $U_n(1) = n + 1$ ,  $T_n(-1) = (-1)^n$ ,  $U_n(-1) = (-1)^n(n + 1)$ ,  $T_{2m}(0) = (-1)^m = U_{2m}(0)$ , and  $T_{2m+1}(0) = 0 = U_{2m+1}(0)$ .

## Additional Reading

1. Dennery, P. and Krzywicki, A. *Mathematics for Physicists*, Harper and Row, 1967. Treats the classical orthogonal polynomials in the spirit of this chapter.
2. Tricomi, F. *Vorlesungen über Orthogonalreihen*, Springer, 1955. The original unified treatment of the classical orthogonal polynomials.

# Fourier Analysis

The single most recurring theme of mathematical physics is Fourier analysis. It shows up, for example, in classical mechanics and the analysis of normal modes, in electromagnetic theory and the frequency analysis of waves, in noise considerations and thermal physics, in quantum theory and the transformation between momentum and coordinate representations, and in relativistic quantum field theory and creation and annihilation operation formalism.

## 8.1 Fourier Series

One way to begin the study of Fourier series and transforms is to invoke a generalization of the Stone-Weierstrass Approximation Theorem (Theorem 5.2.3), which established the completeness of monomials,  $x^k$ . The generalization of Theorem 5.2.3 permits us to find another set of orthogonal functions in terms of which we can expand an arbitrary function. This generalization involves polynomials in more than one variable (For a proof of this theorem, see Simmons [Simm 83, pp 160–161].)

generalized  
Stone-Weierstrass  
theorem

**8.1.1. Theorem.** (generalized Stone-Weierstrass theorem) *Suppose that  $f(x_1, x_2, \dots, x_n)$  is continuous in the domain  $\{a_i \leq x_i \leq b_i\}_{i=1}^n$ . Then it can be expanded in terms of the monomials  $x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$ , where the  $k_i$  are nonnegative integers.*

Now let us consider functions that are periodic and investigate their expansion in terms of elementary periodic functions. We use the generalized Stone-Weierstrass theorem with two variables,  $x$  and  $y$ . A function  $g(x, y)$  can be written as  $g(x, y) = \sum_{k,m=0}^{\infty} a_{km} x^k y^m$ . In this equation,  $x$  and  $y$  can be considered as coordinates in the  $xy$ -plane, which in turn can be written in terms of polar coordinates

$r$  and  $\theta$ . In that case, we obtain

$$f(r, \theta) \equiv g(r \cos \theta, r \sin \theta) = \sum_{k,m=0}^{\infty} a_{km} r^{k+m} \cos^k \theta \sin^m \theta.$$

In particular, if we let  $r = 1$ , we obtain a function of  $\theta$  alone, which upon substitution of complex exponentials for  $\sin \theta$  and  $\cos \theta$  becomes

$$f(\theta) = \sum_{k,m=0}^{\infty} a_{km} \frac{1}{2^k} (e^{i\theta} + e^{-i\theta})^k \frac{1}{(2i)^m} (e^{i\theta} - e^{-i\theta})^m = \sum_{n=-\infty}^{\infty} b_n e^{in\theta}, \quad (8.1)$$

where  $b_n$  is a constant that depends on  $a_{km}$ . The RHS of (8.1) is periodic with period  $2\pi$ ; thus, it is especially suitable for periodic functions  $f(\theta)$  that satisfy the periodicity condition  $f(\theta - \pi) = f(\theta + \pi)$ .

We can also write Equation (8.1) as

$$\begin{aligned} f(\theta) &= b_0 + \sum_{n=1}^{\infty} (b_n e^{in\theta} + b_{-n} e^{-in\theta}) \\ &= b_0 + \sum_{n=1}^{\infty} [ \underbrace{(b_n + b_{-n})}_{\equiv A_n} \cos n\theta + i \underbrace{(b_n - b_{-n})}_{\equiv B_n} \sin n\theta ] \\ &= b_0 + \sum_{n=1}^{\infty} (A_n \cos n\theta + B_n \sin n\theta). \end{aligned} \quad (8.2)$$

If  $f(\theta)$  is real, then  $b_0$ ,  $A_n$ , and  $B_n$  are also real. Equation (8.1) or (8.2) is called the **Fourier series expansion** of  $f(\theta)$ .

Let us now concentrate on the elementary periodic functions  $e^{in\theta}$ . We define the  $\{|e_n\rangle\}_{n=1}^{\infty}$  such that their " $\theta$ th components" are given by

$$\langle \theta | e_n \rangle = \frac{1}{\sqrt{2\pi}} e^{in\theta}, \quad \text{where } \theta \in (-\pi, \pi).$$

These functions—or ket vectors—which belong to  $\mathcal{L}^2(-\pi, \pi)$ , are orthonormal, as can be easily verified. It can also be shown that they are complete. In fact, for functions that are *continuous* on  $(-\pi, \pi)$ , this is a result of the generalized Stone-Weierstrass theorem. It turns out, however, that  $\{|e_n\rangle\}_{n=1}^{\infty}$  is also a complete orthonormal sequence for *piecewise continuous* functions on  $(-\pi, \pi)$ .<sup>1</sup> Therefore, any periodic piecewise continuous function of  $\theta$  can be expressed as a linear combination of these orthonormal vectors. Thus if  $|f\rangle \in \mathcal{L}^2(-\pi, \pi)$ , then

$$|f\rangle = \sum_{n=-\infty}^{\infty} f_n |e_n\rangle, \quad \text{where } f_n = \langle e_n | f \rangle. \quad (8.3)$$

<sup>1</sup>A piecewise continuous function on a finite interval is one that has a finite number of discontinuities in its interval of definition.

Fourier series  
expansion: angular  
expression

We can write this as a functional relation if we take the  $\theta$ th component of both sides:  $\langle \theta | f \rangle = \sum_{n=-\infty}^{\infty} f_n \langle \theta | e_n \rangle$ , or

$$f(\theta) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} f_n e^{in\theta} \tag{8.4}$$

with  $f_n$  given by

$$\begin{aligned} f_n &= \langle e_n | \mathbf{1} | f \rangle = \langle e_n | \left( \int_{-\pi}^{\pi} |\theta\rangle \langle \theta| d\theta \right) | f \rangle = \int_{-\pi}^{\pi} \langle e_n | \theta \rangle \langle \theta | f \rangle d\theta \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-in\theta} f(\theta) d\theta. \end{aligned} \tag{8.5}$$

fundamental cell of a  
periodic function

It is important to note that even though  $f(\theta)$  may be defined only for  $-\pi \leq \theta \leq \pi$ , Equation (8.4) extends the domain of definition of  $f(\theta)$  to all the intervals  $(2k - 1)\pi \leq \theta \leq (2k + 1)\pi$  for all  $k \in \mathbb{Z}$ . Thus, if a function is to be represented by Equation (8.4) without any specification of the interval of definition, it must be periodic in  $\theta$ . For such functions, the interval of their definition can be translated by a factor of  $2\pi$ . Thus,  $f(\theta)$  with  $-\pi \leq \theta \leq \pi$  is equivalent to  $f(\theta - 2m\pi)$  with  $2m\pi - \pi \leq \theta \leq 2m\pi + \pi$ ; both will give the same Fourier series expansion. We shall define periodic functions in their **fundamental cell** such as  $(-\pi, \pi)$ .

"The profound study  
of nature is the most  
fruitful source of  
mathematical  
discoveries."  
Joseph Fourier

**Joseph Fourier** (1768–1830) did very well as a young student of mathematics but had set his heart on becoming an army officer. Denied a commission because he was the son of a tailor, he went to a Benedictine school with the hope that he could continue studying mathematics at its seminary in Paris. The French Revolution changed those plans and set the stage for many of the personal circumstances of Fourier's later years, due in part to his courageous defense of some of its victims, an action that led to his arrest in 1794. He was released later that year, and he enrolled as a student in the Ecole Normale, which opened and closed within a year. His performance there, however, was enough to earn him a position as assistant lecturer (under Lagrange and Monge) in the *Ecole Polytechnique*. He was an excellent mathematical physicist, was a friend of Napoleon (so far as such people have friends), and accompanied him in 1798 to Egypt, where Fourier held various diplomatic and administrative posts while also conducting research. Napoleon took note of his accomplishments and, on Fourier's return to France in 1801, appointed him prefect of the district of Isère, in southeastern France, and in this capacity built the first real road from Grenoble to Turin. He also befriended the boy Champollion, who later deciphered the *Rosetta stone* as the first long step toward understanding the hieroglyphic writing of the ancient Egyptians.



Like other scientists of his time, Fourier took up the flow of heat. The flow was of interest as a practical problem in the handling of metals in industry and as a scientific problem in attempts to determine the temperature in the interior of the earth, the variation



of that temperature with time, and other such questions. He submitted a basic paper on heat conduction to the Academy of Sciences of Paris in 1807. The paper was judged by Lagrange, Laplace, and Legendre. The paper was not published, mainly due to the objections of Lagrange, who had earlier rejected the use of trigonometric series. But the Academy did wish to encourage Fourier to develop his ideas, and so made the problem of the propagation of heat the subject of a grand prize to be awarded in 1812. Fourier submitted a revised paper in 1811, which was judged by the men already mentioned and others. It won the prize but was criticized for its lack of rigor and so was not published at that time in the *Mémoires* of the Academy.

He developed a mastery of clear notation, some of which is still in use today. (The modern integral sign and the placement of the limits of integration near its top and bottom were introduced by Fourier.) It was also his habit to maintain close association between mathematical relations and physically measurable quantities, especially in limiting or asymptotic cases, even performing some of the experiments himself. He was one of the first to begin full incorporation of physical constants into his equations, and made considerable strides toward the modern ideas of units and dimensional analysis.

Fourier continued to work on the subject of heat and, in 1822, published one of the classics of mathematics, *Théorie Analytique de la Chaleur*, in which he made extensive use of the series that now bear his name and incorporated the first part of his 1811 paper practically without change. Two years later he became secretary of the Academy and was able to have his 1811 paper published in its original form in the *Mémoires*.

**Fourier series** were of profound significance in connection with the evolution of the concept of a function, the rigorous theory of definite integrals, and the development of **Hilbert spaces**. Fourier claimed that "arbitrary" graphs can be represented by trigonometric series and should therefore be treated as legitimate functions, and it came as a shock to many that he turned out to be right. The classical definition of the definite integral due to Riemann was first given in his fundamental paper of 1854 on the subject of Fourier series. Hilbert thought of a function as represented by an infinite sequence, the Fourier coefficients of the function.

Fourier himself is one of the fortunate few: his name has become rooted in all civilized languages as an adjective that is well-known to physical scientists and mathematicians in every part of the world.

Functions are not always defined on  $(-\pi, \pi)$ . Let us consider a function  $F(x)$  that is defined on  $(a, b)$  and is periodic with period  $L = b - a$ . We define a new variable,

$$\theta \equiv \frac{2\pi}{L} \left( x - a - \frac{L}{2} \right) \Rightarrow x = \frac{L}{2\pi} \theta + a + \frac{L}{2},$$

and note that  $f(\theta) \equiv F((L/2\pi)\theta + a + L/2)$  has period  $(-\pi, \pi)$  because

$$f(\theta \pm \pi) = F\left(\frac{L}{2\pi}(\theta \pm \pi) + a + \frac{L}{2}\right) = F\left(x \pm \frac{L}{2}\right)$$

and  $F(x + L/2) = F(x - L/2)$ . It follows that we can expand the latter as in Equation (8.4). Using that equation, but writing  $\theta$  in terms of  $x$ , we obtain

Fourier series  
expansion: general  
expression

$$\begin{aligned}
 F(x) &= F\left(\frac{L}{2\pi}\theta + a + \frac{L}{2}\right) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} f_n \exp\left[in\frac{2\pi}{L}\left(x - a - \frac{L}{2}\right)\right] \\
 &= \frac{1}{\sqrt{L}} \sum_{n=-\infty}^{\infty} F_n e^{2n\pi ix/L}, \tag{8.6}
 \end{aligned}$$

where we have introduced<sup>2</sup>  $F_n \equiv \sqrt{L/2\pi} f_n e^{-i(2\pi n/L)(a+L/2)}$ . Using Equation (8.5), we can write

$$\begin{aligned}
 F_n &= \sqrt{\frac{L}{2\pi}} e^{-i(2\pi n/L)(a+L/2)} \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-in\theta} f(\theta) d\theta \\
 &= \frac{\sqrt{L}}{2\pi} e^{-i(2\pi n/L)(a+L/2)} \int_a^{a+L} e^{-i(2\pi n/L)(x-a-L/2)} F(x) \frac{2\pi}{L} dx \\
 &= \frac{1}{\sqrt{L}} \int_a^b e^{-i(2\pi n/L)x} F(x) dx. \tag{8.7}
 \end{aligned}$$

The functions  $\exp(2\pi inx/L)/\sqrt{L}$  are easily seen to be orthonormal as members of  $\mathcal{L}^2(a, b)$ . We can introduce  $\{|e_n\rangle\}_{n=1}^{\infty}$  with the “ $x$ th component” given by  $\langle x|e_n\rangle = (1/\sqrt{L})e^{2\pi inx/L}$ . Then the reader may check that Equations (8.6) and (8.7) can be written as  $|F\rangle = \sum_{n=-\infty}^{\infty} F_n |e_n\rangle$  with  $F_n = \langle n|F\rangle$ .

square wave voltage

**8.1.2. Example.** In the study of electrical circuits, periodic voltage signals of different shapes are encountered. An example is a **square wave** voltage of height  $U_0$ , “duration”  $T$ , and “rest duration”  $T$  [see Figure 8.1(a)]. The potential as a function of time  $V(t)$  can be expanded as a Fourier series. The interval is  $(0, 2T)$  because that is one whole cycle of the potential variation. We therefore use Equation (8.6) and write

$$V(t) = \frac{1}{\sqrt{2T}} \sum_{n=-\infty}^{\infty} V_n e^{2n\pi it/2T}, \quad \text{where} \quad V_n = \frac{1}{\sqrt{2T}} \int_0^{2T} e^{-2n\pi it/2T} V(t) dt.$$

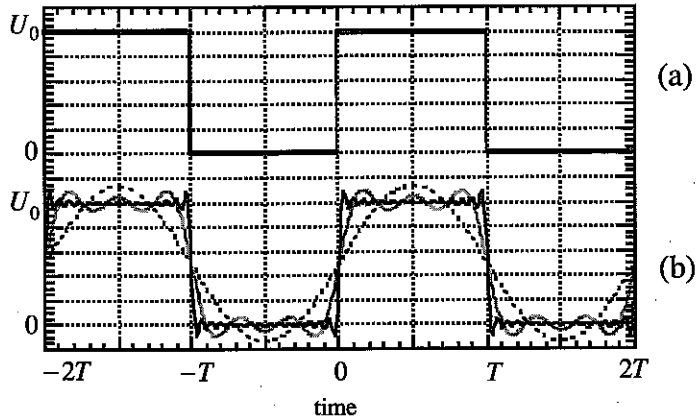
The problem is to find  $V_n$ . This is easily done by substituting

$$V(t) = \begin{cases} U_0 & \text{if } 0 \leq t \leq T, \\ 0 & \text{if } T \leq t \leq 2T \end{cases}$$

in the last integral:

$$\begin{aligned}
 V_n &= \frac{U_0}{\sqrt{2T}} \int_0^T e^{-in\pi t/T} dt = \frac{U_0}{\sqrt{2T}} \left(-\frac{T}{in\pi}\right) [(-1)^n - 1] \quad \text{where } n \neq 0 \\
 &= \begin{cases} 0 & \text{if } n \text{ is even and } n \neq 0, \\ \frac{\sqrt{2T} U_0}{in\pi} & \text{if } n \text{ is odd.} \end{cases}
 \end{aligned}$$

<sup>2</sup>The  $F_n$  are defined such that what they multiply in the expansion are orthonormal in the interval  $(a, b)$ .



**Figure 8.1** (a) The periodic square wave potential. (b) Various approximations to the Fourier series of the square-wave potential. The dashed plot is that of the first term of the series, the thick grey plot keeps 3 terms, and the solid plot 15 terms.

For  $n = 0$ , we obtain  $V_0 = \frac{1}{\sqrt{2T}} \int_0^{2T} V(t) dt = \frac{1}{\sqrt{2T}} \int_0^T U_0 dt = U_0 \sqrt{\frac{T}{2}}$ . Therefore, we can write

$$\begin{aligned} V(t) &= \frac{1}{\sqrt{2T}} \left[ U_0 \sqrt{\frac{T}{2}} + \frac{\sqrt{2T} U_0}{i\pi} \left( \sum_{\substack{n=-\infty \\ n \text{ odd}}}^{-1} \frac{1}{n} e^{in\pi t/T} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{1}{n} e^{in\pi t/T} \right) \right] \\ &= U_0 \left\{ \frac{1}{2} + \frac{1}{i\pi} \left[ \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{1}{-n} e^{-in\pi t/T} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{1}{n} e^{in\pi t/T} \right] \right\} \\ &= U_0 \left\{ \frac{1}{2} + \frac{2}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \sin\left(\frac{[2k+1]\pi t}{T}\right) \right\}. \end{aligned}$$

Figure 8.1(b) shows the graphical representation of the above sum when only a finite number of terms are present. ■

sawtooth voltage

**8.1.3. Example.** Another frequently used voltage is the **sawtooth voltage** [see Figure 8.2(a)]. The equation for  $V(t)$  with period  $T$  is  $V(t) = U_0 t/T$  for  $0 \leq t \leq T$ , and its Fourier representation is

$$V(t) = \frac{1}{\sqrt{T}} \sum_{n=-\infty}^{\infty} V_n e^{2n\pi i t/T}, \quad \text{where} \quad V_n = \frac{1}{\sqrt{T}} \int_0^T e^{-2n\pi i t/T} V(t) dt.$$

Substituting for  $V(t)$  in the integral above yields

$$\begin{aligned} V_n &= \frac{1}{\sqrt{T}} \int_0^T e^{-2\pi int/T} U_0 \frac{t}{T} dt = U_0 T^{-3/2} \int_0^T e^{-2\pi int/T} t dt \\ &= U_0 T^{-3/2} \left( \frac{Tt}{-i2n\pi} e^{-2\pi int/T} \Big|_0^T + \frac{T}{i2n\pi} \underbrace{\int_0^T e^{-2\pi int/T} dt}_{=0} \right) \\ &= U_0 T^{-3/2} \left( \frac{T^2}{-i2n\pi} \right) = -\frac{U_0 \sqrt{T}}{i2n\pi} \quad \text{where } n \neq 0, \\ V_0 &= \frac{1}{\sqrt{T}} \int_0^T V(t) dt = \frac{1}{\sqrt{T}} \int_0^T U_0 \frac{t}{T} dt = \frac{1}{2} U_0 \sqrt{T}. \end{aligned}$$

Thus,

$$\begin{aligned} V(t) &= \frac{1}{\sqrt{T}} \left[ \frac{1}{2} U_0 \sqrt{T} - \frac{U_0 \sqrt{T}}{i2\pi} \left( \sum_{n=-\infty}^{-1} \frac{1}{n} e^{i2n\pi t/T} + \sum_{n=1}^{\infty} \frac{1}{n} e^{i2n\pi t/T} \right) \right] \\ &= U_0 \left\{ \frac{1}{2} - \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin \left( \frac{2n\pi t}{T} \right) \right\}. \end{aligned}$$

Figure 8.2(b) shows the graphical representation of the above series keeping the first few terms. ■

The foregoing examples indicate an important fact about Fourier series. At points of discontinuity (for example,  $t = T$  in the preceding two examples), the value of the function is not defined, but the Fourier series expansion assigns it a value—the average of the two values on the right and left of the discontinuity. For instance, when we substitute  $t = T$  in the series of Example 8.1.3, all the sine terms vanish and we obtain  $V(T) = U_0/2$ , the average of  $U_0$  (on the left) and 0 (on the right). We express this as

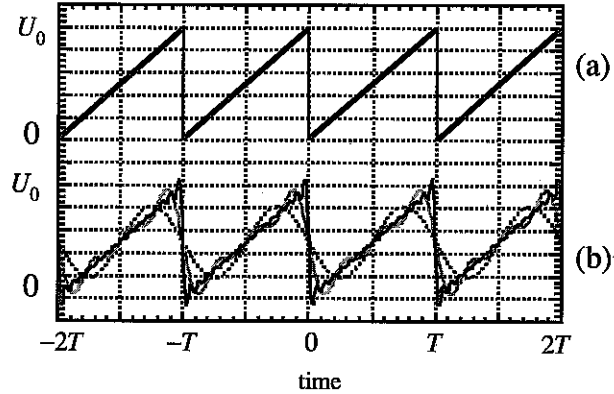
$$V(T) = \frac{1}{2} [V(T-0) + V(T+0)] \equiv \frac{1}{2} \lim_{\epsilon \rightarrow 0} [V(T-\epsilon) + V(T+\epsilon)].$$

This is a general property of Fourier series. In fact, the main theorem of Fourier series, which follows, incorporates this property. (For a proof of this theorem, see [Cour 62].)

**8.1.4. Theorem.** *The Fourier series of a function  $f(\theta)$  that is piecewise continuous in the interval  $(-\pi, \pi)$  converges to*

$$\begin{aligned} &\frac{1}{2} [f(\theta+0) + f(\theta-0)] \quad \text{for } -\pi < \theta < \pi, \\ &\frac{1}{2} [f(\pi) + f(-\pi)] \quad \text{for } \theta = \pm\pi. \end{aligned}$$

Although we used exponential functions to find the Fourier expansion of the two examples above, it is more convenient to start with the trigonometric series



**Figure 8.2** (a) The periodic saw-tooth potential. (b) Various approximations to the Fourier series of the sawtooth potential. The dashed plot is that of the first term of the series, the thick grey plot keeps 3 terms, and the solid plot 15 terms.

when the expansion of a real function is sought. Equation (8.2) already gives such an expansion. All we need to do now is find expressions for  $A_n$  and  $B_n$ . From the definitions of  $A_n$  and the relation between  $b_n$  and  $f_n$  we get

$$\begin{aligned}
 A_n &= b_n + b_{-n} = \frac{1}{\sqrt{2\pi}}(f_n + f_{-n}) \\
 &= \frac{1}{\sqrt{2\pi}} \left( \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-in\theta} f(\theta) d\theta + \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{in\theta} f(\theta) d\theta \right) \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} [e^{-in\theta} + e^{in\theta}] f(\theta) d\theta = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos n\theta f(\theta) d\theta.
 \end{aligned} \tag{8.8}$$

Similarly,

$$\begin{aligned}
 B_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin n\theta f(\theta) d\theta, \\
 b_0 &= \frac{1}{\sqrt{2\pi}} f_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) d\theta \equiv \frac{1}{2} A_0.
 \end{aligned} \tag{8.9}$$

So, for a function  $f(\theta)$  defined in  $(-\pi, \pi)$ , the Fourier trigonometric series is as in Equation (8.2) with the coefficients given by Equations (8.8) and (8.9). For a function  $F(x)$ , defined on  $(a, b)$ , the trigonometric series becomes

$$F(x) = \frac{1}{2} A_0 + \sum_{n=1}^{\infty} \left( A_n \cos \frac{2n\pi x}{L} + B_n \sin \frac{2n\pi x}{L} \right), \tag{8.10}$$

where

$$\begin{aligned} A_n &= \frac{2}{L} \int_a^b \cos\left(\frac{2n\pi x}{L}\right) F(x) dx, \\ B_n &= \frac{2}{L} \int_a^b \sin\left(\frac{2n\pi x}{L}\right) F(x) dx. \end{aligned} \quad (8.11)$$

A convenient rule to remember is that for even (odd) functions—which are necessarily defined on a symmetric interval around the origin—only cosine (sine) terms appear in the Fourier expansion.

**8.1.5. Example.** An alternating current is turned into a direct current by starting with a signal of the form  $V(t) \propto |\sin \omega t|$ , i.e., a harmonic function that is never negative, as shown in Figure 8.3(a). Then by proper electronics, one smooths out the “bumps” so that the output signal is very nearly a direct voltage. Let us Fourier-analyze the above signal. Since  $V(t)$  is even for  $-\pi < \omega t < \pi$ , we expect only cosine terms to be present. If for the time being we use  $\theta$  instead of  $\omega t$ , we can write  $|\sin \theta| = \frac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n \cos n\theta$ , where

$$\begin{aligned} A_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} |\sin \theta| \cos n\theta d\theta = \frac{2}{\pi} \int_0^{\pi} \sin \theta \cos n\theta d\theta \\ &= \frac{2}{\pi} \int_0^{\pi} \frac{1}{2} [\sin(n+1)\theta - \sin(n-1)\theta] d\theta = -\frac{2}{n^2-1} \left[ \frac{(-1)^n + 1}{\pi} \right] \\ &= \begin{cases} -\frac{4}{\pi} \left( \frac{1}{n^2-1} \right) & \text{for } n \text{ even and } n \neq 0, \\ 0 & \text{for } n \text{ odd,} \end{cases} \end{aligned}$$

and  $A_0 = (1/\pi) \int_{-\pi}^{\pi} |\sin \theta| d\theta = 4/\pi$ . The expansion then yields

$$|\sin \omega t| = \frac{2}{\pi} - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos 2k\omega t}{4k^2-1},$$

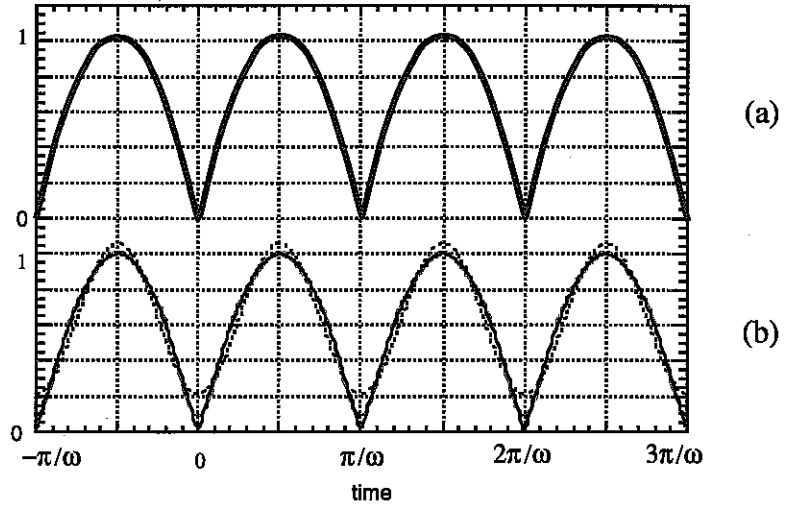
where in the sum we substituted  $2k$  for  $n$ , and  $\omega t$  for  $\theta$ . Figure 8.3(b) shows the graph of the series above when only the first few terms are kept. ■

It is useful to have a representation of the Dirac delta function in terms of the present orthonormal basis of Fourier expansion. First we note that we can represent the delta function in terms of a series in *any* set of orthonormal functions (see Problem 8.23):

$$\delta(x-x') = \sum_n f_n(x) f_n^*(x') w(x). \quad (8.12)$$

Next we use the basis of the Fourier expansion for which  $w(x) = 1$ . We then obtain

$$\delta(x-x') = \sum_{n=-\infty}^{\infty} \frac{e^{2\pi i n x/L}}{\sqrt{L}} \frac{e^{-2\pi i n x'/L}}{\sqrt{L}} = \frac{1}{L} \sum_{n=-\infty}^{\infty} e^{2\pi i n (x-x')/L}.$$



**Figure 8.3** (a) The periodic “monopolar” sine potential. (b) Various approximations to the Fourier series of the “monopolar” sine potential. The dashed plot is that of the first term of the series, the thick grey plot keeps 3 terms, and the solid plot 15 terms.

### 8.1.1 The Gibbs Phenomenon

The plot of the Fourier series expansions in Figures 8.1(b) and 8.2(b) exhibit a feature that is common to all such expansions: At the discontinuity of the periodic function, the truncated Fourier series overestimates the actual function. This is called the **Gibbs phenomenon**, and is the subject of this subsection.

Gibbs phenomenon

Let us approximate the infinite series with a finite sum. Then

$$\begin{aligned} f_N(\theta) &= \frac{1}{\sqrt{2\pi}} \sum_{n=-N}^N f_n e^{in\theta} = \frac{1}{\sqrt{2\pi}} \sum_{n=-N}^N e^{in\theta} \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{-in\theta'} f(\theta') d\theta' \\ &= \frac{1}{2\pi} \int_0^{2\pi} d\theta' f(\theta') \sum_{n=-N}^N e^{in(\theta-\theta')}, \end{aligned}$$

where we substituted Equation (8.5) in the sum and, without loss of generality, changed the interval of integration from  $(-\pi, \pi)$  to  $(0, 2\pi)$ . Problem 8.2 shows that the sum in the last equation is

$$\sum_{n=-N}^N e^{in(\theta-\theta')} = \frac{\sin[(N + \frac{1}{2})(\theta - \theta')]}{\sin[\frac{1}{2}(\theta - \theta')]}.$$

It follows that

$$f_N(\theta) = \frac{1}{2\pi} \int_0^{2\pi} d\theta' f(\theta') \frac{\sin[(N + \frac{1}{2})(\theta - \theta')]}{\sin[\frac{1}{2}(\theta - \theta')]}$$

$$= \frac{1}{2\pi} \int_{-\theta}^{2\pi-\theta} d\phi f(\phi + \theta) \underbrace{\frac{\sin[(N + \frac{1}{2})\phi]}{\sin(\frac{1}{2}\phi)}}_{\equiv S(\phi)} \equiv \frac{1}{2\pi} \int_{-\theta}^{2\pi-\theta} d\phi f(\phi + \theta) S(\phi). \tag{8.13}$$

We want to investigate the behavior of  $f_N$  at a discontinuity of  $f$ . By translating the limits of integration if necessary, we can assume that the discontinuity of  $f$  occurs at a point  $\alpha$  such that  $0 \neq \alpha \neq 2\pi$ . Let us denote the jump at this discontinuity for the function itself by  $\Delta f$ , and for its finite Fourier sum by  $\Delta f_N$ :

$$\Delta f \equiv f(\alpha + \epsilon) - f(\alpha - \epsilon), \quad \Delta f_N \equiv f_N(\alpha + \epsilon) - f_N(\alpha - \epsilon).$$

Then, we have

$$\begin{aligned} \Delta f_N &= \frac{1}{2\pi} \int_{-\alpha-\epsilon}^{2\pi-\alpha-\epsilon} d\phi f(\phi + \alpha + \epsilon) S(\phi) - \frac{1}{2\pi} \int_{-\alpha+\epsilon}^{2\pi-\alpha+\epsilon} d\phi f(\phi + \alpha - \epsilon) S(\phi) \\ &= \frac{1}{2\pi} \left\{ \int_{-\alpha-\epsilon}^{-\alpha+\epsilon} d\phi f(\phi + \alpha + \epsilon) S(\phi) + \int_{-\alpha+\epsilon}^{2\pi-\alpha-\epsilon} d\phi f(\phi + \alpha + \epsilon) S(\phi) \right\} \\ &\quad - \frac{1}{2\pi} \left\{ \int_{-\alpha+\epsilon}^{2\pi-\alpha-\epsilon} d\phi f(\phi + \alpha - \epsilon) S(\phi) + \int_{2\pi-\alpha-\epsilon}^{2\pi-\alpha+\epsilon} d\phi f(\phi + \alpha - \epsilon) S(\phi) \right\} \\ &= \frac{1}{2\pi} \left\{ \int_{-\alpha-\epsilon}^{-\alpha+\epsilon} d\phi f(\phi + \alpha + \epsilon) S(\phi) - \int_{2\pi-\alpha-\epsilon}^{2\pi-\alpha+\epsilon} d\phi f(\phi + \alpha - \epsilon) S(\phi) \right\} \\ &\quad + \frac{1}{2\pi} \int_{-\alpha+\epsilon}^{2\pi-\alpha-\epsilon} d\phi [f(\phi + \alpha + \epsilon) - f(\phi + \alpha - \epsilon)] S(\phi) \end{aligned}$$

The first two integrals give zero because of the small ranges of integration and the continuity of the integrands in those intervals. The integrand of the third integral is almost zero for all values of the range of integration except when  $\phi \approx 0$ . Hence, we can confine the integration to the small interval  $(-\delta, +\delta)$  for which the difference in the square brackets is simply  $\Delta f$ . It now follows that

maximum overshoot  
in Gibbs  
phenomenon  
calculated

$$\Delta f_N(\delta) \approx \frac{\Delta f}{2\pi} \int_{-\delta}^{\delta} \frac{\sin[(N + \frac{1}{2})\phi]}{\sin(\frac{1}{2}\phi)} d\phi \approx \frac{\Delta f}{\pi} \int_0^{\delta} \frac{\sin[(N + \frac{1}{2})\phi]}{\frac{1}{2}\phi} d\phi,$$

where we have emphasized the dependence of  $f_N$  on  $\delta$  and approximated the sine in the denominator by its argument, a good approximation due to the smallness of  $\phi$ . The reader may find the plot of the integrand in Figure 6.2, where it is shown clearly that the major contribution to the integral comes from the interval  $[0, \pi/(N + \frac{1}{2})]$ , where  $\pi/(N + \frac{1}{2})$  is the first zero of the integrand. Furthermore, it is clear that if the upper limit is larger than  $\pi/(N + \frac{1}{2})$ , the result of the integral will decrease, because in each interval of length  $2\pi$ , the area below the horizontal axis is larger than that above. Therefore, if we are interested in the *maximum* overshoot



of the finite sum, we must set the upper limit equal to  $\pi/(N + \frac{1}{2})$ . It follows firstly that the maximum overshoot of the finite sum occurs at  $\pi/(N + \frac{1}{2}) \approx \pi/N$  to the right of the discontinuity. Secondly, the amount of the maximum overshoot is

$$\begin{aligned}
 (\Delta f_N)_{\max} &\approx \frac{2\Delta f}{\pi} \int_0^{\pi/(N+\frac{1}{2})} \frac{\overbrace{\sin[(N+\frac{1}{2})\phi]}^{\equiv x}}{\phi} d\phi \\
 &= \frac{2}{\pi} \Delta f \int_0^{\pi} \frac{\sin x}{x} dx \approx 1.179\Delta f.
 \end{aligned} \tag{8.14}$$

Thus

**8.1.6. Box.** (Gibbs phenomenon) *The finite (large- $N$ ) sum approximation of the discontinuous function overshoots the function itself at a discontinuity by about 18 percent.*

## 8.1.2 Fourier Series in Higher Dimensions

It is instructive to generalize the Fourier series to more than one dimension. This generalization is especially useful in crystallography and solid-state physics, which deal with three-dimensional periodic structures. To generalize to  $N$  dimensions, we first consider a special case in which an  $N$ -dimensional periodic function is a product of  $N$  one-dimensional periodic functions. That is, we take the  $N$  functions

$$f^{(j)}(x) = \frac{1}{\sqrt{L_j}} \sum_{k=-\infty}^{\infty} f_k^{(j)} e^{2i\pi kx/L_j}, \quad j = 1, 2, \dots, N,$$

and multiply them on both sides to obtain

$$F(\mathbf{r}) = f^{(1)}(x_1) f^{(2)}(x_2) \cdots f^{(N)}(x_N) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} F_{\mathbf{k}} e^{i\mathbf{g}_{\mathbf{k}} \cdot \mathbf{r}}, \tag{8.15}$$

where we have used the following new notations:

$$\begin{aligned}
 F(\mathbf{r}) &\equiv f^{(1)}(x_1) f^{(2)}(x_2) \cdots f^{(N)}(x_N), & V &= L_1 L_2 \cdots L_N, \\
 \mathbf{k} &\equiv (k_1, k_2, \dots, k_N), & F_{\mathbf{k}} &\equiv f_{k_1} \cdots f_{k_N}, \\
 \mathbf{g}_{\mathbf{k}} &= 2\pi (k_1/L_1, \dots, k_N/L_N), & \mathbf{r} &= (x_1, x_2, \dots, x_N).
 \end{aligned}$$

We take Equation (8.15) as the definition of the Fourier series for *any* periodic function of  $N$  variables (not just the product of  $N$  functions of a single variable). However, application of (8.15) requires some clarification. In one dimension, the

Wigner-Seitz cell

shape of the smallest region of periodicity is unique. It is simply a line segment of length  $L$ , for example. In two and more dimensions, however, such regions may have a variety of shapes. For instance, in two dimensions, they can be rectangles, pentagons, hexagons, and so forth. Thus, we let  $V$  in Equation (8.15) stand for a primitive cell of the  $N$ -dimensional lattice. This cell is important in solid-state physics, and (in three dimensions) is called the **Wigner-Seitz cell**.

It is customary to absorb the factor  $1/\sqrt{V}$  into  $F_{\mathbf{k}}$ , and write

$$F(\mathbf{r}) = \sum_{\mathbf{k}} F_{\mathbf{k}} e^{i\mathbf{g}_{\mathbf{k}} \cdot \mathbf{r}} \Leftrightarrow F_{\mathbf{k}} = \frac{1}{V} \int_V F(\mathbf{r}) e^{-i\mathbf{g}_{\mathbf{k}} \cdot \mathbf{r}} d^N x, \quad (8.16)$$

where the integral is over a single Wigner-Seitz cell.

Recall that  $F(\mathbf{r})$  is a periodic function of  $\mathbf{r}$ . This means that when  $\mathbf{r}$  is changed by  $\mathbf{R}$ , where  $\mathbf{R}$  is a vector describing the boundaries of a cell, then we should get the same function:  $F(\mathbf{r} + \mathbf{R}) = F(\mathbf{r})$ . When substituted in (8.16), this yields  $F(\mathbf{r} + \mathbf{R}) = \sum_{\mathbf{k}} F_{\mathbf{k}} e^{i\mathbf{g}_{\mathbf{k}} \cdot (\mathbf{r} + \mathbf{R})} = \sum_{\mathbf{k}} e^{i\mathbf{g}_{\mathbf{k}} \cdot \mathbf{R}} F_{\mathbf{k}} e^{i\mathbf{g}_{\mathbf{k}} \cdot \mathbf{r}}$ , which is equal to  $F(\mathbf{r})$  if

$$e^{i\mathbf{g}_{\mathbf{k}} \cdot \mathbf{R}} = 1. \quad (8.17)$$

reciprocal lattice vectors

In three dimensions  $\mathbf{R} = m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2 + m_3 \mathbf{a}_3$ , where  $m_1, m_2$ , and  $m_3$  are integers and  $\mathbf{a}_1, \mathbf{a}_2$ , and  $\mathbf{a}_3$  are *crystal axes*, which are not generally orthogonal. On the other hand,  $\mathbf{g}_{\mathbf{k}} = n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2 + n_3 \mathbf{b}_3$ , where  $n_1, n_2$ , and  $n_3$  are integers, and  $\mathbf{b}_1, \mathbf{b}_2$ , and  $\mathbf{b}_3$  are the **reciprocal lattice vectors** defined by

$$\mathbf{b}_1 = \frac{2\pi(\mathbf{a}_2 \times \mathbf{a}_3)}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}, \quad \mathbf{b}_2 = \frac{2\pi(\mathbf{a}_3 \times \mathbf{a}_1)}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}, \quad \mathbf{b}_3 = \frac{2\pi(\mathbf{a}_1 \times \mathbf{a}_2)}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}.$$

The reader may verify that  $\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi \delta_{ij}$ . Thus

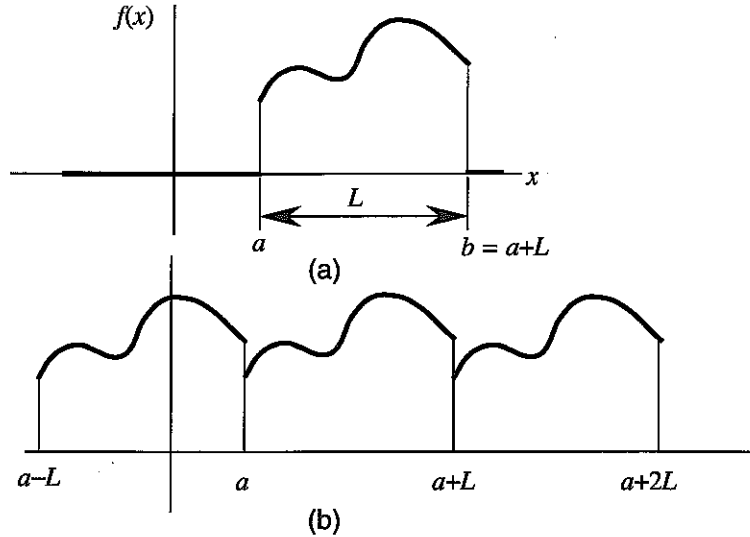
$$\begin{aligned} \mathbf{g}_{\mathbf{k}} \cdot \mathbf{R} &= \left( \sum_{i=1}^3 n_i \mathbf{b}_i \right) \cdot \left( \sum_{j=1}^3 m_j \mathbf{a}_j \right) = \sum_{i,j} n_i m_j \mathbf{b}_i \cdot \mathbf{a}_j \\ &= 2\pi \sum_{j=1}^3 m_j n_j = 2\pi (\text{integer}), \end{aligned}$$

and Equation (8.17) is satisfied.

## 8.2 The Fourier Transform

The Fourier series representation of  $F(x)$  is valid for the entire real line as long as  $F(x)$  is periodic. However, most functions encountered in physical applications are defined in some interval  $(a, b)$  without repetition beyond that interval. It would be useful if we could also expand such functions in some form of Fourier "series."

One way to do this is to start with the periodic series and then let the period go to infinity while extending the domain of the definition of the function. As a



**Figure 8.4** (a) The function we want to represent. (b) The Fourier series representation of the function.

specific case, suppose we are interested in representing a function  $f(x)$  that is defined only for the interval  $(a, b)$  and is assigned the value zero everywhere else [see Figure 8.4(a)]. To begin with, we might try the Fourier series representation, but this will produce a repetition of our function. This situation is depicted in Figure 8.4(b).

Next we may try a function  $g_\Lambda(x)$  defined in the interval  $(a - \Lambda/2, b + \Lambda/2)$ , where  $\Lambda$  is an arbitrary positive number:

$$g_\Lambda(x) = \begin{cases} 0 & \text{if } a - \Lambda/2 < x < a, \\ f(x) & \text{if } a < x < b, \\ 0 & \text{if } b < x < b + \Lambda/2. \end{cases}$$

This function, which is depicted in Figure 8.5, has the Fourier series representation

$$g_\Lambda(x) = \frac{1}{\sqrt{L + \Lambda}} \sum_{n=-\infty}^{\infty} g_{\Lambda,n} e^{2i\pi nx/(L+\Lambda)}, \quad (8.18)$$

where

$$g_{\Lambda,n} = \frac{1}{\sqrt{L + \Lambda}} \int_{a-\Lambda/2}^{b+\Lambda/2} e^{-2i\pi nx/(L+\Lambda)} g_\Lambda(x) dx. \quad (8.19)$$

We have managed to separate various copies of the original periodic function by  $\Lambda$ . It should be clear that if  $\Lambda \rightarrow \infty$ , we can completely isolate the function

and stop the repetition. Let us investigate the behavior of Equations (8.18) and (8.19) as  $\Lambda$  grows without bound. First, we notice that the quantity  $k_n$  defined by  $k_n \equiv 2n\pi/(L + \Lambda)$  and appearing in the exponent becomes almost continuous. In other words, as  $n$  changes by one unit,  $k_n$  changes only slightly. This suggests that the terms in the sum in Equation (8.18) can be lumped together in  $j$  intervals of width  $\Delta n_j$ , giving

$$g_\Lambda(x) \approx \sum_{j=-\infty}^{\infty} \frac{g_\Lambda(k_j)}{\sqrt{L + \Lambda}} e^{ik_j x} \Delta n_j,$$

where  $k_j \equiv 2n_j\pi/(L + \Lambda)$ , and  $g_\Lambda(k_j) \equiv g_{\Lambda,n_j}$ . Substituting  $\Delta n_j = [(L + \Lambda)/2\pi] \Delta k_j$  in the above sum, we obtain

$$g_\Lambda(x) \approx \sum_{j=-\infty}^{\infty} \frac{g_\Lambda(k_j)}{\sqrt{L + \Lambda}} e^{ik_j x} \frac{L + \Lambda}{2\pi} \Delta k_j = \frac{1}{\sqrt{2\pi}} \sum_{j=-\infty}^{\infty} \tilde{g}_\Lambda(k_j) e^{ik_j x} \Delta k_j,$$

where we introduced  $\tilde{g}_\Lambda(k_j)$  defined by  $\tilde{g}_\Lambda(k_j) \equiv \sqrt{(L + \Lambda)/2\pi} g_\Lambda(k_j)$ . It is now clear that the preceding sum approaches an integral in the limit that  $\Lambda \rightarrow \infty$ . In the same limit,  $g_\Lambda(x) \rightarrow f(x)$ , and we have

Fourier integral transforms

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dk, \tag{8.20}$$

where

$$\begin{aligned} \tilde{f}(k) &\equiv \lim_{\Lambda \rightarrow \infty} \tilde{g}_\Lambda(k_j) = \lim_{\Lambda \rightarrow \infty} \sqrt{\frac{L + \Lambda}{2\pi}} g_\Lambda(k_j) \\ &= \lim_{\Lambda \rightarrow \infty} \sqrt{\frac{L + \Lambda}{2\pi}} \frac{1}{\sqrt{L + \Lambda}} \int_{a-\Lambda/2}^{b+\Lambda/2} e^{-ik_j x} g_\Lambda(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx. \end{aligned} \tag{8.21}$$

Equations (8.20) and (8.21) are called the **Fourier integral transforms** of  $\tilde{f}(k)$  and  $f(x)$ , respectively.

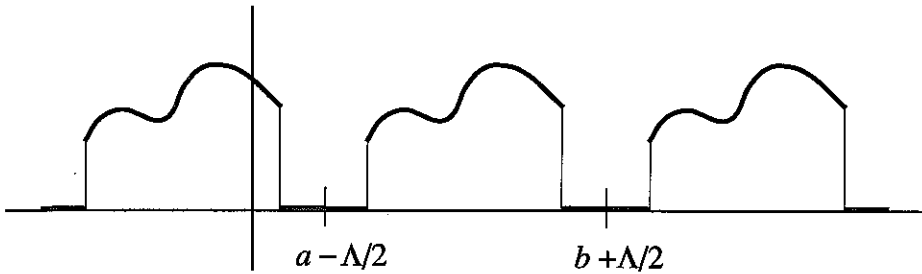
**8.2.1. Example.** Let us evaluate the Fourier transform of the function defined by

$$f(x) = \begin{cases} b & \text{if } |x| < a, \\ 0 & \text{if } |x| > a \end{cases}$$

(see Figure 8.6). From (8.21) we have

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx = \frac{b}{\sqrt{2\pi}} \int_{-a}^a e^{-ikx} dx = \frac{2ab}{\sqrt{2\pi}} \left( \frac{\sin ka}{ka} \right),$$

which is the function encountered (and depicted) in Example 6.1.2.



**Figure 8.5** By introducing the parameter  $\Delta$ , we have managed to separate the copies of the function.

Let us discuss this result in detail. First, note that if  $a \rightarrow \infty$ , the function  $f(x)$  becomes a constant function over the entire real line, and we get

$$\tilde{f}(k) = \frac{2b}{\sqrt{2\pi}} \lim_{a \rightarrow \infty} \frac{\sin ka}{k} = \frac{2b}{\sqrt{2\pi}} \pi \delta(k)$$

by the result of Example 6.1.2. This is the Fourier transform of an everywhere-constant function (see Problem 8.12). Next, let  $b \rightarrow \infty$  and  $a \rightarrow 0$  in such a way that  $2ab$ , which is the area under  $f(x)$ , is 1. Then  $f(x)$  will approach the delta function, and  $\tilde{f}(k)$  becomes

$$\tilde{f}(k) = \lim_{\substack{b \rightarrow \infty \\ a \rightarrow 0}} \frac{2ab}{\sqrt{2\pi}} \frac{\sin ka}{ka} = \frac{1}{\sqrt{2\pi}} \lim_{a \rightarrow 0} \frac{\sin ka}{ka} = \frac{1}{\sqrt{2\pi}}$$

So the Fourier transform of the delta function is the constant  $1/\sqrt{2\pi}$ .

Finally, we note that the width of  $f(x)$  is  $\Delta x = 2a$ , and the width of  $\tilde{f}(k)$  is roughly the distance, on the  $k$ -axis, between its first two roots,  $k_+$  and  $k_-$ , on either side of  $k = 0$ :  $\Delta k = k_+ - k_- = 2\pi/a$ . Thus increasing the width of  $f(x)$  results in a decrease in the width of  $\tilde{f}(k)$ . In other words, when the function is wide, its Fourier transform is narrow. In the limit of infinite width (a constant function), we get infinite sharpness (the delta function). The last two statements are very general. In fact, it can be shown that  $\Delta x \Delta k \geq 1$  for any function  $f(x)$ . When both sides of this inequality are multiplied by the (reduced) Planck constant  $\hbar \equiv h/(2\pi)$ , the result is the celebrated **Heisenberg uncertainty relation**:<sup>3</sup>

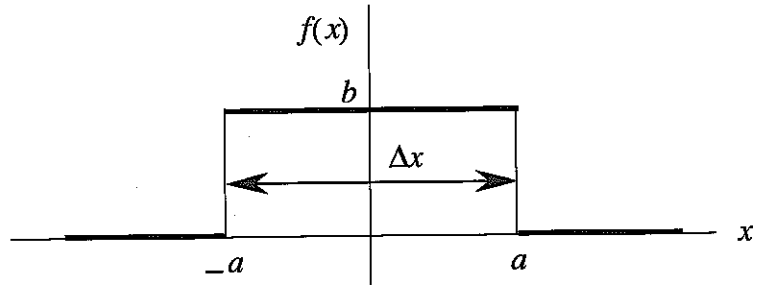
Heisenberg  
uncertainty relation

$$\Delta x \Delta p \geq \hbar,$$

where  $p = \hbar k$  is the momentum of the particle. Having obtained the transform of  $f(x)$ , we can write

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{2b}{\sqrt{2\pi}} \frac{\sin ka}{k} e^{ikx} dk = \frac{b}{\pi} \int_{-\infty}^{\infty} \frac{\sin ka}{k} e^{ikx} dk. \quad \blacksquare$$

<sup>3</sup>In the context of the uncertainty relation, the width of the function—the so-called wave packet—measures the uncertainty in the position  $x$  of a quantum mechanical particle. Similarly, the width of the Fourier transform measures the uncertainty in  $k$ , which is related to momentum  $p$  via  $p = \hbar k$ .



**Figure 8.6** The square “bump” function.

**8.2.2. Example.** Let us evaluate the Fourier transform of a Gaussian  $g(x) = ae^{-bx^2}$  with  $a, b > 0$ :

$$\tilde{g}(k) = \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-b(x^2+ikx/b)} dx = \frac{ae^{-k^2/4b}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-b(x+ik/2b)^2} dx.$$

To evaluate this integral rigorously, we would have to use techniques developed in complex analysis, which are not introduced until Chapter 10 (see Example 10.3.8). However, we can ignore the fact that the exponent is complex, substitute  $y = x + ik/(2b)$ , and write

$$\int_{-\infty}^{\infty} e^{-b[x+ik/(2b)]^2} dx = \int_{-\infty}^{\infty} e^{-by^2} dy = \sqrt{\frac{\pi}{b}}.$$

Thus, we have  $\tilde{g}(k) = \frac{a}{\sqrt{2b}} e^{-k^2/(4b)}$ , which is also a Gaussian.

We note again that the width of  $g(x)$ , which is proportional to  $1/\sqrt{b}$ , is in inverse relation to the width of  $\tilde{g}(k)$ , which is proportional to  $\sqrt{b}$ . We thus have  $\Delta x \Delta k \sim 1$ . ■

Equations (8.20) and (8.21) are reciprocals of one another. However, it is not obvious that they are consistent. In other words, if we substitute (8.20) in the RHS of (8.21), do we get an identity? Let's try this:

$$\begin{aligned} \tilde{f}(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-ikx} \left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k') e^{ik'x} dk' \right] \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} \tilde{f}(k') e^{i(k'-k)x} dk'. \end{aligned}$$

We now change the order of the two integrations:

$$\tilde{f}(k) = \int_{-\infty}^{\infty} dk' \tilde{f}(k') \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{i(k'-k)x} \right].$$

But the expression in the square brackets is the delta function (see Example 6.1.2). Thus, we have  $\tilde{f}(k) = \int_{-\infty}^{\infty} dk' \tilde{f}(k') \delta(k' - k)$ , which is an identity.

As in the case of Fourier series, Equations (8.20) and (8.21) are valid even if  $f$  and  $\tilde{f}$  are piecewise continuous. In that case the Fourier transforms are written as

$$\begin{aligned}\frac{1}{2}[f(x+0) + f(x-0)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dk, \\ \frac{1}{2}[\tilde{f}(k+0) + \tilde{f}(k-0)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx,\end{aligned}\tag{8.22}$$

where each zero on the LHS is an  $\epsilon$  that has gone to its limit.

It is useful to generalize Fourier transform equations to more than one dimension. The generalization is straightforward:

$$\begin{aligned}f(\mathbf{r}) &= \frac{1}{(2\pi)^{n/2}} \int d^n k e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{f}(\mathbf{k}), \\ \tilde{f}(\mathbf{k}) &= \frac{1}{(2\pi)^{n/2}} \int d^n x f(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}.\end{aligned}\tag{8.23}$$

Let us now use the abstract notation of Chapter 6 to get more insight into the preceding results. In the language of Chapter 6, Equation (8.20) can be written as

$$\langle x | f \rangle = \int_{-\infty}^{\infty} \langle k | \tilde{f} \rangle \langle x | k \rangle dk = \langle x | \left( \int_{-\infty}^{\infty} |k\rangle \langle k| dk \right) | \tilde{f} \rangle,\tag{8.24}$$

where we have defined

$$\langle x | k \rangle = \frac{1}{\sqrt{2\pi}} e^{ikx}.\tag{8.25}$$

Equation (8.24) suggests the identification  $|\tilde{f}\rangle \equiv |f\rangle$  as well as the identity

$$\mathbf{1} = \int_{-\infty}^{\infty} |k\rangle \langle k| dk,\tag{8.26}$$

which is the same as (6.1). Equation (6.3) yields

$$\langle k | k' \rangle = \delta(k - k'),\tag{8.27}$$

which upon the insertion of a unit operator gives an integral representation of the delta function:

$$\begin{aligned}\delta(k - k') &= \langle k | \mathbf{1} | k' \rangle = \langle k | \left( \int_{-\infty}^{\infty} |x\rangle \langle x| dx \right) | k' \rangle \\ &= \int_{-\infty}^{\infty} \langle k | x \rangle \langle x | k' \rangle dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{i(k'-k)x}.\end{aligned}$$

Obviously, we can also write  $\delta(x - x') = [1/(2\pi)] \int_{-\infty}^{\infty} dk e^{i(x-x')k}$ .

If more than one dimension is involved, we use

$$\begin{aligned}\delta(\mathbf{k} - \mathbf{k}') &= \frac{1}{(2\pi)^n} \int d^n x e^{i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}}, \\ \delta(\mathbf{r} - \mathbf{r}') &= \frac{1}{(2\pi)^n} \int d^n k e^{i(\mathbf{r} - \mathbf{r}') \cdot \mathbf{k}},\end{aligned}\quad (8.28)$$

with the inner product relations

$$\langle \mathbf{r} | \mathbf{k} \rangle = \frac{1}{(2\pi)^{n/2}} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad \langle \mathbf{k} | \mathbf{r} \rangle = \frac{1}{(2\pi)^{n/2}} e^{-i\mathbf{k} \cdot \mathbf{r}}. \quad (8.29)$$

Equations (8.28) and (8.29) and the identification  $|\tilde{f}\rangle \equiv |f\rangle$  exhibit a striking resemblance between  $|\mathbf{r}\rangle$  and  $|\mathbf{k}\rangle$ . In fact, any given abstract vector  $|f\rangle$  can be expressed either in terms of its  $r$  representation,  $\langle \mathbf{r} | f \rangle = f(\mathbf{r})$ , or in terms of its  $k$  representation,  $\langle \mathbf{k} | f \rangle \equiv \tilde{f}(\mathbf{k})$ . These two representations are completely equivalent, and there is a one-to-one correspondence between the two, given by Equation (8.23). The representation that is used in practice is dictated by the physical application. In quantum mechanics, for instance, most of the time the  $r$  representation, corresponding to the position, is used, because then the operator equations turn into differential equations that are normally linear and easier to solve than the corresponding equations in the  $k$  representation, which is related to the momentum.

**8.2.3. Example.** In this example we evaluate the Fourier transform of the Coulomb potential  $V(r)$  of a point charge  $q$ :  $V(r) = q/r$ . The Fourier transform is important in scattering experiments with atoms, molecules, and solids. As we shall see in the following, the Fourier transform of  $V(r)$  is not defined. However, if we work with the **Yukawa potential**,

$$V_\alpha(r) = \frac{q e^{-\alpha r}}{r}, \quad \alpha > 0,$$

the Fourier transform will be well-defined, and we can take the limit  $\alpha \rightarrow 0$  to recover the Coulomb potential. Thus, we seek the Fourier transform of  $V_\alpha(r)$ .

We are working in three dimensions and therefore may write

$$\tilde{V}_\alpha(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \iiint d^3 x e^{-i\mathbf{k} \cdot \mathbf{r}} \frac{q e^{-\alpha r}}{r}.$$

It is clear from the presence of  $r$  that spherical coordinates are appropriate. We are free to pick any direction as the  $z$ -axis. A simplifying choice in this case is the direction of  $\mathbf{k}$ . So, we let  $\mathbf{k} = |\mathbf{k}| \hat{\mathbf{e}}_z = k \hat{\mathbf{e}}_z$ , or  $\mathbf{k} \cdot \mathbf{r} = kr \cos \theta$ , where  $\theta$  is the polar angle in spherical coordinates. Now we have

$$\tilde{V}_\alpha(\mathbf{k}) = \frac{q}{(2\pi)^{3/2}} \int_0^\infty r^2 dr \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\varphi e^{-ikr \cos \theta} \frac{e^{-\alpha r}}{r}.$$

The  $\varphi$  integration is trivial and gives  $2\pi$ . The  $\theta$  integration is done next:

$$\int_0^\pi \sin \theta e^{-ikr \cos \theta} d\theta = \int_{-1}^1 e^{-ikru} du = \frac{1}{ikr} (e^{ikr} - e^{-ikr}).$$



We thus have

$$\begin{aligned}\tilde{V}_\alpha(\mathbf{k}) &= \frac{q(2\pi)}{(2\pi)^{3/2}} \int_0^\infty dr r^2 \frac{e^{-\alpha r}}{r} \frac{1}{ikr} (e^{ikr} - e^{-ikr}) \\ &= \frac{q}{(2\pi)^{1/2}} \frac{1}{ik} \int_0^\infty dr [e^{(-\alpha+ik)r} - e^{-(\alpha+ik)r}] \\ &= \frac{q}{(2\pi)^{1/2}} \frac{1}{ik} \left( \left. \frac{e^{(-\alpha+ik)r}}{-\alpha+ik} \right|_0^\infty + \left. \frac{e^{-(\alpha+ik)r}}{\alpha+ik} \right|_0^\infty \right).\end{aligned}$$

Note how the factor  $e^{-\alpha r}$  has tamed the divergent behavior of the exponential at  $r \rightarrow \infty$ . This was the reason for introducing it in the first place. Simplifying the last expression yields  $\tilde{V}_\alpha(\mathbf{k}) = (2q/\sqrt{2\pi})(k^2 + \alpha^2)^{-1}$ . The parameter  $\alpha$  is a measure of the range of the potential. It is clear that the larger  $\alpha$  is, the smaller the range. In fact, it was in response to the short range of nuclear forces that Yukawa introduced  $\alpha$ . For electromagnetism, where the range is infinite,  $\alpha$  becomes zero and  $V_\alpha(r)$  reduces to  $V(r)$ . Thus, the Fourier transform of the Coulomb potential is

$$\tilde{V}_{\text{Coul}}(\mathbf{k}) = \frac{2q}{\sqrt{2\pi}} \frac{1}{k^2}.$$

If a charge *distribution* is involved, the Fourier transform will be different. ■

**8.2.4. Example.** The example above deals with the electrostatic potential of a point charge. Let us now consider the case where the charge is distributed over a finite volume. Then the potential is

$$V(\mathbf{r}) = \iiint \frac{q\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d^3x' \equiv q \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d^3x',$$

where  $q\rho(\mathbf{r}')$  is the charge density at  $\mathbf{r}'$ , and we have used a single integral because  $d^3x'$  already indicates the number of integrations to be performed. Note that we have normalized  $\rho(\mathbf{r}')$  so that its integral over the volume is 1. Figure 8.7 shows the geometry of the situation.

Making a change of variables,  $\mathbf{R} \equiv \mathbf{r}' - \mathbf{r}$ , or  $\mathbf{r}' = \mathbf{R} + \mathbf{r}$ , and  $d^3x' = d^3X$ , with  $\mathbf{R} \equiv (X, Y, Z)$ , we get

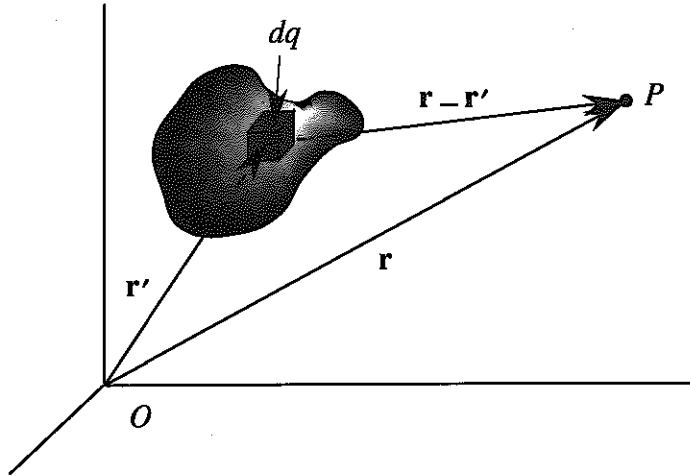
$$\tilde{V}(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \int d^3x e^{-ik \cdot \mathbf{r}} q \int \frac{\rho(\mathbf{R} + \mathbf{r})}{R} d^3X. \quad (8.30)$$

To evaluate Equation (8.30), we substitute for  $\rho(\mathbf{R} + \mathbf{r})$  in terms of its Fourier transform,

$$\rho(\mathbf{R} + \mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int d^3k' \tilde{\rho}(\mathbf{k}') e^{i\mathbf{k}' \cdot (\mathbf{R} + \mathbf{r})}. \quad (8.31)$$

Combining (8.30) and (8.31), we obtain

$$\begin{aligned}\tilde{V}(\mathbf{k}) &= \frac{q}{(2\pi)^3} \int d^3x d^3X d^3k' \frac{e^{i\mathbf{k}' \cdot \mathbf{R}}}{R} \tilde{\rho}(\mathbf{k}') e^{i\mathbf{r} \cdot (\mathbf{k}' - \mathbf{k})} \\ &= q \int d^3X d^3k' \frac{e^{i\mathbf{k}' \cdot \mathbf{R}}}{R} \tilde{\rho}(\mathbf{k}') \underbrace{\left( \frac{1}{(2\pi)^3} \int d^3x e^{i\mathbf{r} \cdot (\mathbf{k}' - \mathbf{k})} \right)}_{\delta(\mathbf{k}' - \mathbf{k})} \\ &= q \tilde{\rho}(\mathbf{k}) \int d^3X \frac{e^{i\mathbf{k} \cdot \mathbf{R}}}{R}.\end{aligned} \quad (8.32)$$



**Figure 8.7** The Fourier transform of the potential of a continuous charge distribution at  $P$  is calculated using this geometry.

What is nice about this result is that the contribution of the charge distribution,  $\bar{\rho}(\mathbf{k})$ , has been completely factored out. The integral, aside from a constant and a change in the sign of  $\mathbf{k}$ , is simply the Fourier transform of the Coulomb potential of a point charge obtained in the previous example. We can therefore write Equation (8.32) as

$$\tilde{V}(\mathbf{k}) = (2\pi)^{3/2} \bar{\rho}(\mathbf{k}) \tilde{V}_{\text{Coul}}(-\mathbf{k}) = \frac{4\pi q \bar{\rho}(\mathbf{k})}{|\mathbf{k}|^2}.$$

This equation is important in analyzing the structure of atomic particles. The Fourier transform  $\tilde{V}(\mathbf{k})$  is directly measurable in scattering experiments. In a typical experiment a (charged) target is probed with a charged point particle (electron). If the analysis of the scattering data shows a deviation from  $1/k^2$  in the behavior of  $\tilde{V}(\mathbf{k})$ , then it can be concluded that the target particle has a charge distribution. More specifically, a plot of  $k^2 \tilde{V}(\mathbf{k})$  versus  $k$  gives the variation of  $\bar{\rho}(\mathbf{k})$ , the **form factor**, with  $k$ . If the resulting graph is a constant, then  $\bar{\rho}(\mathbf{k})$  is a constant, and the target is a point particle [ $\bar{\rho}(\mathbf{k})$  is a constant for point particles, where  $\bar{\rho}(\mathbf{r}') \propto \delta(\mathbf{r} - \mathbf{r}')$ ]. If there is any deviation from a constant function,  $\bar{\rho}(\mathbf{k})$  must have a dependence on  $k$ , and correspondingly, the target particle must have a charge distribution.

The above discussion, when generalized to four-dimensional relativistic space-time, was the basis for a strong argument in favor of the existence of point-like particles—quarks—inside a proton in 1968, when the results of the scattering of high-energy electrons off protons at the Stanford Linear Accelerator Center revealed deviation from a constant for the proton form factor. ■

### 8.2.1 Fourier Transforms and Derivatives

The Fourier transform is very useful for solving differential equations. This is because the derivative operator in  $\mathbf{r}$  space turns into ordinary multiplication in  $\mathbf{k}$

form factor

Fourier transform and the discovery of quarks

space. For example, if we differentiate  $f(\mathbf{r})$  in Equation (8.23) with respect to  $x_j$ , we obtain

$$\begin{aligned}\frac{\partial}{\partial x_j} f(\mathbf{r}) &= \frac{1}{(2\pi)^{n/2}} \int d^n k \frac{\partial}{\partial x_j} e^{i(k_1 x_1 + \dots + k_j x_j + \dots + k_n x_n)} \tilde{f}(\mathbf{k}) \\ &= \frac{1}{(2\pi)^{n/2}} \int d^n k (ik_j) e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{f}(\mathbf{k}).\end{aligned}$$

That is, every time we differentiate with respect to any component of  $\mathbf{r}$ , the corresponding component of  $\mathbf{k}$  “comes down.” Thus, the  $n$ -dimensional gradient is  $\nabla f(\mathbf{r}) = (2\pi)^{-n/2} \int d^n k (i\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{f}(\mathbf{k})$ , and the  $n$ -dimensional Laplacian is  $\nabla^2 f(\mathbf{r}) = (2\pi)^{-n/2} \int d^n k (-k^2) e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{f}(\mathbf{k})$ .

We shall use Fourier transforms extensively in solving differential equations later in the book. Here, we can illustrate the above points with a simple example. Consider the ordinary second-order differential equation

$$C_2 \frac{d^2 y}{dx^2} + C_1 \frac{dy}{dx} + C_0 y = f(x),$$

where  $C_0$ ,  $C_1$ , and  $C_2$  are constants. We can “solve” this equation by simply substituting the following in it:

$$\begin{aligned}y(x) &= \frac{1}{\sqrt{2\pi}} \int dk \tilde{y}(k) e^{ikx}, & \frac{dy}{dx} &= \frac{1}{\sqrt{2\pi}} \int dk \tilde{y}(k) (ik) e^{ikx}, \\ \frac{d^2 y}{dx^2} &= -\frac{1}{\sqrt{2\pi}} \int dk \tilde{y}(k) k^2 e^{ikx}, & f(x) &= \frac{1}{\sqrt{2\pi}} \int dk \tilde{f}(k) e^{ikx}.\end{aligned}$$

This gives

$$\frac{1}{\sqrt{2\pi}} \int dk \tilde{y}(k) (-C_2 k^2 + iC_1 k + C_0) e^{ikx} = \frac{1}{\sqrt{2\pi}} \int dk \tilde{f}(k) e^{ikx}.$$

Equating the coefficients of  $e^{ikx}$  on both sides, we obtain

$$\tilde{y}(k) = \frac{\tilde{f}(k)}{-C_2 k^2 + iC_1 k + C_0}.$$

If we know  $\tilde{f}(k)$  [which can be obtained from  $f(x)$ ], we can calculate  $y(x)$  by Fourier-transforming  $\tilde{y}(k)$ . The resulting integrals are not generally easy to evaluate. In some cases the methods of complex analysis may be helpful; in others numerical integration may be the last resort. However, the real power of the Fourier transform lies in the formal analysis of differential equations.

## 8.2.2 The Discrete Fourier Transform

The preceding remarks alluded to the power of the Fourier transform in solving certain differential equations. If such a solution is combined with numerical techniques, the integrals must be replaced by sums. This is particularly true if our

function is given by a table rather than a mathematical relation, a common feature of numerical analysis. So suppose that we are given a set of measurements performed in equal time intervals of  $\Delta t$ . Suppose that the overall period in which these measurements are done is  $T$ . We are seeking a Fourier transform of this finite set of data. First we write

$$\tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \approx \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{N-1} f(t_n) e^{-i\omega t_n} \Delta t,$$

or, discretizing the frequency as well and writing  $\omega_m = m\Delta\omega$ , with  $\Delta\omega$  to be determined later, we have

$$\tilde{f}(m\Delta\omega) = \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{N-1} f(n\Delta t) e^{-i(m\Delta\omega)n\Delta t} \left(\frac{T}{N}\right). \quad (8.33)$$

Since the Fourier transform is given in terms of a finite sum, let us explore the idea of writing the inverse transform also as a sum. So, multiply both sides of the above equation by  $[e^{i(m\Delta\omega)k\Delta t} / (\sqrt{2\pi})] \Delta\omega$  and sum over  $m$ :

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \sum_{m=0}^{N-1} \tilde{f}(m\Delta\omega) e^{i(m\Delta\omega)k\Delta t} \Delta\omega &= \frac{T\Delta\omega}{2\pi N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} f(n\Delta t) e^{im\Delta\omega\Delta t(k-n)} \\ &= \frac{T\Delta\omega}{2\pi N} \sum_{n=0}^{N-1} f(n\Delta t) \sum_{m=0}^{N-1} e^{im\Delta\omega\Delta t(k-n)}. \end{aligned}$$

Problem 8.2 shows that

$$\sum_{m=0}^{N-1} e^{im\Delta\omega\Delta t(k-n)} = \begin{cases} N & \text{if } k = n, \\ \frac{e^{iN\Delta\omega\Delta t(k-n)} - 1}{e^{i\Delta\omega\Delta t(k-n)} - 1} & \text{if } k \neq n. \end{cases}$$

We want the sum to vanish when  $k \neq n$ . This suggests demanding that  $N\Delta\omega\Delta t(k-n)$  be an integer multiple of  $2\pi$ . Since  $\Delta\omega$  and  $\Delta t$  are to be independent of this (arbitrary) integer (as well as  $k$  and  $n$ ), we must write

$$N\Delta\omega\Delta t(k-n) = 2\pi(k-n) \Rightarrow N\Delta\omega \frac{T}{N} = 2\pi \Rightarrow \Delta\omega = \frac{2\pi}{T}.$$

discrete Fourier  
transforms

With this choice, we have the following **discrete Fourier transforms**:

$$\begin{aligned} \tilde{f}(\omega_j) &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(t_n) e^{-i\omega_j t_n}, \\ f(t_n) &= \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \tilde{f}(\omega_j) e^{i\omega_j t_n}, \quad t_n = n\Delta t, \quad \omega_j = \frac{2\pi j}{T}, \end{aligned} \quad (8.34)$$

where we have redefined the new  $\tilde{f}$  to be  $\sqrt{2\pi N}/T$  times the old  $\tilde{f}$ .

Discrete Fourier transforms are used extensively in numerical calculation of problems in which ordinary Fourier transforms are used. For instance, if a differential equation lends itself to a solution via the Fourier transform as discussed before, then discrete Fourier transforms will give a procedure for finding the solution numerically. Similarly, the frequency analysis of signals is nicely handled by discrete Fourier transforms.

fast Fourier  
transform

It turns out that discrete Fourier analysis is very intensive computationally. Its present status as a popular tool in computational physics is due primarily to a very efficient method of calculation known as the **fast Fourier transform**. In a typical Fourier transform, one has to perform a sum of  $N$  terms for every point. Since there are  $N$  points to transform, the total computational time will be of order  $N^2$ . In the fast Fourier transform, one takes  $N$  to be even and divides the sum into two other sums, one over the even terms and one over the odd terms. Then the computation time will be of order  $2 \times (N/2)^2$ , or half the original calculation. Similarly, if  $N/2$  is even, one can further divide the odd and even sums by two and obtain a computation time of  $4 \times (N/4)^2$ , or a quarter of the original calculation. In general, if  $N = 2^k$ , then by dividing the sums consecutively, we end up with  $N$  transforms to be performed after  $k$  steps. So, the computation time will be  $kN = N \log_2 N$ . For  $N = 128$ , the computation time will be  $100 \log_2 128 = 700$  as opposed to  $128^2 \approx 16,400$ , a reduction by a factor of over 20. The fast Fourier transform is indeed fast!

### 8.2.3 The Fourier Transform of a Distribution

Although one can define the Fourier transform of a distribution in exact analogy to an ordinary function, sometimes it is convenient to define the Fourier transform of the distribution as a linear functional.

Let us ignore the distinction between the two variables  $x$  and  $k$ , and simply define the Fourier transform of a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  as

$$\tilde{f}(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{iut} dt.$$

Now we consider two functions,  $f$  and  $g$ , and note that

$$\begin{aligned} \langle f, \tilde{g} \rangle &\equiv \int_{-\infty}^{\infty} f(u) \tilde{g}(u) du = \int_{-\infty}^{\infty} f(u) \left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{-iut} dt \right] du \\ &= \int_{-\infty}^{\infty} g(t) \left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{-iut} du \right] dt \\ &= \int_{-\infty}^{\infty} g(t) \tilde{f}(t) dt = \langle \tilde{f}, g \rangle. \end{aligned}$$

The following definition is motivated by the last equation.

**8.2.5. Definition.** Let  $\varphi$  be a distribution and let  $f$  be a  $\mathcal{C}_F^\infty$  function whose Fourier transform  $\tilde{f}$  exists and is also a  $\mathcal{C}_F^\infty$  function. Then we define the Fourier transform  $\tilde{\varphi}$  of  $\varphi$  to be the distribution given by

$$\langle \tilde{\varphi}, f \rangle = \langle \varphi, \tilde{f} \rangle.$$

**8.2.6. Example.** The Fourier transform of  $\delta(x)$  is given by

$$\begin{aligned} \langle \delta, f \rangle &= \langle \delta, \tilde{f} \rangle = \tilde{f}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) dt \\ &= \int_{-\infty}^{\infty} \left( \frac{1}{\sqrt{2\pi}} \right) f(t) dt = \left\langle \frac{1}{\sqrt{2\pi}}, f \right\rangle. \end{aligned}$$

Thus,  $\tilde{\delta} = 1/\sqrt{2\pi}$ , as expected.

The Fourier transform of  $\delta(x - x') \equiv \delta_{x'}(x)$  is given by

$$\begin{aligned} \langle \delta_{x'}, f \rangle &= \langle \delta_{x'}, \tilde{f} \rangle = \tilde{f}(x') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-ix't} dt \\ &= \int_{-\infty}^{\infty} \left( \frac{1}{\sqrt{2\pi}} e^{-ix't} \right) f(t) dt. \end{aligned}$$

Thus, if  $\varphi(x) = \delta(x - x')$ , then  $\tilde{\varphi}(t) = (1/\sqrt{2\pi})e^{-ix't}$ . ■

## 8.3 Problems

**8.1.** Consider the function  $f(\theta) = \sum_{m=-\infty}^{\infty} \delta(\theta - 2m\pi)$ .

(a) Show that  $f$  is periodic of period  $2\pi$ .

(b) What is the Fourier series expansion for  $f(\theta)$ .

**8.2.** Break the sum  $\sum_{n=-N}^N e^{in(\theta-\theta')}$  into  $\sum_{n=-N}^{-1} + 1 + \sum_{n=1}^N$ . Use the geometric sum formula

$$\sum_{n=0}^N ar^n = a \frac{r^{N+1} - 1}{r - 1}$$

to obtain

$$\sum_{n=1}^N e^{in(\theta-\theta')} = e^{i(\theta-\theta')} \frac{e^{iN(\theta-\theta')} - 1}{e^{i(\theta-\theta')} - 1} = e^{i\frac{1}{2}(N+1)(\theta-\theta')} \frac{\sin[\frac{1}{2}N(\theta-\theta')]}{\sin[\frac{1}{2}(\theta-\theta')]}.$$

By changing  $n$  to  $-n$  or equivalently,  $(\theta - \theta')$  to  $-(\theta - \theta')$  find a similar sum from  $-N$  to  $-1$ . Now put everything together and use the trigonometric identity

$$2 \cos \alpha \sin \beta = \sin(\alpha + \beta) - \sin(\alpha - \beta)$$

to show that

$$\sum_{n=-N}^N e^{in(\theta-\theta')} = \frac{\sin[(N + \frac{1}{2})(\theta - \theta')]}{\sin[\frac{1}{2}(\theta - \theta')]}.$$

**8.3.** Find the Fourier series expansion of the periodic function defined on its fundamental cell as

$$f(\theta) = \begin{cases} -\frac{1}{2}(\pi + \theta) & \text{if } -\pi \leq \theta < 0, \\ \frac{1}{2}(\pi - \theta) & \text{if } 0 < \theta \leq \pi. \end{cases}$$

**8.4.** Show that  $A_n$  and  $B_n$  in Equation (8.2) are real when  $f(\theta)$  is real.

**8.5.** Find the Fourier series expansion of the periodic function  $f(\theta)$  defined on its fundamental cell,  $(-\pi, \pi)$ , as  $f(\theta) = \cos \alpha\theta$ ,

(a) when  $\alpha$  is an integer.

(b) when  $\alpha$  is not an integer.

**8.6.** Find the Fourier series expansion of the periodic function defined on its fundamental cell,  $(-\pi, \pi)$ , as  $f(\theta) = \theta$ .

**8.7.** Consider the periodic function that is defined on its fundamental cell,  $(-a, a)$ , as  $f(x) = |x|$ .

(a) Find its Fourier series expansion.

(b) Evaluate both sides of the expansion at  $x = 0$ , and show that

$$\pi^2 = 8 \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2}.$$

(c) Show that the infinite series gives the same result as the function when both are evaluated at  $x = a$ .

**8.8.** Let  $f(x) = x$  be a periodic function defined over the interval  $(0, 2a)$ . Find the Fourier series expansion of  $f$ .

**8.9.** Show that the piecewise parabolic “approximation” to  $a^2 \sin(\pi x/a)$  in the interval  $(-a, a)$  given by the function

$$f(x) = \begin{cases} 4x(a+x) & \text{if } -a \leq x \leq 0 \\ 4x(a-x) & \text{if } 0 \leq x \leq a \end{cases}$$

has the Fourier series expansion

$$f(x) = \frac{32a^2}{\pi^3} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^3} \sin \frac{(2n+1)\pi x}{a}.$$

Plot  $f(x)$ ,  $a^2 \sin(\pi x/a)$ , and the series expansion (up to 20 terms) for  $a = 1$  between  $-1$  and  $+1$  on the same graph.

**8.10.** Find the Fourier series expansion of  $f(\theta) = \theta^2$  for  $|\theta| < \pi$ . Then show that

$$\frac{\pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2} \quad \text{and} \quad \frac{\pi^2}{12} = - \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}.$$

**8.11.** Find the Fourier series expansion of

$$f(t) = \begin{cases} \sin \omega t & \text{if } 0 \leq t \leq \pi/\omega, \\ 0 & \text{if } -\pi/\omega \leq t \leq 0. \end{cases}$$

**8.12.** What is the Fourier transform of

- (a) the constant function  $f(x) = C$ , and  
 (b) the Dirac delta function  $\delta(x)$ ?

**8.13.** Show that

- (a) if  $g(x)$  is real, then  $\tilde{g}^*(k) = \tilde{g}(-k)$ , and  
 (b) if  $g(x)$  is even (odd), then  $\tilde{g}(k)$  is also even (odd).

**8.14.** Let  $g_c(x)$  stand for the single function that is nonzero only on a subinterval of the fundamental cell  $(a, a + L)$ . Define the function  $g(x)$  as

$$g(x) = \sum_{j=-\infty}^{\infty} g_c(x - jL).$$

- (a) Show that  $g(x)$  is periodic with period  $L$ .  
 (b) Find its Fourier transform  $\tilde{g}(k)$ , and verify that

$$\tilde{g}(k) = L\tilde{g}_c(k) \sum_{m=-\infty}^{\infty} \delta(kL - 2m\pi).$$

- (c) Find the (inverse) transform of  $\tilde{g}(k)$ , and show that it is the Fourier series of  $g_c(x)$ .

**8.15.** Evaluate the Fourier transform of

$$g(x) = \begin{cases} b - b|x|/a & \text{if } |x| < a, \\ 0 & \text{if } |x| > a. \end{cases}$$

**8.16.** Let  $f(\theta)$  be a periodic function given by  $f(\theta) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ . Find its Fourier transform  $\tilde{f}(t)$ .

**8.17.** Let

$$f(t) = \begin{cases} \sin \omega_0 t & \text{if } |t| < T, \\ 0 & \text{if } |t| > T. \end{cases}$$

Show that

$$\tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}} \left\{ \frac{\sin[(\omega - \omega_0)T]}{\omega - \omega_0} - \frac{\sin[(\omega + \omega_0)T]}{\omega + \omega_0} \right\}.$$

Verify the uncertainty relation  $\Delta\omega\Delta t \approx 4\pi$ .



**8.18.** If  $f(x) = g(x + a)$ , show that  $\tilde{f}(k) = e^{-iak}\tilde{g}(k)$ .

**8.19.** For  $a > 0$  find the Fourier transform of  $f(x) = e^{-a|x|}$ . Is  $\tilde{f}(k)$  symmetric? Is it real? Verify the uncertainty relations.

**8.20.** The displacement of a damped harmonic oscillator is given by

$$f(t) = \begin{cases} Ae^{-\alpha t} e^{i\omega_0 t} & \text{if } t > 0, \\ 0 & \text{if } t < 0. \end{cases}$$

Find  $\tilde{f}(\omega)$  and show that the frequency distribution  $|\tilde{f}(\omega)|^2$  is given by

$$|\tilde{f}(\omega)|^2 = \frac{A^2}{2\pi} \frac{1}{(\omega - \omega_0)^2 + \alpha^2}.$$

convolution theorem

**8.21.** Prove the **convolution theorem**:

$$\int_{-\infty}^{\infty} f(x)g(y-x) dx = \int_{-\infty}^{\infty} \tilde{f}(k)\tilde{g}(k)e^{iky} dk.$$

What will this give when  $y = 0$ ?

Parseval's relation

**8.22.** Prove **Parseval's relation** for Fourier transforms:

$$\int_{-\infty}^{\infty} f(x)g^*(x) dx = \int_{-\infty}^{\infty} \tilde{f}(k)\tilde{g}^*(k) dk.$$

In particular, the norm of a function—with weight function equal to 1—is invariant under Fourier transform.

**8.23.** Use the completeness relation  $\mathbf{1} = \sum_n |n\rangle \langle n|$  and sandwich it between  $|x\rangle$  and  $\langle x'|$  to find an expression for the Dirac delta function in terms of an infinite series of orthonormal functions.

**8.24.** Use a Fourier transform in three dimensions to find a solution of the Poisson equation:  $\nabla^2 \Phi(\mathbf{r}) = -4\pi\rho(\mathbf{r})$ .

**8.25.** For  $\varphi(x) = \delta(x - x')$ , find  $\tilde{\varphi}(y)$ .

**8.26.** Show that  $\tilde{\tilde{f}}(t) = f(-t)$ .

**8.27.** The Fourier transform of a distribution  $\varphi$  is given by

$$\tilde{\varphi}(t) = \sum_{n=0}^{\infty} \frac{1}{n!} \delta'(t - n).$$

What is  $\varphi(x)$ ? Hint: Use  $\tilde{\tilde{\varphi}}(x) = \varphi(-x)$

**8.28.** For  $f(x) = \sum_{k=0}^n a_k x^k$ , show that

$$\tilde{f}(u) = \sqrt{2\pi} \sum_{k=0}^n i^k a_k \delta^{(k)}(u), \quad \text{where} \quad \delta^{(k)}(u) \equiv \frac{d^k}{du^k} \delta(u).$$

## Additional Reading

1. Courant, R. and Hilbert, D. *Methods of Mathematical Physics*, vol. 1, Interscience, 1962. The classic book by two masters. This is a very readable book written specifically for physicists. Its treatment of Fourier series and transforms is very clear.
2. DeVries, P. *A First Course in Computational Physics*, Wiley, 1994. A good discussion of the fast Fourier transform including some illustrative computer programs.
3. Reed, M., and Simon, B. *Fourier Analysis, Self-Adjointness*, Academic Press, 1980. Second volume of a four-volume series. A comprehensive exposition of Fourier analysis with emphasis on operator theory.
4. Richtmyer, R. *Principles of Advanced Mathematical Physics*, Springer-Verlag, 1978. A two-volume book on mathematical physics written in a formal style, but very useful due to its comprehensiveness and the large number of examples drawn from physics. Chapter 4 discusses Fourier analysis and distributions.

Part III

---

# Complex Analysis



# Complex Calculus

Complex analysis, just like real analysis, deals with questions of continuity, convergence of series, differentiation, integration, and so forth. The reader is assumed to have been exposed to the *algebra* of complex numbers.

## 9.1 Complex Functions

A complex function is a map  $f : \mathbb{C} \rightarrow \mathbb{C}$ , and we write  $f(z) = w$ , where both  $z$  and  $w$  are complex numbers.<sup>1</sup> The map  $f$  can be geometrically thought of as a correspondence between two complex planes, the  $z$ -plane and the  $w$ -plane. The  $w$ -plane has a real axis and an imaginary axis, which we can call  $u$  and  $v$ , respectively. Both  $u$  and  $v$  are real functions of the coordinates of  $z$ , i.e.,  $x$  and  $y$ . Therefore, we may write

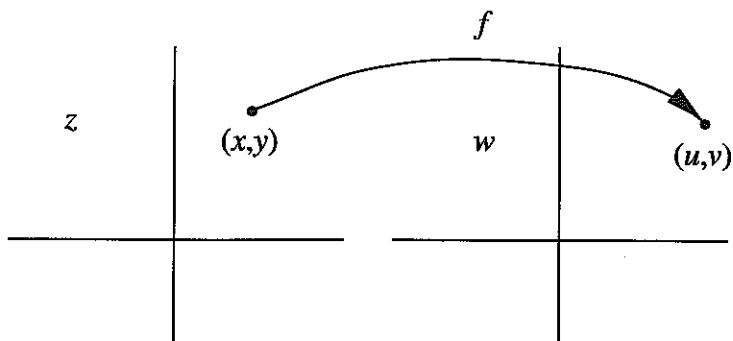
$$f(z) = u(x, y) + iv(x, y). \quad (9.1)$$

This equation gives a unique point  $(u, v)$  in the  $w$ -plane for each point  $(x, y)$  in the  $z$ -plane (see Figure 9.1). Under  $f$ , regions of the  $z$ -plane are mapped onto regions of the  $w$ -plane. For instance, a curve in the  $z$ -plane may be mapped into a curve in the  $w$ -plane. The following example illustrates this point.

**9.1.1. Example.** Let us investigate the behavior of a couple of elementary complex functions. In particular, we shall look at the way a line  $y = mx$  in the  $z$ -plane is mapped into curves in the  $w$ -plane.

---

<sup>1</sup>Strictly speaking, we should write  $f : S \rightarrow \mathbb{C}$  where  $S$  is a *subset* of the complex plane. The reason is that most functions are not defined for the entire set of complex numbers, so that the domain of such functions is not necessarily  $\mathbb{C}$ . We shall specify the domain only when it is absolutely necessary. Otherwise, we use the generic notation  $f : \mathbb{C} \rightarrow \mathbb{C}$ , even though  $f$  is defined only on a subset of  $\mathbb{C}$ .



**Figure 9.1** A map from the  $z$ -plane to the  $w$ -plane.

(a) For  $w = f(z) = z^2$ , we have

$$w = (x + iy)^2 = x^2 - y^2 + 2ixy,$$

with  $u(x, y) = x^2 - y^2$  and  $v(x, y) = 2xy$ . For  $y = mx$ , i.e., for a line in the  $z$ -plane with slope  $m$ , these equations yield  $u = (1 - m^2)x^2$  and  $v = 2mx^2$ . Eliminating  $x$  in these equations, we find  $v = [2m/(1 - m^2)]u$ . This is a line passing through the origin of the  $w$ -plane [see Figure 9.2(a)]. Note that the angle the image line makes with the real axis of the  $w$ -plane is twice the angle the original line makes with the  $x$ -axis. (Show this!). (b) The function  $w = f(z) = e^z = e^{x+iy}$  gives  $u(x, y) = e^x \cos y$  and  $v(x, y) = e^x \sin y$ . Substituting  $y = mx$ , we obtain  $u = e^x \cos mx$  and  $v = e^x \sin mx$ . Unlike part (a), we cannot eliminate  $x$  to find  $v$  as an explicit function of  $u$ . Nevertheless, the last pair of equations are *parametric equations* of a curve, which we can plot in a  $uv$ -plane as shown in Figure 9.2(b). ■

Limits of complex functions are defined in terms of absolute values. Thus,  $\lim_{z \rightarrow a} f(z) = w_0$  means that given any real number  $\epsilon > 0$ , we can find a corresponding real number  $\delta > 0$  such that  $|f(z) - w_0| < \epsilon$  whenever  $|z - a| < \delta$ . Similarly, we say that a function  $f$  is **continuous** at  $z = a$  if  $\lim_{z \rightarrow a} f(z) = f(a)$ .

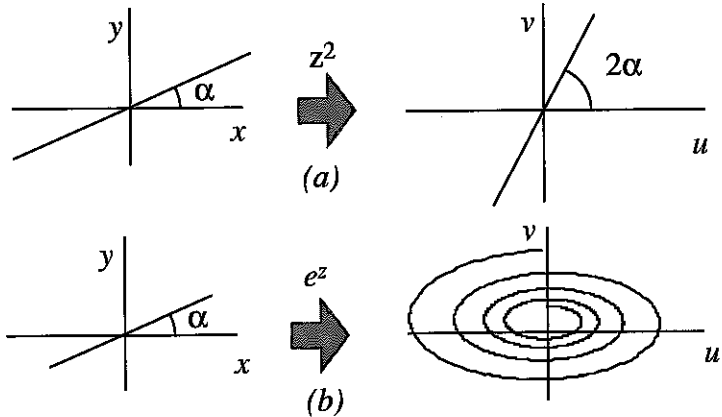
## 9.2 Analytic Functions

The derivative of a complex function is defined as usual:

**9.2.1. Definition.** Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be a complex function. The derivative of  $f$  at  $z_0$  is

$$\left. \frac{df}{dz} \right|_{z_0} = \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z},$$

provided that the limit exists and is independent of  $\Delta z$ .



**Figure 9.2** (a) The map  $z^2$  takes a line with slope angle  $\alpha$  and maps it to a line with twice the angle in the  $w$ -plane. (b) The map  $e^z$  takes the same line and maps it to a spiral in the  $w$ -plane.

In this definition “independent of  $\Delta z$ ” means independent of  $\Delta x$  and  $\Delta y$  (the components of  $\Delta z$ ) and, therefore, independent of the direction of approach to  $z_0$ . The restrictions of this definition apply to the real case as well. For instance, the derivative of  $f(x) = |x|$  at  $x = 0$  does not exist<sup>2</sup> because it approaches  $+1$  from the right and  $-1$  from the left.

It can easily be shown that all the formal rules of differentiation that apply to the real case also apply to the complex case. For example, if  $f$  and  $g$  are differentiable, then  $f \pm g$ ,  $fg$ , and—as long as  $g$  is not zero— $f/g$  are also differentiable, and their derivatives are given by the usual rules of differentiation.

Example illustrating path dependence of derivative

**9.2.2. Example.** Let us examine the derivative of  $f(z) = x^2 + 2iy^2$  at  $z = 1 + i$ :

$$\begin{aligned} \left. \frac{df}{dz} \right|_{z=1+i} &= \lim_{\Delta z \rightarrow 0} \frac{f(1+i+\Delta z) - f(1+i)}{\Delta z} \\ &= \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{(1+\Delta x)^2 + 2i(1+\Delta y)^2 - 1 - 2i}{\Delta x + i\Delta y} \\ &= \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{2\Delta x + 4i\Delta y + (\Delta x)^2 + 2i(\Delta y)^2}{\Delta x + i\Delta y} \end{aligned}$$

Let us approach  $z = 1 + i$  along the line  $y - 1 = m(x - 1)$ . Then  $\Delta y = m\Delta x$ , and the limit yields

$$\left. \frac{df}{dz} \right|_{z=1+i} = \lim_{\Delta x \rightarrow 0} \frac{2\Delta x + 4im\Delta x + (\Delta x)^2 + 2im^2(\Delta x)^2}{\Delta x + im\Delta x} = \frac{2 + 4im}{1 + im}$$

<sup>2</sup>One can rephrase this and say that the derivative exists, but not in terms of ordinary functions, rather, in terms of *generalized functions*—in this case  $\theta(x)$ —discussed in Chapter 6.

It follows that we get infinitely many values for the derivative depending on the value we assign to  $m$ , i.e., depending on the direction along which we approach  $1 + i$ . Thus, the derivative does not exist at  $z = 1 + i$ . ■

It is clear from the definition that differentiability puts a severe restriction on  $f(z)$  because it requires the limit to be the same for *all paths* going through  $z_0$ . Furthermore, differentiability is a *local* property: To test whether or not a function  $f(z)$  is differentiable at  $z_0$ , we move away from  $z_0$  by a small amount  $\Delta z$  and check the existence of the limit in Definition 9.2.1.

What are the conditions under which a complex function is differentiable? For  $f(z) = u(x, y) + iv(x, y)$ , Definition 9.2.1 yields

$$\left. \frac{df}{dz} \right|_{z_0} = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \left\{ \frac{u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0)}{\Delta x + i\Delta y} + i \frac{v(x_0 + \Delta x, y_0 + \Delta y) - v(x_0, y_0)}{\Delta x + i\Delta y} \right\}.$$

If this limit is to exist for all paths, it must exist for the two particular paths on which  $\Delta y = 0$  (parallel to the  $x$ -axis) and  $\Delta x = 0$  (parallel to the  $y$ -axis). For the first path we get

$$\begin{aligned} \left. \frac{df}{dz} \right|_{z_0} &= \lim_{\Delta x \rightarrow 0} \frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0)}{\Delta x} \\ &\quad + i \lim_{\Delta x \rightarrow 0} \frac{v(x_0 + \Delta x, y_0) - v(x_0, y_0)}{\Delta x} = \left. \frac{\partial u}{\partial x} \right|_{(x_0, y_0)} + i \left. \frac{\partial v}{\partial x} \right|_{(x_0, y_0)}. \end{aligned}$$

For the second path ( $\Delta x = 0$ ), we obtain

$$\begin{aligned} \left. \frac{df}{dz} \right|_{z_0} &= \lim_{\Delta y \rightarrow 0} \frac{u(x_0, y_0 + \Delta y) - u(x_0, y_0)}{i\Delta y} \\ &\quad + i \lim_{\Delta y \rightarrow 0} \frac{v(x_0, y_0 + \Delta y) - v(x_0, y_0)}{i\Delta y} = -i \left. \frac{\partial u}{\partial y} \right|_{(x_0, y_0)} + \left. \frac{\partial v}{\partial y} \right|_{(x_0, y_0)}. \end{aligned}$$

If  $f$  is to be differentiable at  $z_0$ , the derivatives along the two paths must be equal. Equating the real and imaginary parts of both sides of this equation and ignoring the subscript  $z_0$  ( $x_0, y_0$ , or  $z_0$  is arbitrary), we obtain

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \quad (9.2)$$

These two conditions, which are necessary for the differentiability of  $f$ , are called the **Cauchy–Riemann conditions**.

An alternative way of writing the Cauchy–Riemann (C-R) conditions is obtained by making the substitution<sup>3</sup>  $x = \frac{1}{2}(z + z^*)$  and  $y = \frac{1}{2i}(z - z^*)$  in  $u(x, y)$

Cauchy–Riemann  
conditions

<sup>3</sup>We use  $z^*$  to indicate the complex conjugate of  $z$ . Occasionally we may use  $\bar{z}$ .



and  $v(x, y)$ , using the chain rule to write Equation (9.2) in terms of  $z$  and  $z^*$ , substituting the results in  $\frac{\partial f}{\partial z^*} = \frac{\partial u}{\partial z^*} + i \frac{\partial v}{\partial z^*}$  and showing that Equation (9.2) is equivalent to the single equation  $\partial f / \partial z^* = 0$ . This equation says that

**9.2.3. Box.** *If  $f$  is to be differentiable, it must be independent of  $z^*$ .*

Expression for the  
derivative of a  
differentiable  
complex function

If the derivative of  $f$  exists, the arguments leading to Equation (9.2) imply that the derivative can be expressed as

$$\frac{df}{dz} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}. \quad (9.3)$$

The C-R conditions assure us that these two equations are equivalent.

The following example illustrates the differentiability of complex functions.

**9.2.4. Example.** Let us determine whether or not the following functions are differentiable:

(a) We have already established that  $f(z) = x^2 + 2iy^2$  is not differentiable at  $z = 1 + i$ . We can now show that it has no derivative at any point in the complex plane (except at the origin). This is easily seen by noting that  $u = x^2$  and  $v = 2y^2$ , and that  $\partial u / \partial x = 2x \neq \partial v / \partial y = 4y$ , and the first Cauchy–Riemann condition is not satisfied. The second C-R condition is satisfied, but that is not enough.

We can also write  $f(z)$  in terms of  $z$  and  $z^*$ :

$$f(z) = \left[ \frac{1}{2}(z + z^*) \right]^2 + 2i \left[ \frac{1}{2i}(z - z^*) \right]^2 = \frac{1}{4}(1 - 2i)(z^2 + z^{*2}) + \frac{1}{2}(1 + 2i)zz^*.$$

$f(z)$  has an explicit dependence on  $z^*$ . Therefore, it is not differentiable.

(b) Now consider  $f(z) = x^2 - y^2 + 2ixy$ , for which  $u = x^2 - y^2$  and  $v = 2xy$ . The C-R conditions become  $\partial u / \partial x = 2x = \partial v / \partial y$  and  $\partial u / \partial y = -2y = -\partial v / \partial x$ . Thus,  $f(z)$  may be differentiable. Recall that the C-R conditions are only *necessary* conditions; we have not shown (but we will, shortly) that they are also *sufficient*.

To check the dependence of  $f$  on  $z^*$ , substitute  $x = (z + z^*)/2$  and  $y = (z - z^*)/(2i)$  in  $u$  and  $v$  to show that  $f(z) = z^2$ , and thus there is no  $z^*$  dependence.

(c) Let  $u(x, y) = e^x \cos y$  and  $v(x, y) = e^x \sin y$ . Then  $\partial u / \partial x = e^x \cos y = \partial v / \partial y$  and  $\partial u / \partial y = -e^x \sin y = -\partial v / \partial x$ , and the C-R conditions are satisfied. Also,

$$f(z) = e^x \cos y + i e^x \sin y = e^x (\cos y + i \sin y) = e^x e^{iy} = e^{x+iy} = e^z,$$

and there is no  $z^*$  dependence. ■

The requirement of differentiability is very restrictive: The derivative must exist along infinitely many paths. On the other hand, the C-R conditions seem deceptively mild: They are derived for only two paths. Nevertheless, the two paths are, in fact, true representatives of all paths; that is, the C-R conditions are not only necessary, but also sufficient:

**9.2.5. Theorem.** *The function  $f(z) = u(x, y) + iv(x, y)$  is differentiable in a region of the complex plane if and only if the Cauchy–Riemann conditions,*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

*(or, equivalently,  $\partial f/\partial z^* = 0$ ), are satisfied and all first partial derivatives of  $u$  and  $v$  are continuous in that region. In that case*

$$\frac{df}{dz} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}.$$

*Proof.* We have already shown the “only if” part. To show the “if” part, note that if the derivative exists at all, it must equal (9.3). Thus, we have to show that

$$\lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}$$

or, equivalently, that

$$\left| \frac{f(z + \Delta z) - f(z)}{\Delta z} - \left( \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) \right| < \epsilon \quad \text{whenever} \quad |\Delta z| < \delta.$$

By definition,

$$f(z + \Delta z) - f(z) = u(x + \Delta x, y + \Delta y) + iv(x + \Delta x, y + \Delta y) - u(x, y) - iv(x, y).$$

Since  $u$  and  $v$  have continuous first partial derivatives, we can write

$$u(x + \Delta x, y + \Delta y) = u(x, y) + \frac{\partial u}{\partial x} \Delta x + \frac{\partial u}{\partial y} \Delta y + \epsilon_1 \Delta x + \delta_1 \Delta y,$$

$$v(x + \Delta x, y + \Delta y) = v(x, y) + \frac{\partial v}{\partial x} \Delta x + \frac{\partial v}{\partial y} \Delta y + \epsilon_2 \Delta x + \delta_2 \Delta y,$$

where  $\epsilon_1, \epsilon_2, \delta_1,$  and  $\delta_2$  are real numbers that approach zero as  $\Delta x$  and  $\Delta y$  approach zero. Using these expressions, we can write

$$\begin{aligned} f(z + \Delta z) - f(z) &= \left( \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) \Delta x + i \left( -i \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} \right) \Delta y \\ &\quad + (\epsilon_1 + i\epsilon_2) \Delta x + (\delta_1 + i\delta_2) \Delta y \\ &= \left( \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) (\Delta x + i \Delta y) + \epsilon \Delta x + \delta \Delta y, \end{aligned}$$

where  $\epsilon \equiv \epsilon_1 + i\epsilon_2, \delta \equiv \delta_1 + i\delta_2,$  and we used the C-R conditions in the last step. Dividing both sides by  $\Delta z = \Delta x + i \Delta y,$  we get

$$\frac{f(z + \Delta z) - f(z)}{\Delta z} - \left( \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) = \epsilon \frac{\Delta x}{\Delta z} + \delta \frac{\Delta y}{\Delta z}.$$

By the triangle inequality,  $|RHS| \leq |\epsilon_1 + i\epsilon_2| + |\delta_1 + i\delta_2|$ . This follows from the fact that  $|\Delta x|/|\Delta z|$  and  $|\Delta y|/|\Delta z|$  are both equal to at most 1. The  $\epsilon$  and  $\delta$  terms can be made as small as desired by making  $\Delta z$  small enough. We have thus established that when the C-R conditions hold, the function  $f$  is differentiable.  $\square$

---

**Augustin-Louis Cauchy (1789–1857)** was one of the most influential French mathematicians of the nineteenth century. He began his career as a military engineer, but when his health broke down in 1813 he followed his natural inclination and devoted himself wholly to mathematics.



In mathematical productivity Cauchy was surpassed only by Euler, and his collected works fill 27 fat volumes. He made substantial contributions to number theory and determinants; is considered to be the originator of the theory of finite groups; and did extensive work in astronomy, mechanics, optics, and the theory of elasticity.

His greatest achievements, however, lay in the field of analysis. Together with his contemporaries Gauss and Abel, he was a pioneer in the rigorous treatment of limits, continuous functions, derivatives, integrals, and infinite series. Several of the basic tests for the convergence of series are associated with his name. He also provided the first existence proof for solutions of differential equations, gave the first proof of the convergence of a Taylor series, and was the first to feel the need for a careful study of the convergence behavior of Fourier series (see Chapter 8). However, his most important work was in the theory of functions of a complex variable, which in essence he created and which has continued to be one of the dominant branches of both pure and applied mathematics. In this field, Cauchy's integral theorem and Cauchy's integral formula are fundamental tools without which modern analysis could hardly exist (see Chapter 9).

Unfortunately, his personality did not harmonize with the fruitful power of his mind. He was an arrogant royalist in politics and a self-righteous, preaching, pious believer in religion—all this in an age of republican skepticism—and most of his fellow scientists disliked him and considered him a smug hypocrite. It might be fairer to put first things first and describe him as a great mathematician who happened also to be a sincere but narrow-minded bigot.

---

analyticity and singularity; regular and singular points; entire functions

**9.2.6. Definition.** A function  $f : \mathbb{C} \rightarrow \mathbb{C}$  is called **analytic** at  $z_0$  if it is differentiable at  $z_0$  and at all other points in some neighborhood of  $z_0$ . A point at which  $f$  is analytic is called a **regular point** of  $f$ . A point at which  $f$  is not analytic is called a **singular point** or a **singularity** of  $f$ . A function for which all points in  $\mathbb{C}$  are regular is called an **entire function**.

**9.2.7. Example.** DERIVATIVES OF SOME FUNCTIONS

(a)  $f(z) = z$ .

Here  $u = x$  and  $v = y$ ; the C-R conditions are easily shown to hold, and for any  $z$ , we have  $df/dz = \partial u/\partial x + i\partial v/\partial x = 1$ . Therefore, the derivative exists at all points of the complex

plane. (b)  $f(z) = z^2$ .

Here  $u = x^2 - y^2$  and  $v = 2xy$ ; the C-R conditions hold, and for all points  $z$  of the complex plane, we have  $df/dz = \partial u/\partial x + i\partial v/\partial x = 2x + i2y = 2z$ . Therefore,  $f(z)$  is differentiable at all points. (c)  $f(z) = z^n$  for  $n \geq 1$ .

We can use mathematical induction and the fact that the product of two entire functions is an entire function to show that  $\frac{d}{dz}(z^n) = nz^{n-1}$ . (d)  $f(z) = a_0 + a_1z + \cdots + a_{n-1}z^{n-1} + a_nz^n$ ,

where  $a_i$  are arbitrary constants. That  $f(z)$  is entire follows directly from part (c) and the fact that the sum of two entire functions is entire. (e)  $f(z) = 1/z$ . The derivative can be found to be  $f'(z) = -1/z^2$ , which does not exist for  $z = 0$ . Thus,  $z = 0$  is a singularity of  $f(z)$ . However, any other point is a regular point of  $f$ . (f)  $f(z) = |z|^2$ . Using the definition of the derivative, we obtain

$$\frac{\Delta f}{\Delta z} = \frac{|z + \Delta z|^2 - |z|^2}{\Delta z} = \frac{(z + \Delta z)(z^* + \Delta z^*) - zz^*}{\Delta z} = z^* + \Delta z^* + z \frac{\Delta z^*}{\Delta z}.$$

For  $z = 0$ ,  $\Delta f/\Delta z = \Delta z^*$ , which goes to zero as  $\Delta z \rightarrow 0$ . Therefore,  $df/dz = 0$  at  $z = 0$ .<sup>4</sup> However, if  $z \neq 0$ , the limit of  $\Delta f/\Delta z$  will depend on how  $z$  is approached. Thus,  $df/dz$  does not exist if  $z \neq 0$ . This shows that  $|z|^2$  is differentiable only at  $z = 0$  and nowhere else in its neighborhood. It also shows that even if the real (here,  $u = x^2 + y^2$ ) and imaginary (here,  $v = 0$ ) parts of a complex function have continuous partial derivatives of all orders at a point, the function may not be differentiable there. (g)  $f(z) = 1/\sin z$ : This gives  $df/dz = -\cos z/\sin^2 z$ . Thus,  $f$  has infinitely many (isolated) singular points at  $z = \pm n\pi$  for  $n = 0, 1, 2, \dots$  ■

### 9.2.8. Example. THE COMPLEX EXPONENTIAL FUNCTION

In this example, we find the (unique) function  $f: \mathbb{C} \rightarrow \mathbb{C}$  that has the following three properties:

- (a)  $f$  is single-valued and analytic for all  $z$ ,
- (b)  $df/dz = f(z)$ , and
- (c)  $f(z_1 + z_2) = f(z_1)f(z_2)$ .

Property (b) shows that if  $f(z)$  is well behaved, then  $df/dz$  is also well behaved. In particular, if  $f(z)$  is defined for all values of  $z$ , then  $f$  must be entire.

For  $z_1 = 0 = z_2$ , property (c) yields  $f(0) = [f(0)]^2 \Rightarrow f(0) = 1$ , or  $f(0) = 0$ . On the other hand,

$$\frac{df}{dz} = \lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{f(z)f(\Delta z) - f(z)}{\Delta z} = f(z) \lim_{\Delta z \rightarrow 0} \frac{f(\Delta z) - 1}{\Delta z}.$$

Property (b) now implies that

$$\lim_{\Delta z \rightarrow 0} \frac{f(\Delta z) - 1}{\Delta z} = 1 \Rightarrow f'(0) = 1 \quad \text{and} \quad f(0) = 1.$$

<sup>4</sup>Although the derivative of  $|z|^2$  exists at  $z = 0$ , it is not analytic there (or anywhere else). To be analytic at a point, a function must have derivatives at *all points* in some neighborhood of the given point.

The first implication follows from the definition of derivative, and the second from the fact that the only other choice, namely  $f(0) = 0$ , would yield  $-\infty$  for the limit.

Now, we write  $f(z) = u(x, y) + iv(x, y)$ , for which property (b) becomes

$$\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = u + iv \Rightarrow \frac{\partial u}{\partial x} = u, \quad \frac{\partial v}{\partial x} = v.$$

These equations have the most general solution  $u(x, y) = a(y)e^x$  and  $v(x, y) = b(y)e^x$ , where  $a(y)$  and  $b(y)$  are the “constants” of integration. The Cauchy–Riemann conditions now yield  $a(y) = db/dy$  and  $da/dy = -b(y)$ , whose most general solution is  $a(y) = A \cos y + B \sin y$ ,  $b(y) = A \sin y - B \cos y$ . On the other hand,  $f(0) = 1$  yields  $u(0, 0) = 1$  and  $v(0, 0) = 0$ , implying that  $a(0) = 1$ ,  $b(0) = 0$  or  $A = 1$ ,  $B = 0$ . We therefore conclude that

$$f(z) = a(y)e^x + ib(y)e^x = e^x(\cos y + i \sin y) = e^x e^{iy} = e^z.$$

Both  $e^x$  and  $e^{iy}$  are well-defined in the entire complex plane. Hence,  $e^z$  is defined and differentiable over all  $\mathbb{C}$ ; therefore, it is entire. ■

Example 9.2.7 shows that any polynomial in  $z$  is entire. Example 9.2.8 shows that the exponential function  $e^z$  is also entire. Therefore, any product and/or sum of polynomials and  $e^z$  will also be entire. We can build other entire functions. For instance,  $e^{iz}$  and  $e^{-iz}$  are entire functions; therefore, the trigonometric functions, defined as

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i} \quad \text{and} \quad \cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad (9.4)$$

are also entire functions. Problem 9.5 shows that  $\sin z$  and  $\cos z$  have only *real* zeros. The hyperbolic functions can be defined similarly:

$$\sinh z = \frac{e^z - e^{-z}}{2} \quad \text{and} \quad \cosh z = \frac{e^z + e^{-z}}{2}. \quad (9.5)$$

Although the sum and the product of entire functions are entire, the ratio, in general, is not. For instance, if  $f(z)$  and  $g(z)$  are polynomials of degrees  $m$  and  $n$ , respectively, then for  $n > 0$ , the ratio  $f(z)/g(z)$  is not entire, because at the zeros of  $g(z)$ —which always exist and we assume that it is not a zero of  $f(z)$ —the derivative is not defined.

The functions  $u(x, y)$  and  $v(x, y)$  of an analytic function have an interesting property that the following example investigates.

**9.2.9. Example.** The family of curves  $u(x, y) = \text{constant}$  is perpendicular to the family of curves  $v(x, y) = \text{constant}$  at each point of the complex plane where  $f(z) = u + iv$  is analytic.

This can easily be seen by looking at the normal to the curves. The normal to the curve  $u(x, y) = \text{constant}$  is simply  $\nabla u = (\partial u/\partial x, \partial u/\partial y)$ . Similarly, the normal to the curve  $v(x, y) = \text{constant}$  is  $\nabla v = (\partial v/\partial x, \partial v/\partial y)$ . Taking the dot product of these two normals, we obtain

$$(\nabla u) \cdot (\nabla v) = \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} = \frac{\partial u}{\partial x} \left( -\frac{\partial u}{\partial y} \right) + \frac{\partial u}{\partial y} \left( \frac{\partial u}{\partial x} \right) = 0$$

by the C-R conditions. ■

### 9.3 Conformal Maps

The real and imaginary parts of an analytic function separately satisfy the two-dimensional Laplace's equation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \quad (9.6)$$

This can easily be verified from the C-R conditions. Laplace's equation in three dimensions,

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0,$$

describes the electrostatic potential  $\Phi$  in a charge-free region of space. In a typical electrostatic problem the potential  $\Phi$  is given at certain boundaries (usually conducting surfaces), and its value at every point in space is sought. There are numerous techniques for solving such problems, and some of them will be discussed later in the book. However, some of these problems have a certain degree of symmetry that reduces them to two-dimensional problems. In such cases, the theory of analytic functions can be extremely helpful.

The symmetry mentioned above is cylindrical symmetry, where the potential is known a priori to be independent of the  $z$ -coordinate (the axis of symmetry). This situation occurs when conductors are cylinders and—if there are charge distributions in certain regions of space—the densities are  $z$ -independent. In such cases,  $\partial\Phi/\partial z = 0$ , and the problem reduces to a two-dimensional one.

harmonic functions

Functions satisfying Laplace's equation are called **harmonic functions**. Thus, the electrostatic potential is a three-dimensional harmonic function, and the potential for a cylindrically symmetric charge distribution and boundary condition is a two-dimensional harmonic function. Since the real and the imaginary parts of a complex analytic function are also harmonic, techniques of complex analysis are sometimes useful in solving electrostatic problems with cylindrical symmetry.<sup>5</sup>

complex potential

To illustrate the connection between electrostatics and complex analysis, consider a long straight filament with a constant linear charge density  $\lambda$ . It is shown in introductory electromagnetism that the potential  $\Phi$  (disregarding the arbitrary constant that determines the reference potential) is given, in cylindrical coordinates, by  $\Phi = 2\lambda \ln \rho = 2\lambda \ln[(x^2 + y^2)^{1/2}] = 2\lambda \ln |z|$ . Since  $\Phi$  satisfies Laplace's equation, we conclude that  $\Phi$  *could* be the real part of an analytic function  $w(z)$ , which we call the **complex potential**. Example 9.2.9, plus the fact that the curves  $u = \Phi = \text{constant}$  are circles, imply that the constant- $v$  curves are rays, i.e.,  $v \propto \varphi$ . Choosing the constant of proportionality as  $2\lambda$ , we obtain

$$w(z) = 2\lambda \ln \rho + i2\lambda \varphi = 2\lambda \ln(\rho e^{i\varphi}) = 2\lambda \ln z$$

<sup>5</sup>We use electrostatics because it is more familiar to physics students. Engineering students are familiar with steady state heat transfer as well, which also involves Laplace's equation, and therefore is amenable to this technique.

It is useful to know the complex potential of more than one filament of charge. To find such a potential we must first find  $w(z)$  for a line charge when it is displaced from the origin. If the line is located at  $z_0 = x_0 + iy_0$ , then it is easy to show that  $w(z) = 2\lambda \ln(z - z_0)$ . If there are  $n$  line charges located at  $z_1, z_2, \dots, z_n$ , then

$$w(z) = 2 \sum_{k=1}^n \lambda_k \ln(z - z_k). \quad (9.7)$$

The function  $w(z)$  can be used directly to solve a number of electrostatic problems involving simple charge distributions and conductor arrangements. Some of these are illustrated in problems at the end of this chapter. Instead of treating  $w(z)$  as a complex potential, let us look at it as a map from the  $z$ -plane (or  $xy$ -plane) to the  $w$ -plane (or  $uv$ -plane). In particular, the equipotential curves (circles) are mapped onto *lines parallel to the  $v$ -axis* in the  $w$ -plane. This is so because equipotential curves are defined by  $u = \text{constant}$ . Similarly, the constant- $v$  curves are mapped onto horizontal lines in the  $w$ -plane.

This is an enormous simplification of the geometry. Straight lines, especially when they are parallel to axes, are by far simpler geometrical objects than circles,<sup>6</sup> especially if the circles are not centered at the origin. So let us consider two complex “worlds.” One is represented by the  $xy$ -plane and denoted by  $z$ . The other, the “prime world,” is represented<sup>7</sup> by  $z'$ , and its real and imaginary parts by  $x'$  and  $y'$ . We start in  $z$ , where we need to find a physical quantity such as the electrostatic potential  $\Phi(x, y)$ . If the problem is too complicated in the  $z$ -world, we transfer it to the  $z'$ -world, in which it may be easily solvable; we solve the problem there (in terms of  $x'$  and  $y'$ ) and then transfer back to the  $z$ -world ( $x$  and  $y$ ). The mapping that relates  $z$  and  $z'$  must be cleverly chosen. Otherwise, there is no guarantee that the problem will simplify.

Two conditions are necessary for the above strategy to work. First, the differential equation describing the physics must not get more complicated with the transfer to  $z'$ . Since Laplace’s equation is already of the simplest type, the  $z'$ -world must also respect Laplace’s equation. Second, and more importantly, the mapping must preserve the angles between curves. This is necessary because we want the equipotential curves and the field lines to be perpendicular in both worlds. A mapping that preserves the angle between two curves at a given point is called a **conformal mapping**. We already have such mappings at our disposal, as the following proposition shows.

**9.3.1. Proposition.** *Let  $\gamma_1$  and  $\gamma_2$  be curves in the complex  $z$ -plane that intersect at a point  $z_0$  at an angle  $\alpha$ . Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be a mapping given by  $f(z) = z' = x' + iy'$  that is analytic at  $z_0$ . Let  $\gamma'_1$  and  $\gamma'_2$  be the images of  $\gamma_1$  and  $\gamma_2$  under this mapping, which intersect at an angle  $\alpha'$ . Then,*

<sup>6</sup>This statement is valid only in Cartesian coordinates. But these are precisely the coordinates we are using in this discussion.

<sup>7</sup>We are using  $z'$  instead of  $w$ , and  $(x', y')$  instead of  $(u, v)$ .

(a)  $\alpha' = \alpha$ , that is, the mapping  $f$  is conformal, if  $(dz'/dz)_{z_0} \neq 0$ .

(b) If  $f$  is harmonic in  $(x, y)$ , it is also harmonic in  $(x', y')$ .

*Proof.* We sketch the proof of the first part. The details, as well as the proof of the second part, involve partial differentiation and the chain rule and are left for the reader. The angle between the two curves is obtained by taking the inner product of the two unit vectors tangent to the curves at  $z_0$ . A small displacement along  $\gamma_i$  can be written as  $\hat{e}_x \Delta x_i + \hat{e}_y \Delta y_i$  for  $i = 1, 2$ , and the unit vectors as

$$\hat{e}_i = \frac{\hat{e}_x \Delta x_i + \hat{e}_y \Delta y_i}{\sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}} \quad \text{for } i = 1, 2.$$

Therefore,

$$\hat{e}_1 \cdot \hat{e}_2 = \frac{\Delta x_1 \Delta x_2 + \Delta y_1 \Delta y_2}{\sqrt{(\Delta x_1)^2 + (\Delta y_1)^2} \sqrt{(\Delta x_2)^2 + (\Delta y_2)^2}}.$$

Similarly, in the prime plane, we have

$$\hat{e}'_1 \cdot \hat{e}'_2 = \frac{\Delta x'_1 \Delta x'_2 + \Delta y'_1 \Delta y'_2}{\sqrt{(\Delta x'_1)^2 + (\Delta y'_1)^2} \sqrt{(\Delta x'_2)^2 + (\Delta y'_2)^2}},$$

where  $x' = u(x, y)$  and  $y' = v(x, y)$  and  $u$  and  $v$  are the real and imaginary parts of the analytic function  $f$ . Using the relations

$$\Delta x'_i = \frac{\partial u}{\partial x} \Delta x_i + \frac{\partial u}{\partial y} \Delta y_i, \quad \Delta y'_i = \frac{\partial v}{\partial x} \Delta x_i + \frac{\partial v}{\partial y} \Delta y_i, \quad i = 1, 2,$$

and the Cauchy–Riemann conditions, the reader may verify that  $\hat{e}'_1 \cdot \hat{e}'_2 = \hat{e}_1 \cdot \hat{e}_2$ .  $\square$

The following are some examples of conformal mappings.

- translation (a)  $z' = z + a$ , where  $a$  is an arbitrary complex constant. This is simply a **translation** of the  $z$ -plane.
- dilation (b)  $z' = bz$ , where  $b$  is an arbitrary complex constant. This is a **dilation** whereby distances are dilated by a factor  $|b|$ . A graph in the  $z$ -plane is mapped onto a *similar* (congruent) graph in the  $z'$ -plane that will be reduced ( $|b| < 1$ ) or enlarged ( $|b| > 1$ ) by a factor of  $|b|$ .
- inversion (c)  $z' = 1/z$ . This is called an **inversion**. Example 9.3.2 will show that under such a mapping, circles are mapped onto circles or straight lines.
- (d) Combining the preceding three transformations yields the general mapping

$$z' = \frac{az + b}{cz + d}, \tag{9.8}$$

which is conformal if  $cz + d \neq 0 \neq dz'/dz$ . The latter conditions are equivalent to  $ad - bc \neq 0$ .



**9.3.2. Example.** A circle of radius  $r$  whose center is at  $a$  in the  $z$ -plane is described by the equation  $|z - a| = r$ . When transforming to the  $z'$ -plane under inversion, this equation becomes  $|1/z' - a| = r$ , or  $|1 - az'| = r|z'|$ . Squaring both sides and simplifying yields  $(r^2 - |a|^2)|z'|^2 + 2\operatorname{Re}(az') - 1 = 0$ . In terms of Cartesian coordinates, this becomes

$$(r^2 - |a|^2)(x'^2 + y'^2) + 2(a_r x' - a_i y') - 1 = 0, \tag{9.9}$$

where  $a \equiv a_r + ia_i$ . We now consider two cases:

1.  $r \neq |a|$ : Divide by  $r^2 - |a|^2$  and complete the squares to get

$$\left(x' + \frac{a_r}{r^2 - |a|^2}\right)^2 + \left(y' - \frac{a_i}{r^2 - |a|^2}\right)^2 - \frac{a_r^2 + a_i^2}{(r^2 - |a|^2)^2} - \frac{1}{r^2 - |a|^2} = 0,$$

or defining  $a'_r \equiv -a_r/(r^2 - |a|^2)$ ,  $a'_i \equiv a_i/(r^2 - |a|^2)$ , and  $r' \equiv r/|r^2 - |a|^2|$ , we have  $(x' - a'_r)^2 + (y' - a'_i)^2 = r'^2$ , which can also be written as

$$|z' - a'| = r', \quad a' = a'_r + ia'_i = \frac{a^*}{|a|^2 - r^2}.$$

This is a circle in the  $z'$ -plane with center at  $a'$  and radius of  $r'$ .

2.  $r = |a|$ : Then Equation (9.9) reduces to  $a_r x' - a_i y' = \frac{1}{2}$ , which is the equation of a line.

If we use the transformation  $z' = 1/(z - c)$  instead of  $z' = 1/z$ , then  $|z - a| = r$  becomes  $|1/z' - (a - c)| = r$ , and all the above analysis will go through exactly as before, except that  $a$  is replaced by  $a - c$ . ■

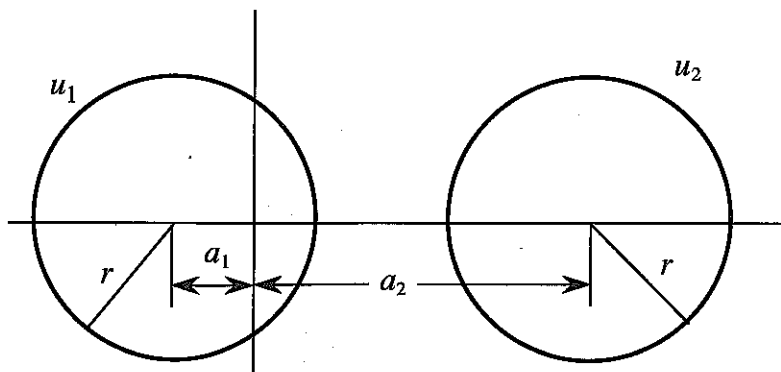
homographic transformations

Mappings of the form given in Equation (9.8) are called **homographic transformations**. A useful property of such transformations is that they can map an infinite region of the  $z$ -plane onto a finite region of the  $z'$ -plane. In fact, points with very large values of  $z$  are mapped onto a neighborhood of the point  $z' = a/c$ . Of course, this argument goes both ways: Equation (9.8) also maps a neighborhood of  $-d/c$  in the  $z$ -plane onto large regions of the  $z'$ -plane. The usefulness of homographic transformations is illustrated in the following example.

**9.3.3. Example.** Consider two cylindrical conductors of equal radius  $r$ , held at potentials  $u_1$  and  $u_2$ , respectively, whose centers are  $D$  units of length apart. Choose the  $x$ - and  $y$ -axes such that the centers of the cylinders are located on the  $x$ -axis at distances  $a_1$  and  $a_2$  from the origin, as shown in Figure 9.3. Let us find the electrostatic potential produced by such a configuration in the  $xy$ -plane.

We know from elementary electrostatics that the problem becomes very simple if the two cylinders are concentric (and, of course, of different radii). Thus, we try to map the two circles onto two concentric circles in the  $z'$ -plane such that the infinite region outside the two circles in the  $z$ -plane gets mapped onto the finite annular region between the two concentric circles in the  $z'$ -plane. We then (easily) find the potential in the  $z'$ -plane, and transfer it back to the  $z$ -plane.

The most general mapping that may be able to do the job is that given by Equation (9.8). However, it turns out that we do not have to be this general. In fact, the special case



**Figure 9.3** In the  $z$ -plane, we see two equal cylinders whose centers are separated.

$z' = 1/(z - c)$  in which  $c$  is a *real* constant will be sufficient. So,  $z = (1/z') + c$ , and the circles  $|z - a_k| = r$  for  $k = 1, 2$  will be mapped onto the circles  $|z' - a'_k| = r'_k$ , where (by Example 9.3.2)  $a'_k = (a_k - c)/[(a_k - c)^2 - r^2]$  and  $r'_k = r/|(a_k - c)^2 - r^2|$ .

Can we arrange the parameters so that the circles in the  $z'$ -plane are concentric, i.e., that  $a'_1 = a'_2$ ? The answer is yes. We set  $a'_1 = a'_2$  and solve for  $a_2$  in terms of  $a_1$ . The result is either the trivial solution  $a_2 = a_1$ , or  $a_2 = c - r^2/(a_1 - c)$ . If we place the origin of the  $z$ -plane at the center of the first cylinder, then  $a_1 = 0$  and  $a_2 = D = c + r^2/c$ . We can also find  $a'_1$  and  $a'_2$ :  $a'_1 = a'_2 \equiv a' = -c/(c^2 - r^2)$ , and the geometry of the problem is as shown in Figure 9.4.

For such a geometry the potential at a point in the annular region is given by  $\Phi' = A \ln \rho + B = A \ln |z' - a'| + B$ , where  $A$  and  $B$  are real constants determined by the conditions  $\Phi'(r'_1) = u_1$  and  $\Phi'(r'_2) = u_2$ , which yields

$$A = \frac{u_1 - u_2}{\ln(r'_1/r'_2)} \quad \text{and} \quad B = \frac{u_2 \ln r'_1 - u_1 \ln r'_2}{\ln(r'_1/r'_2)}.$$

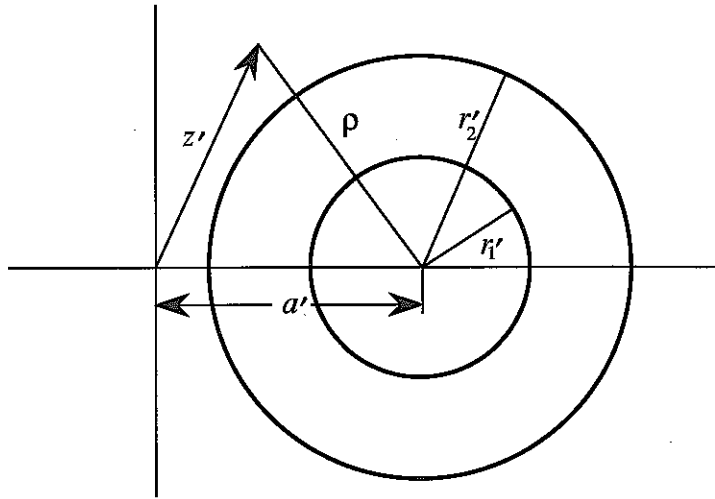
The potential  $\Phi'$  is the real part of the complex function<sup>8</sup>

$$F(z') = A \ln(z' - a') + B,$$

which is analytic except at  $z' = a'$ , a point lying outside the region of interest. We can now go back to the  $z$ -plane by substituting  $z' = 1/(z - c)$  to obtain

$$G(z) = A \ln \left( \frac{1}{z - c} - a' \right) + B,$$

<sup>8</sup>Writing  $z = |z|e^{i\theta}$ , we note that  $\ln z = \ln |z| + i\theta$ , so that the real part of a complex log function is the log of the absolute value.



**Figure 9.4** In the  $z'$ -plane, we see two concentric unequal cylinders.

whose real part is the potential in the  $z$ -plane:

$$\begin{aligned}\Phi(x, y) &= \operatorname{Re}[G(z)] = A \ln \left| \frac{1 - a'z + a'c}{z - c} \right| + B \\ &= A \ln \left| \frac{(1 + a'c - a'x) - ia'y}{(x - c) + iy} \right| + B \\ &= \frac{A}{2} \ln \left[ \frac{(1 + a'c - a'x)^2 + a'^2 y^2}{(x - c)^2 + y^2} \right] + B.\end{aligned}$$

This is the potential we want. ■

## 9.4 Integration of Complex Functions

The derivative of a complex function is an important concept and, as the previous section demonstrated, provides a powerful tool in physical applications. The concept of integration is even more important. In fact, we will see in the next section that derivatives can be written in terms of integrals. We will study integrals of complex functions in detail in this section.

The definite integral of a complex function is defined in analogy to that of a real function:

$$\int_{\alpha_1}^{\alpha_2} f(z) dz = \lim_{\substack{N \rightarrow \infty \\ \Delta z_i \rightarrow 0}} \sum_{i=1}^N f(z_i) \Delta z_i,$$

where  $\Delta z_i$  is a small segment, situated at  $z_i$ , of the curve that connects the complex number  $\alpha_1$  to the complex number  $\alpha_2$  in the  $z$ -plane. Since there are infinitely many ways of connecting  $\alpha_1$  to  $\alpha_2$ , it is possible to obtain different values for the integral for different paths.

One encounters a similar situation when one tries to evaluate the line integral of a vector field. In fact, we can turn the integral of a complex function into a line integral as follows. We substitute  $f(z) = u + iv$  and  $dz = dx + idy$  in the integral to obtain

$$\int_{\alpha_1}^{\alpha_2} f(z) dz = \int_{\alpha_1}^{\alpha_2} (u dx - v dy) + i \int_{\alpha_1}^{\alpha_2} (v dx + u dy).$$

If we define the two-dimensional vectors  $A_1 \equiv (u, -v)$  and  $A_2 \equiv (v, u)$ , then we get  $\int_{\alpha_1}^{\alpha_2} f(z) dz = \int_{\alpha_1}^{\alpha_2} A_1 \cdot dr + i \int_{\alpha_1}^{\alpha_2} A_2 \cdot dr$ . It follows from Stokes' theorem (or Green's theorem, since the vectors lie in a plane) that the integral of  $f$  is path-independent only if both  $A_1$  and  $A_2$  have vanishing curls. This in turn follows if and only if  $u$  and  $v$  satisfy the C-R conditions, and this is exactly what is needed for  $f(z)$  to be analytic.

Path-independence of a line integral of a vector  $A$  is equivalent to the vanishing of the integral along a closed path, and the latter is equivalent to the vanishing of  $\nabla \times A = 0$  at every point of the region bordered by the closed path. The preceding discussion is encapsulated in an important theorem, which we shall state shortly. First, however, it is worthwhile to become familiar with some terminology used frequently in complex analysis.

curve defined

1. A **curve** is a map  $\gamma : [a, b] \rightarrow \mathbb{C}$  from the real interval into the complex plane given by  $\gamma(t) = \gamma_r(t) + i\gamma_i(t)$ , where  $a \leq t \leq b$ , and  $\gamma_r$  and  $\gamma_i$  are the real and imaginary parts of  $\gamma$ ;  $\gamma(a)$  is called the **initial point** of the curve and  $\gamma(b)$  its **final point**.
2. A **simple arc**, or a Jordan arc, is a curve that does not cross itself, i.e.,  $\gamma$  is injective (or one to one), so that  $\gamma(t_1) \neq \gamma(t_2)$  when  $t_1 \neq t_2$ .
3. A **path** is a finite collection  $\{\gamma_1, \gamma_2, \dots, \gamma_n\}$  of simple arcs such that the initial point of  $\gamma_{k+1}$  coincides with the final point of  $\gamma_k$ .
4. A **smooth arc** is a curve for which  $d\gamma/dt = d\gamma_r/dt + id\gamma_i/dt$  exists and is nonzero for  $t \in [a, b]$ .

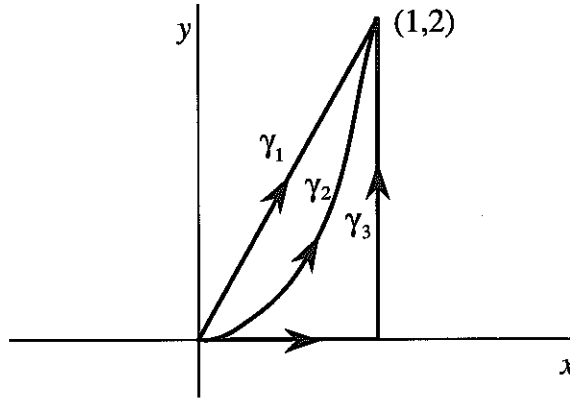
contour defined

5. A **contour** is a path whose arcs are smooth. When the initial point of  $\gamma_1$  coincides with the final point of  $\gamma_n$ , the contour is said to be a **simple closed contour**.

Cauchy-Goursat theorem

**9.4.1. Theorem.** (Cauchy-Goursat theorem) *Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be analytic on a simple closed contour  $C$  and at all points inside  $C$ . Then*

$$\oint_C f(z) dz = 0.$$



**Figure 9.5** The three different paths of integration corresponding to the integrals  $I_1$ ,  $I'_1$ ,  $I_2$ , and  $I'_2$ .

#### 9.4.2. Example. EXAMPLES OF DEFINITE INTEGRALS

(a) Let us evaluate the integral  $I_1 = \int_{\gamma_1} z dz$  where  $\gamma_1$  is the straight line drawn from the origin to the point  $(1, 2)$  (see Figure 9.5). Along such a line  $y = 2x$  and, using  $t$  for  $x$ ,  $\gamma_1(t) = t + 2it$  where  $0 \leq t \leq 1$ ; so

$$I_1 = \int_{\gamma_1} z dz = \int_0^1 (t + 2it)(dt + 2idt) = \int_0^1 (-3tdt + 4itdt) = -\frac{3}{2} + 2i.$$

For a different path  $\gamma_2$ , along which  $y = 2x^2$ , we get  $\gamma_2(t) = t + 2it^2$  where  $0 \leq t \leq 1$ , and

$$I'_1 = \int_{\gamma_2} z dz = \int_0^1 (t + 2it^2)(dt + 4itdt) = -\frac{3}{2} + 2i.$$

Therefore,  $I_1 = I'_1$ . This is what is expected from the Cauchy–Goursat theorem because the function  $f(z) = z$  is analytic on the two paths and in the region bounded by them.

(b) To find  $I_2 \equiv \int_{\gamma_1} z^2 dz$  with  $\gamma_1$  as in part (a), substitute for  $z$  in terms of  $t$ :

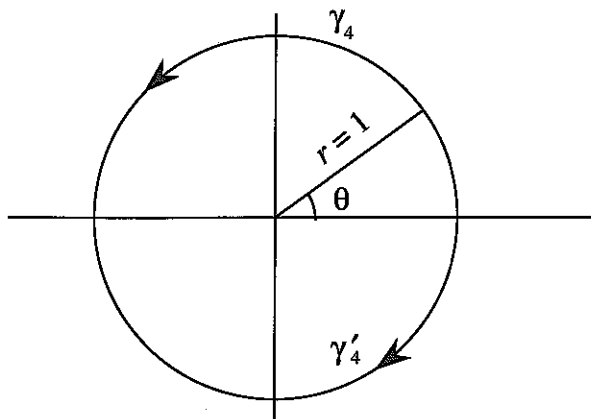
$$I_2 = \int_{\gamma_1} (t + 2it)^2(dt + 2idt) = (1 + 2i)^3 \int_0^1 t^2 dt = -\frac{11}{3} - \frac{2}{3}i.$$

Next we compare  $I_2$  with  $I'_2 = \int_{\gamma_3} z^2 dz$  where  $\gamma_3$  is as shown in Figure 9.5. This path can be described by

$$\gamma_3(t) = \begin{cases} t & \text{for } 0 \leq t \leq 1, \\ 1 + i(t - 1) & \text{for } 1 \leq t \leq 3. \end{cases}$$

Therefore,

$$I'_2 = \int_0^1 t^2 dt + \int_1^3 [1 + i(t - 1)]^2 (idt) = \frac{1}{3} - 4 - \frac{2}{3}i = -\frac{11}{3} - \frac{2}{3}i,$$



**Figure 9.6** The two semicircular paths for calculating  $I_3$  and  $I'_3$ .

which is identical to  $I_2$ , once again because the function is analytic on  $\gamma_1$  and  $\gamma_3$  as well as in the region bounded by them.

(c) Now consider  $I_3 \equiv \int_{\gamma_4} dz/z$  where  $\gamma_4$  is the upper semicircle of unit radius, as shown in Figure 9.6. A parametric equation for  $\gamma_4$  can be given in terms of  $\theta$ :

$$\gamma_4(\theta) = \cos \theta + i \sin \theta = e^{i\theta} \Rightarrow dz = ie^{i\theta} d\theta, \quad 0 \leq \theta \leq \pi.$$

Thus, we obtain

$$I_3 = \int_0^\pi \frac{1}{e^{i\theta}} ie^{i\theta} d\theta = i\pi.$$

On the other hand,

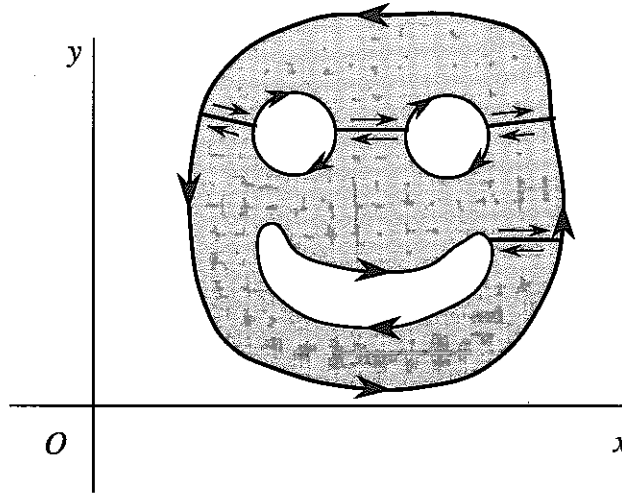
$$I'_3 = \int_{\gamma'_4} \frac{1}{z} dz = \int_{2\pi}^\pi \frac{1}{e^{i\theta}} ie^{i\theta} d\theta = -i\pi.$$

Here the two integrals are not equal. From  $\gamma_4$  and  $\gamma'_4$  we can construct a counterclockwise simple closed contour  $C$ , along which the integral of  $f(z) = 1/z$  becomes  $\oint_C dz/z = I_3 - I'_3 = 2i\pi$ . That the integral is not zero is a consequence of the fact that  $1/z$  is *not* analytic at all points of the region bounded by the closed contour  $C$ . ■

The Cauchy–Goursat theorem applies to more complicated regions. When a region contains points at which  $f(z)$  is not analytic, those points can be avoided by redefining the region and the contour. Such a procedure requires an agreement on the direction we will take.

convention for  
positive sense of  
integration around a  
closed contour

**Convention.** When integrating along a closed contour, we agree to move along the contour in such a way that the enclosed region lies to our left. An integration that follows this convention is called integration in the *positive sense*. Integration performed in the opposite direction acquires a minus sign.



**Figure 9.7** A complicated contour can be broken up into simpler ones. Note that the boundaries of the “eyes” and the “mouth” are forced to be traversed in the (negative) clockwise direction.

For a simple closed contour, movement in the counterclockwise direction yields integration in the positive sense. However, as the contour becomes more complicated, this conclusion breaks down. Figure 9.7 shows a complicated path enclosing a region (shaded) in which the integrand is analytic. Note that it is possible to traverse a portion of the region twice in opposite directions without affecting the integral, which may be a sum of integrals for different pieces of the contour. Also note that the “eyes” and the “mouth” are traversed clockwise! This is necessary because of the convention above. A region such as that shown in Figure 9.7, in which holes are “punched out,” is called **multiply connected**. In contrast, a **simply connected** region is one in which every simple closed contour encloses only points of the region.

simply and multiply connected regions

One important consequence of the Cauchy–Goursat theorem is the following:

Cauchy Integral Formula

**9.4.3. Theorem.** (Cauchy integral formula) *Let  $f$  be analytic on and within a simple closed contour  $C$  integrated in the positive sense. Let  $z_0$  be any interior point to  $C$ . Then*

$$f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0} dz$$

To prove the Cauchy integral formula (CIF), we need the following lemma.

**9.4.4. Lemma.** (Darboux inequality) *Suppose  $f : \mathbb{C} \rightarrow \mathbb{C}$  is continuous and bounded on a path  $\gamma$ , i.e., there exists a positive number  $M$  such that  $|f(z)| \leq M$*

for all values  $z \in \gamma$ . Then

$$\left| \int_{\gamma} f(z) dz \right| \leq ML_{\gamma},$$

where  $L_{\gamma}$  is the length of the path of integration.

*Proof.*

$$\begin{aligned} \left| \int_{\gamma} f(z) dz \right| &= \left| \lim_{\substack{N \rightarrow \infty \\ \Delta z_i \rightarrow 0}} \sum_{i=1}^N f(z_i) \Delta z_i \right| = \lim_{\substack{N \rightarrow \infty \\ \Delta z_i \rightarrow 0}} \left| \sum_{i=1}^N f(z_i) \Delta z_i \right| \\ &\leq \lim_{\substack{N \rightarrow \infty \\ \Delta z_i \rightarrow 0}} \sum_{i=1}^N |f(z_i) \Delta z_i| = \lim_{\substack{N \rightarrow \infty \\ \Delta z_i \rightarrow 0}} \sum_{i=1}^N |f(z_i)| |\Delta z_i| \\ &\leq M \lim_{\substack{N \rightarrow \infty \\ \Delta z_i \rightarrow 0}} \sum_{i=1}^N |\Delta z_i| = ML_{\gamma}. \end{aligned}$$

The first inequality follows from the triangle inequality, the second from the boundedness of  $f$ , and the last equality follows from the definition of the length of a path.  $\square$

Now we are ready to prove the Cauchy integral formula.

*Proof of CIF.* Consider the shaded region in Figure 9.8, which is bounded by  $C$ , by  $\gamma_0$  (a circle of arbitrarily small radius  $\delta$  centered at  $z_0$ ), and by  $L_1$  and  $L_2$ , two straight line segments infinitesimally close to one another (we can, in fact, assume that  $L_1$  and  $L_2$  are right on top of one another; however, they are separated in the figure for clarity). Let us use  $C'$  to denote the union of all these curves.

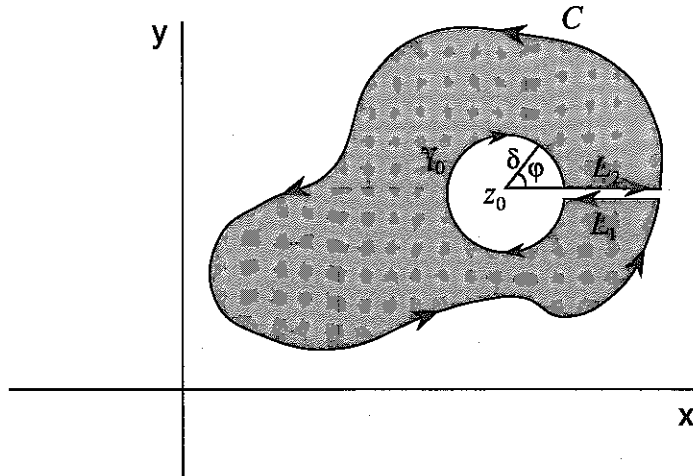
Since  $f(z)/(z - z_0)$  is analytic everywhere on the contour  $C'$  and inside the shaded region, we can write

$$\begin{aligned} 0 &= \frac{1}{2\pi i} \oint_{C'} \frac{f(z)}{z - z_0} dz \tag{9.10} \\ &= \frac{1}{2\pi i} \left[ \oint_C \frac{f(z)}{z - z_0} dz + \oint_{\gamma_0} \frac{f(z)}{z - z_0} dz + \underbrace{\oint_{L_1} \frac{f(z)}{z - z_0} dz + \oint_{L_2} \frac{f(z)}{z - z_0} dz}_{=0} \right]. \end{aligned}$$

The contributions from  $L_1$  and  $L_2$  cancel because they are integrals along the same line segment in opposite directions. Let us evaluate the contribution from the infinitesimal circle  $\gamma_0$ . First we note that because  $f(z)$  is continuous (differentiability implies continuity), we can write

$$\left| \frac{f(z) - f(z_0)}{z - z_0} \right| = \frac{|f(z) - f(z_0)|}{|z - z_0|} = \frac{|f(z) - f(z_0)|}{\delta} < \frac{\epsilon}{\delta}$$





**Figure 9.8** The integrand is analytic within and on the boundary of the shaded region. It is always possible to construct contours that exclude all singular points.

for  $z \in \gamma_0$ , where  $\epsilon$  is a small positive number. We now apply the Darboux inequality and write

$$\left| \oint_{\gamma_0} \frac{f(z) - f(z_0)}{z - z_0} dz \right| < \frac{\epsilon}{\delta} 2\pi\delta = 2\pi\epsilon.$$

This means that the integral goes to zero as  $\delta \rightarrow 0$ , or

$$\oint_{\gamma_0} \frac{f(z)}{z - z_0} dz = \oint_{\gamma_0} \frac{f(z_0)}{z - z_0} dz = f(z_0) \oint_{\gamma_0} \frac{dz}{z - z_0}.$$

We can easily calculate the integral on the RHS by noting that  $z - z_0 = \delta e^{i\varphi}$  and that  $\gamma_0$  has a *clockwise* direction:

$$\oint_{\gamma_0} \frac{dz}{z - z_0} = - \int_0^{2\pi} \frac{i\delta e^{i\varphi} d\varphi}{\delta e^{i\varphi}} = -2\pi i \Rightarrow \oint_{\gamma_0} \frac{f(z)}{z - z_0} dz = -2\pi i f(z_0).$$

Substituting this in (9.10) yields the desired result.  $\square$

**9.4.5. Example.** We can use the CIF to evaluate the integrals

$$I_1 = \oint_{C_1} \frac{z^2 dz}{(z^2 + 3)^2(z - i)}, \quad I_2 = \oint_{C_2} \frac{(z^2 - 1) dz}{(z - \frac{1}{2})(z^2 - 4)^3},$$

$$I_3 = \oint_{C_3} \frac{e^{z/2} dz}{(z - i\pi)(z^2 - 20)^4},$$

where  $C_1, C_2,$  and  $C_3$  are circles centered at the origin with radii  $r_1 = 3/2, r_2 = 1,$  and  $r_3 = 4.$

For  $I_1$  we note that  $f(z) = z^2/(z^2 + 3)^2$  is analytic within and on  $C_1,$  and  $z_0 = i$  lies in the interior of  $C_1.$  Thus,

$$I_1 = \oint_{C_1} \frac{f(z)dz}{z - i} = 2\pi i f(i) = 2\pi i \frac{i^2}{(i^2 + 3)^2} = -i \frac{\pi}{2}.$$

Similarly,  $f(z) = (z^2 - 1)/(z^2 - 4)^3$  for the integral  $I_2$  is analytic on and within  $C_2,$  and  $z_0 = 1/2$  is an interior point of  $C_2.$  Thus, the CIF gives

$$I_2 = \oint_{C_2} \frac{f(z)dz}{z - \frac{1}{2}} = 2\pi i f(1/2) = \frac{32\pi}{1125} i.$$

For the last integral,  $f(z) = e^{z/2}/(z^2 - 20)^4,$  and the interior point is  $z_0 = i\pi:$

$$I_3 = \oint_{C_3} \frac{f(z)dz}{z - i\pi} = 2\pi i f(i\pi) = -\frac{2\pi}{(\pi^2 + 20)^4}.$$

The Cauchy integral formula gives the value of an analytic function at every point inside a simple closed contour when it is given the value of the function only at points on the contour. It seems as though an analytic function is not free to change inside a region once its value is fixed on the contour enclosing that region.

Explanation of why the Cauchy integral formula works!

There is an analogous situation in electrostatics: The specification of the potential at the boundaries, such as the surfaces of conductors, automatically determines the potential at any other point in the region of space bounded by the conductors. This is the content of the uniqueness theorem used in electrostatic boundary value problems. However, the electrostatic potential  $\Phi$  is bound by another condition, Laplace's equation; and the combination of Laplace's equation and the boundary conditions furnishes the uniqueness of  $\Phi.$  Similarly, the real and imaginary parts of an analytic function separately satisfy Laplace's equation in two dimensions! Thus, it should come as no surprise that the value of an analytic function on a boundary (contour) determines the function at all points inside the boundary.

## 9.5 Derivatives as Integrals

The Cauchy Integral Formula is a very powerful tool for working with analytic functions. One of the applications of this formula is in evaluating the derivatives of such functions. It is convenient to change the dummy integration variable to  $\xi$  and write the CIF as

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\xi) d\xi}{\xi - z}, \tag{9.11}$$

where  $C$  is a simple closed contour in the  $\xi$ -plane and  $z$  is a point within  $C.$  As preparation for defining the derivative of an analytic function, we need the following result.

**9.5.1. Proposition.** Let  $\gamma$  be any path—a contour, for example—and  $g$  a continuous function on that path. The function  $f(z)$  defined by

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{g(\xi) d\xi}{\xi - z}$$

is analytic at every point  $z \notin \gamma$ .

*Proof.* The proof follows immediately from differentiation of the integral:

$$\frac{df}{dz} = \frac{1}{2\pi i} \frac{d}{dz} \int_{\gamma} \frac{g(\xi) d\xi}{\xi - z} = \frac{1}{2\pi i} \int_{\gamma} g(\xi) d\xi \frac{d}{dz} \left( \frac{1}{\xi - z} \right) = \frac{1}{2\pi i} \int_{\gamma} \frac{g(\xi) d\xi}{(\xi - z)^2}.$$

This is defined for all values of  $z$  not on  $\gamma$ .<sup>9</sup> Thus,  $f(z)$  is analytic there. □

We can generalize the formula above to the  $n$ th derivative, and obtain

$$\frac{d^n f}{dz^n} = \frac{n!}{2\pi i} \int_{\gamma} \frac{g(\xi) d\xi}{(\xi - z)^{n+1}}.$$

Applying this result to an analytic function expressed by Equation (9.11), we obtain the following important theorem.

derivative of an analytic function given in terms of an integral

**9.5.2. Theorem.** The derivatives of all orders of an analytic function  $f(z)$  exist in the domain of analyticity of the function and are themselves analytic in that domain. The  $n$ th derivative of  $f(z)$  is given by

$$f^{(n)}(z) = \frac{d^n f}{dz^n} = \frac{n!}{2\pi i} \oint_C \frac{f(\xi) d\xi}{(\xi - z)^{n+1}} \tag{9.12}$$

**9.5.3. Example.** Let us apply Equation (9.12) directly to some simple functions. In all cases, we will assume that the contour is a circle of radius  $r$  centered at  $z$ .

(a) Let  $f(z) = K$ , a constant. Then, for  $n = 1$  we have

$$\frac{df}{dz} = \frac{1}{2\pi i} \oint_C \frac{K d\xi}{(\xi - z)^2}.$$

Since  $\xi$  is on the circle  $C$  centered at  $z$ ,  $\xi - z = re^{i\theta}$  and  $d\xi = rie^{i\theta} d\theta$ . So we have

$$\frac{df}{dz} = \frac{1}{2\pi i} \int_0^{2\pi} \frac{Kire^{i\theta} d\theta}{(re^{i\theta})^2} = \frac{K}{2\pi r} \int_0^{2\pi} e^{-i\theta} d\theta = 0.$$

(b) Given  $f(z) = z$ , its first derivative will be

$$\begin{aligned} \frac{df}{dz} &= \frac{1}{2\pi i} \oint_C \frac{\xi d\xi}{(\xi - z)^2} = \frac{1}{2\pi i} \int_0^{2\pi} \frac{(z + re^{i\theta})ire^{i\theta} d\theta}{(re^{i\theta})^2} \\ &= \frac{1}{2\pi} \left( \frac{z}{r} \int_0^{2\pi} e^{-i\theta} d\theta + \int_0^{2\pi} d\theta \right) = \frac{1}{2\pi} (0 + 2\pi) = 1. \end{aligned}$$

<sup>9</sup>The interchange of differentiation and integration requires justification. Such an interchange can be done if the integral has some restrictive properties. We shall not concern ourselves with such details. In fact, one can achieve the same result by using the definition of derivatives and the usual properties of integrals.

(c) Given  $f(z) = z^2$ , for the first derivative Equation (9.12) yields

$$\begin{aligned} \frac{df}{dz} &= \frac{1}{2\pi i} \oint_C \frac{\xi^2 d\xi}{(\xi - z)^2} = \frac{1}{2\pi i} \int_0^{2\pi} \frac{(z + re^{i\theta})^2 ire^{i\theta} d\theta}{(re^{i\theta})^2} \\ &= \frac{1}{2\pi} \int_0^{2\pi} [z^2 + (re^{i\theta})^2 + 2zre^{i\theta}] (re^{i\theta})^{-1} d\theta \\ &= \frac{1}{2\pi} \left( \frac{z^2}{r} \int_0^{2\pi} e^{-i\theta} d\theta + r \int_0^{2\pi} e^{i\theta} d\theta + 2z \int_0^{2\pi} d\theta \right) = 2z. \end{aligned}$$

It can be shown that, in general,  $(d/dz)z^m = mz^{m-1}$ . The proof is left as Problem 9.30.  $\blacksquare$

The CIF is a central formula in complex analysis, and we shall see its significance in much of the later development of complex analysis. For now, let us demonstrate its usefulness in proving a couple of important properties of analytic functions.

**9.5.4. Proposition.** *The absolute value of an analytic function  $f(z)$  cannot have a local maximum within the region of analyticity of the function.*

*Proof.* Let  $S \subset \mathbb{C}$  be the region of analyticity of  $f$ . Suppose  $z_0 \in S$  were a local maximum. Then we could find a circle  $\gamma_0$  of small enough radius  $\delta$ , centered at  $z_0$ , such that  $|f(z_0)| > |f(z)|$  for all  $z$  on  $\gamma_0$ . We now show that this cannot happen. Using the CIF, and noting that  $z - z_0 = \delta e^{i\theta}$ , we have

$$\begin{aligned} |f(z_0)| &= \left| \frac{1}{2\pi i} \oint_{\gamma_0} \frac{f(z)}{z - z_0} dz \right| = \frac{1}{2\pi} \left| \int_0^{2\pi} \frac{f(z)}{\delta e^{i\theta}} i\delta e^{i\theta} d\theta \right| \\ &\leq \frac{1}{2\pi} \int_0^{2\pi} |f(z)| d\theta \leq \frac{1}{2\pi} \int_0^{2\pi} M d\theta = M, \end{aligned}$$

where  $M$  is the maximum value of  $|f(z)|$  for  $z \in \gamma_0$ . This inequality says that there is at least one point  $z$  on the circle  $\gamma_0$  (the point at which the maximum of  $|f(z)|$  is attained) such that  $|f(z_0)| \leq |f(z)|$ . This contradicts our assumption. Therefore, there can be no local maximum within  $S$ .  $\square$

**9.5.5. Proposition.** *A bounded entire function is necessarily a constant.*

*Proof.* We show that the derivative of such a function is zero. Consider

$$\frac{df}{dz} = \frac{1}{2\pi i} \oint_C \frac{f(\xi) d\xi}{(\xi - z)^2}.$$

Since  $f$  is an entire function, the closed contour  $C$  can be chosen to be a very large circle of radius  $R$  with center at  $z$ . Taking the absolute value of both sides yields

$$\begin{aligned} \left| \frac{df}{dz} \right| &= \frac{1}{2\pi} \left| \int_0^{2\pi} \frac{f(z)}{(Re^{i\theta})^2} iRe^{i\theta} d\theta \right| \\ &\leq \frac{1}{2\pi} \int_0^{2\pi} \frac{|f(z)|}{R} d\theta \leq \frac{1}{2\pi} \int_0^{2\pi} \frac{M}{R} d\theta = \frac{M}{R}, \end{aligned}$$

where  $M$  is the maximum of the function in the complex plane. Now, as  $R \rightarrow \infty$ , the derivative goes to zero, and the function must be a constant.  $\square$

Proposition 9.5.5 is a very powerful statement about analytic functions. There are many interesting and nontrivial *real* functions that are bounded and have derivatives of all orders on the entire real line. For instance,  $e^{-x^2}$  is such a function. No such freedom exists for *complex* analytic functions. Any nontrivial analytic function is either not bounded (goes to infinity somewhere on the complex plane) or not entire (it is not analytic at some point(s) of the complex plane).

fundamental theorem  
of algebra proved

A consequence of Proposition 9.5.5 is the **fundamental theorem of algebra**, which states that any polynomial of degree  $n \geq 1$  has  $n$  roots (some of which may be repeated). In other words, the polynomial

$$p(x) = a_0 + a_1x + \cdots + a_nx^n \quad \text{for } n \geq 1$$

can be factored completely as  $p(x) = c(x - z_1)(x - z_2) \cdots (x - z_n)$  where  $c$  is a constant and the  $z_i$  are, in general, complex numbers.

To see how Proposition 9.5.5 implies the fundamental theorem of algebra, we let  $f(z) = 1/p(z)$  and assume the contrary, i.e., that  $p(z)$  is never zero for any (finite)  $z \in \mathbb{C}$ . Then  $f(z)$  is bounded and analytic for all  $z \in \mathbb{C}$ , and Proposition 9.5.5 says that  $f(z)$  is a constant. This is obviously wrong. Thus, there must be at least one  $z$ , say  $z = z_1$ , for which  $p(z)$  is zero. So, we can factor out  $(z - z_1)$  from  $p(z)$  and write  $p(z) = (z - z_1)q(z)$  where  $q(z)$  is of degree  $n - 1$ . Applying the above argument to  $q(z)$ , we have  $p(z) = (z - z_1)(z - z_2)r(z)$  where  $r(z)$  is of degree  $n - 2$ . Continuing in this way, we can factor  $p(z)$  into linear factors. The last polynomial will be a constant (a polynomial of degree zero) which we have denoted by  $c$ .

The primitive (indefinite integral) of an analytic function can be defined using definite integrals just as in the real case. Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be analytic in a region  $S$  of the complex plane. Let  $z_0$  and  $z$  be two points in  $S$ , and define<sup>10</sup>  $F(z) \equiv \int_{z_0}^z f(\xi) d\xi$ . We can show that  $F(z)$  is the primitive of  $f(z)$  by showing that

$$\lim_{\Delta z \rightarrow 0} \left| \frac{F(z + \Delta z) - F(z)}{\Delta z} - f(z) \right| = 0.$$

We leave the details as a problem for the reader.

**9.5.6. Proposition.** *Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be analytic in a region  $S$  of  $\mathbb{C}$ . Then at every point  $z \in S$ , there exists an analytic function  $F : \mathbb{C} \rightarrow \mathbb{C}$  such that*

$$\frac{dF}{dz} = f(z).$$

<sup>10</sup>Note that the integral is path-independent due to the analyticity of  $f$ . Thus,  $F$  is well-defined.

In the sketch of the proof of Proposition 9.5.6, we used only the continuity of  $f$  and the fact that the integral was well-defined. These two conditions are sufficient to establish the analyticity of  $F$  and  $f$ , since the latter is the derivative of the former. The following theorem, due to Morera, states this fact and is the converse of the Cauchy–Goursat theorem.

**9.5.7. Theorem.** (Morera's theorem) *Let a function  $f : \mathbb{C} \rightarrow \mathbb{C}$  be continuous in a simply connected region  $S$ . If for each simple closed contour  $C$  in  $S$  we have  $\oint_C f(\xi) d\xi = 0$ , then  $f$  is analytic throughout  $S$ .*

## 9.6 Taylor and Laurent Series

The expansion of functions in terms of polynomials or monomials is important in calculus and was emphasized in the analysis of Chapter 5. We now apply this concept to analytic functions.

### 9.6.1 Properties of Series

The reader is assumed to have some familiarity with complex series. Nevertheless, we state (without proof) the most important properties of complex series before discussing Taylor and Laurent series.

absolute convergence A complex series is said to **converge absolutely** if the *real* series  $\sum_{k=0}^{\infty} |z_k| = \sum_{k=0}^{\infty} \sqrt{x_k^2 + y_k^2}$  converges. Clearly, absolute convergence implies convergence.

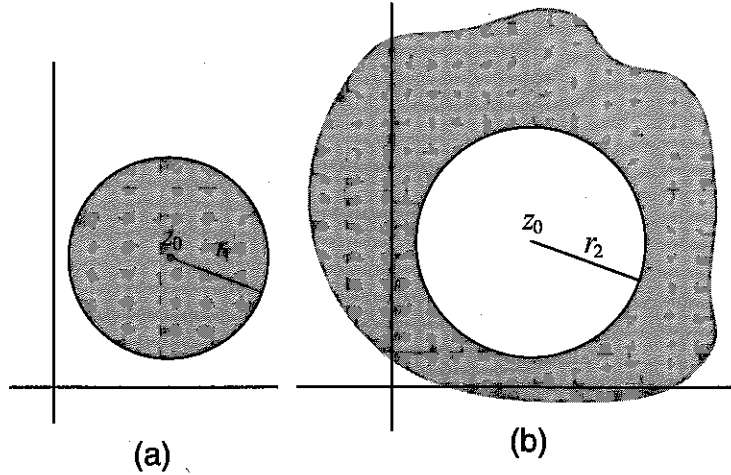
power series **9.6.1. Proposition.** *If the power series  $\sum_{k=0}^{\infty} a_k(z - z_0)^k$  converges for  $z_1 \neq z_0$ , then it converges absolutely for every value of  $z$  such that  $|z - z_0| < |z_1 - z_0|$ . Similarly if the power series  $\sum_{k=0}^{\infty} b_k/(z - z_0)^k$  converges for  $z_2 \neq z_0$ , then it converges absolutely for every value of  $z$  such that  $|z - z_0| > |z_2 - z_0|$ .*

A geometric interpretation of this proposition is that if a power series—with positive powers—converges for a point at a distance  $r_1$  from  $z_0$ , then it converges for *all interior* points of the circle whose center is  $z_0$ , and whose radius is  $r_1$ . Similarly, if a power series—with negative powers—converges for a point at a distance  $r_2$  from  $z_0$ , then it converges for *all exterior* points of the circle whose center is  $z_0$  and whose radius is  $r_2$  (see Figure 9.9). Generally speaking, positive powers are used for points inside a circle and negative powers for points outside it.

circle of convergence The largest circle about  $z_0$  such that the first power series of Proposition 9.6.1 converges is called the **circle of convergence** of the power series. The proposition implies that the series cannot converge at *any* point outside the circle of convergence. (Why?)

In determining the convergence of a power series

$$S(z) \equiv \sum_{n=0}^{\infty} a_n(z - z_0)^n, \quad (9.13)$$



**Figure 9.9** (a) Power series with positive exponents converge for the interior points of a circle. (b) Power series with negative exponents converge for the exterior points of a circle.

we look at the behavior of the sequence of partial sums

$$S_N(z) \equiv \sum_{n=0}^N a_n(z - z_0)^n.$$

Convergence of (9.13) implies that for any  $\epsilon > 0$ , there exists an integer  $N_\epsilon$  such that

$$|S(z) - S_N(z)| < \epsilon \quad \text{whenever } N > N_\epsilon.$$

uniform convergence explained

In general, the integer  $N_\epsilon$  may be dependent on  $z$ ; that is, for different values of  $z$ , we may be forced to pick different  $N_\epsilon$ 's. When  $N_\epsilon$  is independent of  $z$ , we say that the convergence is **uniform**.

power series are uniformly convergent and analytic

**9.6.2. Theorem.** *The power series  $S(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$  is **uniformly convergent** for all points within its circle of convergence and represents a function that is analytic there.*

By substituting the reciprocal of  $(z - z_0)$  in the power series, we can show that if  $\sum_{k=0}^{\infty} b_k/(z - z_0)^k$  is convergent in the annulus  $r_2 < |z - z_0| < r_1$ , then it is uniformly convergent for all  $z$  in that annulus.

power series can be differentiated and integrated term by term

**9.6.3. Theorem.** *A convergent power series can be differentiated and integrated term by term; that is, if  $S(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ , then*

$$\frac{dS(z)}{dz} = \sum_{n=1}^{\infty} n a_n(z - z_0)^{n-1}, \quad \int_{\gamma} S(z) dz = \sum_{n=0}^{\infty} a_n \int_{\gamma} (z - z_0)^n dz$$

for any path  $\gamma$  lying in the circle of convergence of the power series.

## 9.6.2 Taylor and Laurent Series

We now state and prove the two main theorems of this section. A Taylor series consists of terms with only positive powers. A Laurent series allows for negative powers as well.

**9.6.4. Theorem.** (Taylor series) *Let  $f$  be analytic throughout the interior of a circle  $C_0$  having radius  $r_0$  and centered at  $z_0$ . Then at each point  $z$  inside  $C_0$ ,*

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n. \quad (9.14)$$

*That is, the power series converges to  $f(z)$  when  $|z - z_0| < r_0$ .*

*Proof.* From the CIF and the fact that  $z$  is inside  $C_0$ , we have

$$f(z) = \frac{1}{2\pi i} \oint_{C_0} \frac{f(\xi)}{\xi - z} d\xi.$$

On the other hand,

$$\begin{aligned} \frac{1}{\xi - z} &= \frac{1}{\xi - z_0 + z_0 - z} = \frac{1}{(\xi - z_0) \left(1 - \frac{z - z_0}{\xi - z_0}\right)} \\ &= \frac{1}{\xi - z_0} \frac{1}{1 - \frac{z - z_0}{\xi - z_0}} = \frac{1}{\xi - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{\xi - z_0}\right)^n. \end{aligned}$$

The last equality follows from the fact that  $|(z - z_0)/(\xi - z_0)| < 1$ —because  $z$  is inside the circle  $C_0$  and  $\xi$  is on it—and from the sum of a geometric series. Substituting in the CIF and using Theorem 9.5.2, we obtain the result.  $\square$

**Maclaurin series** For  $z_0 = 0$  we obtain the **Maclaurin series**:

$$f(z) = f(0) + f'(0)z + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n$$

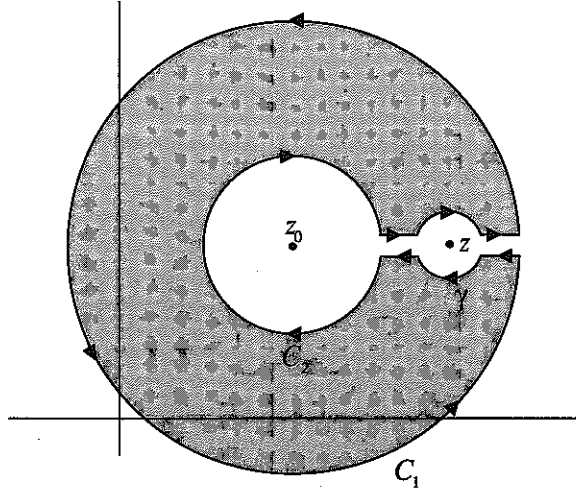
The Taylor expansion requires analyticity of the function at all points interior to the circle  $C_0$ . On many occasions there may be a point inside  $C_0$  at which the function is not analytic. The Laurent series accommodates such cases.

**9.6.5. Theorem.** (Laurent series) *Let  $C_1$  and  $C_2$  be circles of radii  $r_1$  and  $r_2$ , both centered at  $z_0$  in the  $z$ -plane with  $r_1 > r_2$ . Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be analytic on  $C_1$  and  $C_2$  and throughout  $S$ , the annular region between the two circles. Then, at each point  $z \in S$ ,  $f(z)$  is given by*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \quad \text{where} \quad a_n = \frac{1}{2\pi i} \oint_C \frac{f(\xi)}{(\xi - z_0)^{n+1}} d\xi$$

*and  $C$  is any contour within  $S$  that encircles  $z_0$ .*





**Figure 9.10** The annular region within and on whose contour the expanded function is analytic.

*Proof.* Let  $\gamma$  be a small closed contour in  $S$  enclosing  $z$ , as shown in Figure 9.10. For the composite contour  $C'$  the Cauchy–Goursat theorem gives

$$0 = \oint_{C'} \frac{f(\xi)}{\xi - z} d\xi = \oint_{C_1} \frac{f(\xi)}{\xi - z} d\xi - \oint_{C_2} \frac{f(\xi)}{\xi - z} d\xi - \oint_{\gamma} \frac{f(\xi)}{\xi - z} d\xi,$$

where the  $\gamma$  and  $C_2$  integrations are negative because their interior lies to our right as we traverse them. The  $\gamma$  integral is simply  $2\pi i f(z)$  by the CIF. Thus, we obtain

$$2\pi i f(z) = \oint_{C_1} \frac{f(\xi)}{\xi - z} d\xi - \oint_{C_2} \frac{f(\xi)}{\xi - z} d\xi. \tag{9.15}$$

Now we use the same trick we used in deriving the Taylor expansion. Since  $z$  is located in the annular region,  $r_2 < |z - z_0| < r_1$ . We have to keep this in mind when expanding the fractions. In particular, for  $\xi \in C_1$  we want the  $\xi$  term in the denominator, and for  $\xi \in C_2$  we want it in the numerator. Substituting such expansions in Equation (9.15) yields

$$\begin{aligned} 2\pi i f(z) &= \sum_{n=0}^{\infty} (z - z_0)^n \oint_{C_1} \frac{f(\xi) d\xi}{(\xi - z_0)^{n+1}} \\ &\quad + \sum_{n=0}^{\infty} \frac{1}{(z - z_0)^{n+1}} \oint_{C_2} f(\xi) (\xi - z_0)^n d\xi. \end{aligned} \tag{9.16}$$

Now we consider an arbitrary contour  $C$  in  $S$  that encircles  $z_0$ . Figure 9.11 shows a region bounded by a contour composed of  $C_1$  and  $C$ . In this region

$f(\xi)/(\xi - z_0)^{n+1}$  is analytic (because  $\xi$  can never equal  $z_0$ ). Thus, the integral over the composite contour must vanish by the Cauchy–Goursat theorem. It follows that the integral over  $C_1$  is equal to that over  $C$ . A similar argument shows that the  $C_2$  integral can also be replaced by an integral over  $C$ . We let  $n + 1 = -m$  in the second sum of Equation (9.16) to transform it into

$$\sum_{m=-1}^{-\infty} \frac{1}{(z - z_0)^{-m}} \oint_C f(\xi)(\xi - z_0)^{-m-1} d\xi = \sum_{m=-\infty}^{-1} (z - z_0)^m \oint_C \frac{f(\xi) d\xi}{(\xi - z_0)^{m+1}}.$$

Changing the dummy index back to  $n$  and substituting the result in Equation (9.16) yields

$$2\pi i f(z) = \sum_{n=0}^{\infty} (z - z_0)^n \oint_C \frac{f(\xi)}{(\xi - z_0)^{n+1}} d\xi + \sum_{n=-\infty}^{-1} (z - z_0)^n \oint_C \frac{f(\xi)}{(\xi - z_0)^{n+1}} d\xi.$$

We can now combine the sums and divide both sides by  $2\pi i$  to get the desired expansion. □

The Laurent expansion is convergent as long as  $r_2 < |z - z_0| < r_1$ . In particular, if  $r_2 = 0$ , and if the function is analytic throughout the interior of the larger circle, then  $a_n$  will be zero for  $n = -1, -2, \dots$  because  $f(\xi)/(\xi - z_0)^{n+1}$  will be analytic for negative  $n$ , and the integral will be zero by the Cauchy–Goursat theorem. Thus, only positive powers of  $(z - z_0)$  will be present in the series, and we will recover the Taylor series, as we should.

It is clear that we can expand  $C_1$  and shrink  $C_2$  until we encounter a point at which  $f$  is no longer analytic. This is obvious from the construction of the proof, in which only the analyticity in the annular region is important, not its size. Thus, we can include all the possible analytic points by expanding  $C_1$  and shrinking  $C_2$ .

**9.6.6. Example.** Let us expand some functions in terms of series. For an entire function there is no point in the entire complex plane at which it is not analytic. Thus, only positive powers of  $(z - z_0)$  will be present, and we will have a Taylor expansion that is valid for all values of  $z$ .

(a) Let us expand  $e^z$  around  $z_0 = 0$ . The  $n$ th derivative of  $e^z$  is  $e^z$ . Thus,  $f^{(n)}(0) = 1$ , and Taylor (Maclaurin) expansion gives

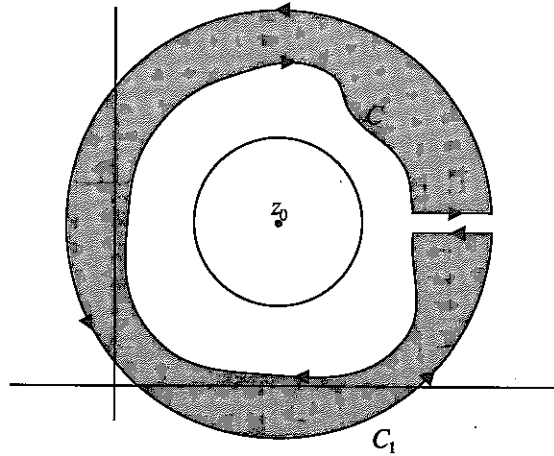
$$e^z = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

(b) The Maclaurin series for  $\sin z$  is obtained by noting that

$$\left. \frac{d^n}{dz^n} \sin z \right|_{z=0} = \begin{cases} 0 & \text{if } n \text{ is even,} \\ (-1)^{(n-1)/2} & \text{if } n \text{ is odd} \end{cases}$$

and substituting this in the Maclaurin expansion:

$$\sin z = \sum_{n \text{ odd}} (-1)^{(n-1)/2} \frac{z^n}{n!} = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!}.$$



**Figure 9.11** The arbitrary contour in the annular region used in the Laurent expansion.

Similarly, we can obtain

$$\cos z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}, \quad \sinh z = \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!}, \quad \cosh z = \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!}.$$

(c) The function  $1/(1+z)$  is not entire, so the region of its convergence is limited. Let us find the Maclaurin expansion of this function. The function is analytic within all circles of radii  $r < 1$ . At  $r = 1$  we encounter a singularity, the point  $z = -1$ . Thus, the series converges for all points<sup>11</sup>  $z$  for which  $|z| < 1$ . For such points we have

$$f^{(n)}(0) = \frac{d^n}{dz^n} [(1+z)^{-1}] \Big|_{z=0} = (-1)^n n!.$$

Thus,

$$\frac{1}{1+z} = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n = \sum_{n=0}^{\infty} (-1)^n z^n. \quad \blacksquare$$

Taylor and Laurent series allow us to express an analytic function as a power series. For a Taylor series of  $f(z)$ , the expansion is routine because the coefficient of its  $n$ th term is simply  $f^{(n)}(z_0)/n!$ , where  $z_0$  is the center of the circle of convergence. When a Laurent series is applicable, however, the  $n$ th coefficient is not, in general, easy to evaluate. Usually it can be found by inspection and certain manipulations of other known series. But if we use such an intuitive approach

<sup>11</sup>As remarked before, the series diverges for *all* points outside the circle  $|z| = 1$ . This does not mean that the function cannot be represented by a series for points outside the circle. On the contrary, we shall see shortly that Laurent series, with *negative* powers of  $z - z_0$  are designed precisely for such a purpose.

to determine the coefficients, can we be sure that we have obtained the correct Laurent series? The following theorem answers this question.

Laurent series is unique

**9.6.7. Theorem.** *If the series  $\sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$  converges to  $f(z)$  at all points in some annular region about  $z_0$ , then it is the unique Laurent series expansion of  $f(z)$  in that region.*

*Proof.* Multiply both sides of  $f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$  by

$$\frac{1}{2\pi i (z - z_0)^{k+1}},$$

integrate the result along a contour  $C$  in the annular region, and use the easily verifiable fact that

$$\frac{1}{2\pi i} \oint_C \frac{dz}{(z - z_0)^{k-n+1}} = \delta_{kn}$$

to obtain

$$\frac{1}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{k+1}} dz = a_k.$$

Thus, the coefficient in the power series of  $f$  is precisely the coefficient in the Laurent series, and the two must be identical.  $\square$

We will look at some examples that illustrate the abstract ideas developed in the preceding collection of theorems and propositions. However, we can consider a much broader range of examples if we know the arithmetic of power series. The following theorem giving arithmetical manipulations with power series is not difficult to prove (see [Chur 74]).

You can add, subtract, and multiply convergent power series

**9.6.8. Theorem.** *Let the two power series  $f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$  and  $g(z) = \sum_{n=-\infty}^{\infty} b_n(z - z_0)^n$  be convergent within some annular region  $r_2 < |z - z_0| < r_1$ . Then the sum  $\sum_{n=-\infty}^{\infty} (a_n + b_n)(z - z_0)^n$  converges to  $f(z) + g(z)$ , and the product*

$$\sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} a_n b_m (z - z_0)^{m+n} \equiv \sum_{k=-\infty}^{\infty} c_k (z - z_0)^k$$

*converges to  $f(z)g(z)$  for  $z$  interior to the annular region. Furthermore, if  $g(z) \neq 0$  for some neighborhood of  $z_0$ , then the series obtained by long division of  $\sum_{n=-\infty}^{\infty} a_n(z - z_0)^n$  by  $\sum_{m=-\infty}^{\infty} b_m(z - z_0)^m$  converges to  $f(z)/g(z)$  in that neighborhood.*

This theorem, in essence, says that converging power series can be manipulated as though they were finite sums (polynomials). Such manipulations are extremely useful when dealing with Taylor and Laurent expansions in which the straightforward calculation of coefficients may be tedious. The following examples illustrate the power of infinite-series arithmetic.

**9.6.9. Example.** To expand the function  $f(z) = \frac{2+3z}{z^2+z^3}$  in a Laurent series about  $z = 0$ , rewrite it as

$$\begin{aligned} f(z) &= \frac{1}{z^2} \left( \frac{2+3z}{1+z} \right) = \frac{1}{z^2} \left( 3 - \frac{1}{1+z} \right) = \frac{1}{z^2} \left( 3 - \sum_{n=0}^{\infty} (-1)^n z^n \right) \\ &= \frac{1}{z^2} (3 - 1 + z - z^2 + z^3 - \dots) = \frac{2}{z^2} + \frac{1}{z} - 1 + z - z^2 + \dots \end{aligned}$$

This series converges for  $0 < |z| < 1$ . We note that negative powers of  $z$  are also present.<sup>12</sup> Using the notation of Theorem 9.6.5, we have  $a_{-2} = 2, a_{-1} = 1, a_n = 0$  for  $n \leq -3$ , and  $a_n = (-1)^{n+1}$  for  $n \geq 0$ . ■

**9.6.10. Example.** The function  $f(z) = 1/(4z - z^2)$  is the ratio of two entire functions. Therefore, by Theorem 9.6.8, it is analytic everywhere except at the zeros of its denominator,  $z = 0$  and  $z = 4$ . For the annular region (here  $r_2$  of Theorem 9.6.5 is zero)  $0 < |z| < 4$ , we expand  $f(z)$  in the Laurent series around  $z = 0$ . Instead of actually calculating  $a_n$ , we first note that

$$f(z) = \frac{1}{4z} \left( \frac{1}{1 - z/4} \right).$$

The second factor can be expanded in a geometric series because  $|z/4| < 1$ :

$$\frac{1}{1 - z/4} = \sum_{n=0}^{\infty} \left( \frac{z}{4} \right)^n = \sum_{n=0}^{\infty} 4^{-n} z^n.$$

Dividing this by  $4z$ , and noting that  $z = 0$  is the only zero of  $4z$  and is excluded from the annular region, we obtain the expansion

$$f(z) = \sum_{n=0}^{\infty} 4^{-n} \frac{z^n}{4z} = \sum_{n=0}^{\infty} 4^{-n-1} z^{n-1}.$$

Although we derived this series using manipulations of other series, the uniqueness of series representations assures us that this is *the* Laurent series for the indicated region.

How can we represent  $f(z)$  in the region for which  $|z| > 4$ ? This region is exterior to the circle  $|z| = 4$ , so we expect negative powers of  $z$ . To find the Laurent expansion we write

$$f(z) = -\frac{1}{z^2} \left( \frac{1}{1 - 4/z} \right)$$

and note that  $|4/z| < 1$  for points exterior to the larger circle. The second factor can be written as a geometric series:

$$\frac{1}{1 - 4/z} = \sum_{n=0}^{\infty} \left( \frac{4}{z} \right)^n = \sum_{n=0}^{\infty} 4^n z^{-n}.$$

<sup>12</sup>This is a reflection of the fact that the function is not analytic inside the entire circle  $|z| = 1$ ; it blows up at  $z = 0$ .

Dividing by  $-z^2$ , which is nonzero in the region exterior to the larger circle, yields

$$f(z) = -\sum_{n=0}^{\infty} 4^n z^{-n-2}. \quad \blacksquare$$

**9.6.11. Example.** The function  $f(z) = z/[(z-1)(z-2)]$  has a Taylor expansion around the origin for  $|z| < 1$ . To find this expansion, we write<sup>13</sup>

$$f(z) = -\frac{1}{z-1} + \frac{2}{z-2} = \frac{1}{1-z} - \frac{1}{1-z/2}.$$

Expanding both fractions in geometric series (both  $|z|$  and  $|z/2|$  are less than 1), we obtain  $f(z) = \sum_{n=0}^{\infty} z^n - \sum_{n=0}^{\infty} (z/2)^n$ . Adding the two series—using Theorem 9.6.8—yields

$$f(z) = \sum_{n=0}^{\infty} (1 - 2^{-n})z^n \quad \text{for } |z| < 1.$$

This is the unique Taylor expansion of  $f(z)$  within the circle  $|z| = 1$ .

For  $1 < |z| < 2$  we have a Laurent series. To obtain this series, write

$$f(z) = \frac{1/z}{1/z-1} - \frac{1}{1-z/2} = -\frac{1}{z} \left( \frac{1}{1-1/z} \right) - \frac{1}{1-z/2}.$$

Since both fractions on the RHS converge in the annular region ( $|1/z| < 1$ ,  $|z/2| < 1$ ), we get

$$\begin{aligned} f(z) &= -\frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{1}{z}\right)^n - \sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^n = -\sum_{n=0}^{\infty} z^{-n-1} - \sum_{n=0}^{\infty} 2^{-n} z^n \\ &= -\sum_{n=-1}^{-\infty} z^n - \sum_{n=0}^{\infty} 2^{-n} z^n = \sum_{n=-\infty}^{\infty} a_n z^n, \end{aligned}$$

where  $a_n = -1$  for  $n < 0$  and  $a_n = -2^{-n}$  for  $n \geq 0$ . This is the unique Laurent expansion of  $f(z)$  in the given region.

Finally, for  $|z| > 2$  we have only negative powers of  $z$ . We obtain the expansion in this region by rewriting  $f(z)$  as follows:

$$f(z) = -\frac{1/z}{1-1/z} + \frac{2/z}{1-2/z}.$$

Expanding the fractions yields

$$f(z) = -\sum_{n=0}^{\infty} z^{-n-1} + \sum_{n=0}^{\infty} 2^{n+1} z^{-n-1} = \sum_{n=0}^{\infty} (2^{n+1} - 1)z^{-n-1}.$$

This is again the unique expansion of  $f(z)$  in the region  $|z| > 2$ . \blacksquare

<sup>13</sup>We could, of course, evaluate the derivatives of all orders of the function at  $z = 0$  and use Maclaurin's formula. However, the present method gives the same result much more quickly.

**9.6.12. Example.** Define  $f(z)$  as

$$f(z) = \begin{cases} (1 - \cos z)/z^2 & \text{for } z \neq 0, \\ \frac{1}{2} & \text{for } z = 0. \end{cases}$$

We can show that  $f(z)$  is an entire function.

Since  $1 - \cos z$  and  $z^2$  are entire functions, their ratio is analytic everywhere except at the zeros of its denominator. The only such zero is  $z = 0$ . Thus, Theorem 9.6.8 implies that  $f(z)$  is analytic everywhere except possibly at  $z = 0$ . To see the behavior of  $f(z)$  at  $z = 0$ , we look at its Maclaurin series:

$$1 - \cos z = 1 - \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!},$$

which implies that

$$\frac{1 - \cos z}{z^2} = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{z^{2n-2}}{(2n)!} = \frac{1}{2} - \frac{z^2}{4!} + \frac{z^4}{6!} - \dots.$$

The expansion on the RHS shows that the value of the series at  $z = 0$  is  $\frac{1}{2}$ , which, by definition, is  $f(0)$ . Thus, the series converges for all  $z$ , and Theorem 9.6.2 says that  $f(z)$  is entire. ■

A Laurent series can give information about the integral of a function around a closed contour in whose interior the function may not be analytic. In fact, the coefficient of the first negative power in a Laurent series is given by

$$a_{-1} = \frac{1}{2\pi i} \oint_C f(\xi) d\xi. \quad (9.17)$$

Thus, to find the integral of a (nonanalytic) function around a closed contour surrounding  $z_0$ , we write the Laurent series for the function and read off the coefficient of the  $1/(z - z_0)$  term.

**9.6.13. Example.** As an illustration of this idea, let us evaluate the integral  $I = \oint_C dz/[z^2(z-2)]$ , where  $C$  is the circle of radius 1 centered at the origin. The function is analytic in the annular region  $0 < |z| < 2$ . We can therefore expand it as a Laurent series about  $z = 0$  in that region:

$$\begin{aligned} \frac{1}{z^2(z-2)} &= -\frac{1}{2z^2} \left( \frac{1}{1-z/2} \right) = -\frac{1}{2z^2} \sum_{n=0}^{\infty} \left( \frac{z}{2} \right)^n \\ &= -\frac{1}{2} \left( \frac{1}{z^2} \right) - \frac{1}{4} \left( \frac{1}{z} \right) - \frac{1}{8} - \dots. \end{aligned}$$

Thus,  $a_{-1} = -\frac{1}{4}$ , and  $\oint_C dz/[z^2(z-2)] = 2\pi i a_{-1} = -i\pi/2$ . A direct evaluation of the integral is nontrivial. In fact, we will see later that to find certain integrals, it is advantageous to cast them in the form of a contour integral and use either Equation (9.17) or a related equation. ■

Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be analytic at  $z_0$ . Then by definition, there exists a *neighborhood* of  $z_0$  in which  $f$  is analytic. In particular, we can find a circle  $|z - z_0| = r > 0$  in whose interior  $f$  has a Taylor expansion.

zero of order  $k$  **9.6.14. Definition.** *Let*

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n \equiv \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Then  $f$  is said to have a **zero of order  $k$**  at  $z_0$  if  $f^{(n)}(z_0) = 0$  for  $n = 0, 1, \dots, k - 1$  but  $f^{(k)}(z_0) \neq 0$ .

In that case  $f(z) = (z - z_0)^k \sum_{n=0}^{\infty} a_{k+n} (z - z_0)^n$ , where  $a_k \neq 0$  and  $|z - z_0| < r$ . We define  $g(z)$  as

$$g(z) = \sum_{n=0}^{\infty} a_{k+n} (z - z_0)^n \quad \text{where } |z - z_0| < r$$

and note that  $g(z_0) = a_k \neq 0$ . Convergence of the series on the RHS implies that  $g(z)$  is continuous at  $z_0$ . Consequently, for each  $\epsilon > 0$ , there exists  $\delta$  such that  $|g(z) - a_k| < \epsilon$  whenever  $|z - z_0| < \delta$ . If we choose  $\epsilon = |a_k|/2$ , then, for some  $\delta_0 > 0$ ,  $|g(z) - a_k| < |a_k|/2$  whenever  $|z - z_0| < \delta_0$ . Thus, as long as  $z$  is inside the circle  $|z - z_0| < \delta_0$ ,  $g(z)$  cannot vanish (because if it did the first inequality would imply that  $|a_k| < |a_k|/2$ ). We therefore have the following result.

the zeros of an analytic function are isolated

**9.6.15. Theorem.** *Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be analytic at  $z_0$  and  $f(z_0) = 0$ . Then there exists a neighborhood of  $z_0$  throughout which  $f$  has no other zeros unless  $f$  is identically zero there. Thus, the zeros of an analytic function are **isolated**.*

simple zero

When  $k = 1$ , we say that  $z_0$  is a **simple zero** of  $f$ . To find the order of the zero of a function at a point, we differentiate the function, evaluate the derivative at that point, and continue the process until we obtain a nonzero value for the derivative.

**9.6.16. Example.** (a) The zeros of  $\cos z$ , which are  $z = (2k + 1)\pi/2$ , are all simple, because

$$\left. \frac{d}{dz} \cos z \right|_{z=(2k+1)\pi/2} = -\sin \left[ (2k + 1) \frac{\pi}{2} \right] \neq 0.$$

(b) To find the order of the zero of  $f(z) = e^z - 1 - z - z^2/2$  at  $z = 0$ , we differentiate  $f(z)$  and evaluate  $f'(0)$ :

$$f'(0) = (e^z - 1 - z)_{z=0} = 0.$$

Differentiating again gives  $f''(0) = (e^z - 1)_{z=0} = 0$ . Differentiating once more yields  $f'''(0) = (e^z)_{z=0} = 1$ . Thus, the zero is of order 3. ■



## 9.7 Problems

**9.1.** Show that the function  $w = 1/z$  maps the straight line  $y = a$  in the  $z$ -plane onto a circle in the  $w$ -plane with radius  $1/(2|a|)$  and center  $(0, 1/(2a))$ .

**9.2.** (a) Using the chain rule, find  $\partial f/\partial z^*$  and  $\partial f/\partial z$  in terms of partial derivatives with respect to  $x$  and  $y$ .

(b) Evaluate  $\partial f/\partial z^*$  and  $\partial f/\partial z$  assuming that the Cauchy–Riemann conditions hold.

**9.3.** Show that when  $z$  is represented by polar coordinates, the derivative of a function  $f(z)$  can be written as

$$\frac{df}{dz} = e^{-i\theta} \left( \frac{\partial U}{\partial r} + i \frac{\partial V}{\partial r} \right),$$

where  $U$  and  $V$  are the real and imaginary parts of  $f(z)$  written in polar coordinates. What are the C-R conditions in polar coordinates? Hint: Start with the C-R conditions in Cartesian coordinates and apply the chain rule to them using  $x = r \cos \theta$  and  $y = r \sin \theta$ .

**9.4.** Show that  $d/dz(\ln z) = 1/z$ . Hint: Find  $u(x, y)$  and  $v(x, y)$  for  $\ln z$  and differentiate them.

**9.5.** Show that  $\sin z$  and  $\cos z$  have only real roots.

**9.6.** Show that

(a) the sum and the product of two entire functions are entire, and

(b) the ratio of two entire functions is analytic everywhere except at the zeros of the denominator.

**9.7.** Given that  $u = 2\lambda \ln[(x^2 + y^2)^{1/2}]$ , show that  $v = 2\lambda \tan^{-1}(y/x)$ , where  $u$  and  $v$  are the real and imaginary parts of an analytic function  $w(z)$ .

**9.8.** If  $w(z)$  is any complex potential, show that its (complex) derivative gives the components of the electric field.

**9.9.** (a) Show that the flux through an element of area  $da$  of the lateral surface of a cylinder (with arbitrary cross section) is  $d\phi = dz(|\mathbf{E}| ds)$  where  $ds$  is an arc length along the equipotential surface.

(b) Prove that  $|\mathbf{E}| = |dw/dz| = \partial v/\partial s$  where  $v$  is the imaginary part of the complex potential, and  $s$  is the parameter describing the length along the equipotential curves.

(c) Combine (a) and (b) to get

$$\text{flux per unit } z\text{-length} = \frac{\phi}{z_2 - z_1} = v(P_2) - v(P_1)$$

for any two points  $P_1$  and  $P_2$  on the cross-sectional curve of the lateral surface. Conclude that the total flux per unit  $z$ -length through a cylinder (with arbitrary

cross section) is  $[v]$ , the total change in  $v$  as one goes around the curve.

(d) Using Gauss's law, show that the capacitance per unit length for the capacitor consisting of the two conductors with potentials  $u_1$  and  $u_2$  is

$$c \equiv \frac{\text{charge per unit length}}{\text{potential difference}} = \frac{[v]/4\pi}{|u_2 - u_1|}.$$

**9.10.** Using Equation (9.7)

(a) find the equipotential curves (curves of constant  $u$ ) and curves of constant  $v$  for two line charges of equal magnitude and opposite signs located at  $y = a$  and  $y = -a$  in the  $xy$ -plane.

(b) Show that

$$z = a \left( \sin \frac{v}{2\lambda} + i \sinh \frac{u}{2\lambda} \right) / \left( \cosh \frac{u}{2\lambda} - \cos \frac{v}{2\lambda} \right)$$

by solving Equation (9.7) for  $z$  and simplifying.

(c) Show that the equipotential curves are circles in the  $xy$ -plane of radii  $a / \sinh(u/2\lambda)$  with centers at  $(0, a \coth(u/2\lambda))$ , and that the curves of constant  $v$  are circles of radii  $a / \sin(v/2\lambda)$  with centers at  $(a \cot(v/2\lambda), 0)$ .

**9.11.** In this problem, you will find the capacitance per unit length of two cylindrical conductors of radii  $R_1$  and  $R_2$  the distance between whose centers is  $D$  by looking for two line charge densities  $+\lambda$  and  $-\lambda$  such that the two cylinders are two of the equipotential surfaces. From Problem 9.10, we have

$$R_i = \frac{a}{\sinh(u_i/2\lambda)}, \quad y_i = a \coth(u_i/2\lambda), \quad i = 1, 2,$$

where  $y_1$  and  $y_2$  are the locations of the centers of the two conductors on the  $y$ -axis (which we assume to connect the two centers).

(a) Show that  $D = |y_1 - y_2| = \left| R_1 \cosh \frac{u_1}{2\lambda} - R_2 \cosh \frac{u_2}{2\lambda} \right|$ .

(b) Square both sides and use  $\cosh(a - b) = \cosh a \cosh b - \sinh a \sinh b$  and the expressions for the  $R$ 's and the  $y$ 's given above to obtain

$$\cosh \left( \frac{u_1 - u_2}{2\lambda} \right) = \left| \frac{R_1^2 + R_2^2 - D^2}{2R_1 R_2} \right|.$$

(c) Now find the capacitance per unit length. Consider the special case of two concentric cylinders.

(d) Find the capacitance per unit length of a cylinder and a plane, by letting one of the radii, say  $R_1$ , go to infinity while  $h \equiv R_1 - D$  remains fixed.

9.12. Use Equations (9.4) and (9.5) to establish the following identities.

$$\begin{array}{ll}
 (a) \operatorname{Re}(\sin z) = \sin x \cosh y, & \operatorname{Im}(\sin z) = \cos x \sinh y. \\
 (b) \operatorname{Re}(\cos z) = \cos x \cosh y, & \operatorname{Im}(\cos z) = -\sin x \sinh y. \\
 (c) \operatorname{Re}(\sinh z) = \sinh x \cos y, & \operatorname{Im}(\sinh z) = \cosh x \sin y. \\
 (d) \operatorname{Re}(\cosh z) = \cosh x \cos y, & \operatorname{Im}(\cosh z) = \sinh x \sin y. \\
 (e) |\sin z|^2 = \sin^2 x + \sinh^2 y, & |\cos z|^2 = \cos^2 x + \sinh^2 y. \\
 (f) |\sinh z|^2 = \sinh^2 x + \sin^2 y, & |\cosh z|^2 = \sinh^2 x + \cos^2 y.
 \end{array}$$

9.13. Find all the zeros of  $\sinh z$  and  $\cosh z$ .

9.14. Verify the following hyperbolic identities.

$$\begin{array}{l}
 (a) \cosh^2 z - \sinh^2 z = 1. \\
 (b) \cosh(z_1 + z_2) = \cosh z_1 \cosh z_2 + \sinh z_1 \sinh z_2. \\
 (c) \sinh(z_1 + z_2) = \sinh z_1 \cosh z_2 + \cosh z_1 \sinh z_2. \\
 (d) \cosh 2z = \cosh^2 z + \sinh^2 z, \quad \sinh 2z = 2 \sinh z \cosh z. \\
 (e) \tanh(z_1 + z_2) = \frac{\tanh z_1 + \tanh z_2}{1 + \tanh z_1 \tanh z_2}.
 \end{array}$$

9.15. Show that

$$(a) \tanh\left(\frac{z}{2}\right) = \frac{\sinh x + i \sin y}{\cosh x + \cos y}, \quad (b) \coth\left(\frac{z}{2}\right) = \frac{\sinh x - i \sin y}{\cosh x - \cos y}.$$

9.16. Find all values of  $z$  such that

$$(a) e^z = -3, \quad (b) e^z = 1 + i\sqrt{3}, \quad (c) e^{2z-1} = 1.$$

9.17. Show that  $|e^{-z}| < 1$  if and only if  $\operatorname{Re}(z) > 0$ .

9.18. Show that both the real and the imaginary parts of an analytic function are harmonic.

9.19. Show that each of the following functions—call each one  $u(x, y)$ —is harmonic, and find the function's harmonic partner,  $v(x, y)$ , such that  $u(x, y) + iv(x, y)$  is analytic.

$$\begin{array}{lll}
 (a) x^3 - 3xy^2. & (b) e^x \cos y. & (c) \frac{x}{x^2 + y^2} \quad \text{where } x^2 + y^2 \neq 0. \\
 (d) e^{-2y} \cos 2x. & (e) e^{y^2 - x^2} \cos 2xy. & \\
 (f) e^x (x \cos y - y \sin y) + 2 \sinh y \sin x + x^3 - 3xy^2 + y. & &
 \end{array}$$

9.20. Prove the following identities.

$$(a) \cos^{-1} z = -i \ln(z \pm \sqrt{z^2 - 1}). \quad (b) \sin^{-1} z = -i \ln[iz \pm \sqrt{1 - z^2}].$$

$$(c) \tan^{-1} z = \frac{1}{2i} \ln \left( \frac{i - z}{i + z} \right). \quad (d) \cosh^{-1} z = \ln(z \pm \sqrt{z^2 - 1}).$$

$$(e) \sinh^{-1} z = \ln(z \pm \sqrt{z^2 + 1}). \quad (f) \tanh^{-1} z = \frac{1}{2} \ln \left( \frac{1 + z}{1 - z} \right).$$

9.21. Find the curve defined by each of the following equations.

$$(a) z = 1 - it, \quad 0 \leq t \leq 2. \quad (b) z = t + it^2, \quad -\infty < t < \infty.$$

$$(c) z = a(\cos t + i \sin t), \quad \frac{\pi}{2} \leq t \leq \frac{3\pi}{2}. \quad (d) z = t + \frac{i}{t}, \quad -\infty < t < 0.$$

9.22. Provide the details of the proof of part (a) of Proposition 9.3.1. Prove part (b) by showing that if  $f(z) = z' = x' + iy'$  is analytic and  $\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = 0$ , then

$$\frac{\partial^2 \Phi}{\partial x'^2} + \frac{\partial^2 \Phi}{\partial y'^2} = 0.$$

9.23. Let  $f(t) = u(t) + iv(t)$  be a (piecewise) continuous complex-valued function of a real variable  $t$  defined in the interval  $a \leq t \leq b$ . Show that if  $F(t) = U(t) + iV(t)$  is a function such that  $dF/dt = f(t)$ , then

$$\int_a^b f(t) dt = F(b) - F(a).$$

This is the fundamental theorem of calculus for complex variables.

9.24. Find the value of the integral  $\int_C [(z+2)/z] dz$ , where  $C$  is (a) the semicircle  $z = 2e^{i\theta}$ , for  $0 \leq \theta \leq \pi$ , (b) the semicircle  $z = 2e^{i\theta}$ , for  $\pi \leq \theta \leq 2\pi$ , and (c) the circle  $z = 2e^{i\theta}$ , for  $-\pi \leq \theta \leq \pi$ .

9.25. Evaluate the integral  $\int_\gamma dz/(z-1-i)$  where  $\gamma$  is (a) the line joining  $z_1 = 2i$  and  $z_2 = 3$ , and (b) the broken path from  $z_1$  to the origin and from there to  $z_2$ .

9.26. Evaluate the integral  $\int_C z^m (z^*)^n dz$ , where  $m$  and  $n$  are integers and  $C$  is the circle  $|z| = 1$  taken counterclockwise.

9.27. Let  $C$  be the boundary of the square with vertices at the points  $z = 0$ ,  $z = 1$ ,  $z = 1 + i$ , and  $z = i$  with counterclockwise direction. Evaluate

$$\oint_C (5z + 2) dz \quad \text{and} \quad \oint_C e^{\pi z^*} dz.$$

**9.28.** Let  $C_1$  be a simple closed contour. Deform  $C_1$  into a new contour  $C_2$  in such a way that  $C_1$  does not encounter any singularity of an analytic function  $f$  in the process. Show that

$$\oint_{C_1} f(z) dz = \oint_{C_2} f(z) dz.$$

That is, the contour can always be deformed into simpler shapes (such as a circle) and the integral evaluated.

**9.29.** Use the result of the previous problem to show that

$$\oint_C \frac{dz}{z-1-i} = 2\pi i \text{ and } \oint_C (z-1-i)^{m-1} dz = 0 \text{ for } m = \pm 1, \pm 2, \dots$$

when  $C$  is the boundary of a square with vertices at  $z = 0$ ,  $z = 2$ ,  $z = 2 + 2i$ , and  $z = 2i$ , taken counterclockwise.

**9.30.** Use Equation (9.12) and the binomial expansion to show that

$$\frac{d}{dz}(z^m) = mz^{m-1}.$$

**9.31.** Evaluate  $\oint_C dz/(z^2 - 1)$  where  $C$  is the circle  $|z| = 3$  integrated in the positive sense. Hint: Deform  $C$  into a contour  $C'$  that bypasses the singularities of the integrand.

**9.32.** Show that when  $f$  is analytic within and on a simple closed contour  $C$  and  $z_0$  is not on  $C$ , then

$$\oint_C \frac{f'(z) dz}{z - z_0} = \oint_C \frac{f(z) dz}{(z - z_0)^2}.$$

**9.33.** Let  $C$  be the boundary of a square whose sides lie along the lines  $x = \pm 3$  and  $y = \pm 3$ . For the positive sense of integration, evaluate each of the following integrals.

$$(a) \oint_C \frac{e^{-z}}{z - i\pi/2} dz. \quad (b) \oint_C \frac{e^z}{z(z^2 + 10)} dz. \quad (c) \oint_C \frac{\cos z}{(z - \frac{\pi}{4})(z^2 - 10)} dz.$$

$$(d) \oint_C \frac{\sinh z}{z^4} dz. \quad (e) \oint_C \frac{\cosh z}{z^4} dz. \quad (f) \oint_C \frac{\cos z}{z^3} dz.$$

$$(g) \oint_C \frac{\cos z}{(z - i\pi/2)^2} dz. \quad (h) \oint_C \frac{e^z}{(z - i\pi)^2} dz. \quad (i) \oint_C \frac{\cos z}{z + i\pi} dz.$$

$$(j) \oint_C \frac{e^z}{z^2 - 5z + 4} dz. \quad (k) \oint_C \frac{\sinh z}{(z - i\pi/2)^2} dz. \quad (l) \oint_C \frac{\cosh z}{(z - i\pi/2)^2} dz.$$

$$(m) \oint_C \frac{\tan z}{(z - \alpha)^2} dz \quad \text{for } -3 < \alpha < 3. \quad (n) \oint_C \frac{z^2}{(z - 2)(z^2 - 10)} dz.$$

9.34. Let  $C$  be the circle  $|z - i| = 3$  integrated in the positive sense. Find the value of each of the following integrals.

$$\begin{aligned} (a) \oint_C \frac{e^z}{z^2 + \pi^2} dz, & \quad (b) \oint_C \frac{\sinh z}{(z^2 + \pi^2)^2} dz, & \quad (c) \oint_C \frac{dz}{z^2 + 9}. \\ (d) \oint_C \frac{dz}{(z^2 + 9)^2}, & \quad (e) \oint_C \frac{\cosh z}{(z^2 + \pi^2)^3} dz, & \quad (f) \oint_C \frac{z^2 - 3z + 4}{z^2 - 4z + 3} dz. \end{aligned}$$

9.35. Show that Legendre polynomials (for  $|x| < 1$ ) can be represented as

$$P_n(x) = \frac{(-1)^n}{2^n(2\pi i)} \oint_C \frac{(1 - z^2)^n}{(z - x)^{n+1}} dz,$$

where  $C$  is the unit circle around the origin.

9.36. Let  $f$  be analytic within and on the circle  $\gamma_0$  given by  $|z - z_0| = r_0$  and integrated in the positive sense. Show that **Cauchy's inequality** holds:

$$|f^{(n)}(z_0)| \leq \frac{n!M}{r_0^n},$$

where  $M$  is the maximum value of  $|f(z)|$  on  $\gamma_0$ .

9.37. Expand  $\sinh z$  in a Taylor series about the point  $z = i\pi$ .

9.38. What is the largest circle within which the Maclaurin series for  $\tanh z$  converges to  $\tanh z$ ?

9.39. Find the (unique) Laurent expansion of each of the following functions about the origin for its entire region of analyticity.

$$\begin{aligned} (a) \frac{1}{(z-2)(z-3)}, & \quad (b) z \cos(z^2), & \quad (c) \frac{1}{z^2(1-z)}, & \quad (d) \frac{\sinh z - z}{z^4}. \\ (e) \frac{1}{(1-z)^3}, & \quad (f) \frac{1}{z^2 - 1}, & \quad (g) \frac{z^2 - 4}{z^2 - 9}, & \quad (h) \frac{1}{(z^2 - 1)^2}. \\ (i) \frac{z}{z-1}. \end{aligned}$$

9.40. Show that the following functions are entire.

$$(a) f(z) = \begin{cases} \frac{e^{2z} - 1}{z^2} - \frac{2}{z} & \text{for } z \neq 0, \\ 2 & \text{for } z = 0. \end{cases} \quad (b) f(z) = \begin{cases} \frac{\sin z}{z} & \text{for } z \neq 0, \\ 1 & \text{for } z = 0. \end{cases}$$

$$(c) f(z) = \begin{cases} \frac{\cos z}{z^2 - \pi^2/4} & \text{for } z \neq \pm\pi/2, \\ -1/\pi & \text{for } z = \pm\pi/2. \end{cases}$$

9.41. Let  $f$  be analytic at  $z_0$  and  $f(z_0) = f'(z_0) = \cdots = f^{(k)}(z_0) = 0$ . Show that the following function is analytic at  $z_0$ :

$$g(z) = \begin{cases} \frac{f(z)}{(z - z_0)^{k+1}} & \text{for } z \neq z_0, \\ \frac{f^{(k+1)}(z_0)}{(k+1)!} & \text{for } z = z_0. \end{cases}$$

9.42. Obtain the first few nonzero terms of the Laurent series expansion of each of the following functions about the origin. Also find the integral of the function along a small simple closed contour encircling the origin.

$$\begin{array}{llll} (a) \frac{1}{\sin z} & (b) \frac{1}{1 - \cos z} & (c) \frac{z}{1 - \cosh z} & (d) \frac{z^2}{z - \sin z} \\ (e) \frac{z^4}{6z + z^3 - 6 \sinh z} & (f) \frac{1}{z^2 \sin z} & (g) \frac{1}{e^z - 1} & \end{array}$$

### Additional Reading

1. Churchill, R. and Verhey, R. *Complex Variables and Applications*, 3rd ed., McGraw-Hill, 1974. An introductory text on complex variables with many examples and exercises.
2. Lang, S. *Complex Analysis*, 2nd ed., Springer-Verlag, 1985. An extremely well-written book by a master expositor. Although the book has a formal tone, the clarity of exposition and the use of many examples make this book very readable.

# Calculus of Residues

One of the most powerful tools made available by complex analysis is the theory of residues, which makes possible the routine evaluation of certain definite integrals that are impossible to calculate otherwise. The derivation, application, and analysis of this tool constitute the main focus of this chapter. In the preceding chapter we saw examples in which integrals were related to expansion coefficients of Laurent series. Here we will develop a systematic way of evaluating both real and complex integrals.

## 10.1 Residues

isolated singularity

Recall that a singular point  $z_0$  of  $f : \mathbb{C} \rightarrow \mathbb{C}$  is a point at which  $f$  fails to be analytic. If in addition, there is some neighborhood of  $z_0$  in which  $f$  is analytic at every point (except of course at  $z_0$  itself), then  $z_0$  is called an **isolated singularity** of  $f$ . Almost all the singularities we have encountered so far have been isolated singularities. However, we will see later—when discussing multivalued functions—that singularities that are not isolated do exist.

Let  $z_0$  be an isolated singularity of  $f$ . Then there exists an  $r > 0$  such that within the “annular” region  $0 < |z - z_0| < r$ , the function  $f$  has the Laurent expansion

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n \equiv \sum_{n=0}^{\infty} a_n(z - z_0)^n + \frac{b_1}{z - z_0} + \frac{b_2}{(z - z_0)^2} + \cdots$$

where

$$a_n = \frac{1}{2\pi i} \oint_C \frac{f(\xi) d\xi}{(\xi - z_0)^{n+1}} \quad \text{and} \quad b_n = \frac{1}{2\pi i} \oint_C f(\xi)(\xi - z_0)^{n-1} d\xi.$$



In particular,

$$b_1 = \frac{1}{2\pi i} \oint_C f(\xi) d\xi, \quad (10.1)$$

residue defined

where  $C$  is any simple closed contour around  $z_0$ , traversed in the positive sense, on and interior to which  $f$  is analytic except at the point  $z_0$  itself. The complex number  $b_1$ , which is essentially the integral of  $f(z)$  along the contour, is called the **residue** of  $f$  at the isolated singular point  $z_0$ . It is important to note that the residue is independent of the contour  $C$  as long as  $z_0$  is the only isolated singular point within  $C$ .

**Pierre Alphonse Laurent** (1813–1854) graduated from the Ecole Polytechnique near the top of his class and became a second lieutenant in the engineering corps. On his return from the war in Algeria, he took part in the effort to improve the port at Le Havre, spending six years there directing various parts of the project. Laurent's superior officers admired the breadth of his practical experience and the good judgment it afforded the young engineer. During this period he wrote his first scientific paper, on the calculus of variations, and submitted it to the French Academy of Sciences for the grand prix in mathematics. Unfortunately the competition had already closed (although the judges had not yet declared a winner), and Laurent's submission was not successful. However, the paper so impressed Cauchy that he recommended its publication, also without success.

The paper for which Laurent is most well known suffered a similar fate. In it he described a more general form of a theorem earlier proven by Cauchy for the power series expansion of a function. Laurent realized that one could generalize this result to hold in any annular region between two singular or discontinuous points by using both positive and negative powers in the series, thus allowing treatment of regions beyond the first singular or discontinuous point. Again, Cauchy argued for the paper's publication without success. The passage of time provided a more just reward, however, and the use of Laurent series became a fundamental tool in complex analysis.

Laurent later worked in the theory of light waves and contended with Cauchy over the interpretation of the differential equations the latter had formulated to explain the behavior of light. Little came of his work in this area, however, and Laurent died at the age of forty-two, a captain serving on the committee on fortifications in Paris. His widow pressed to have two more of his papers read to the Academy, only one of which was published.

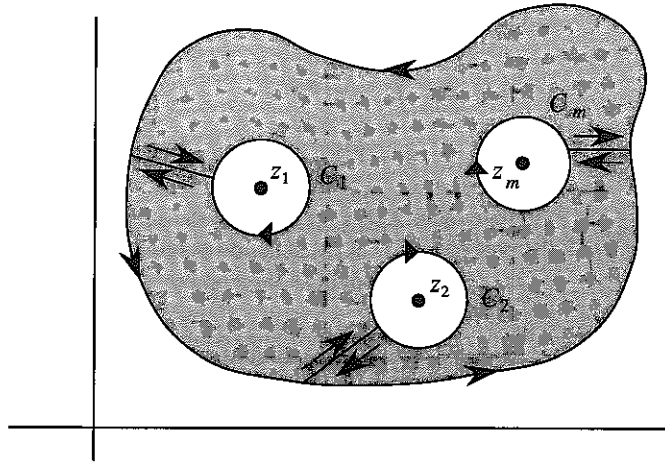
We use the notation  $\text{Res}[f(z_0)]$  to denote the residue of  $f$  at the isolated singular point  $z_0$ . Equation (10.1) can then be written as

$$\oint_C f(z) dz = 2\pi i \text{Res}[f(z_0)].$$

What if there are several isolated singular points within the simple closed contour  $C$ ? The following theorem provides the answer.

residue theorem

**10.1.1. Theorem.** (the residue theorem) *Let  $C$  be a positively oriented simple*



**Figure 10.1** Singularities are avoided by going around them.

closed contour within and on which a function  $f$  is analytic except at a finite number of isolated singular points  $z_1, z_2, \dots, z_m$  interior to  $C$ . Then

$$\oint_C f(z) dz = 2\pi i \sum_{k=1}^m \text{Res}[f(z_k)]. \tag{10.2}$$

*Proof.* Let  $C_k$  be the positively traversed circle around  $z_k$ . Then Figure 10.1 and the Cauchy-Goursat theorem yield

$$0 = \oint_{C'} f(z) dz = - \oint_{\text{circles}} f(z) dz + \oint_{\text{parallel lines}} f(z) dz + \oint_C f(z) dz,$$

where  $C'$  is the union of all the contours, and the minus sign on the first integral is due to the interiors of all circles lie to our right as we traverse their boundaries. The contributions of the parallel lines cancel out, and we obtain

$$\oint_C f(z) dz = \sum_{k=1}^m \oint_{C_k} f(z) dz = \sum_{k=1}^m 2\pi i \text{Res}[f(z_k)],$$

where in the last step the definition of residue at  $z_k$  has been used. □

**10.1.2. Example.** Let us evaluate the integral  $\oint_C (2z - 3) dz/[z(z - 1)]$  where  $C$  is the circle  $|z| = 2$ . There are two isolated singularities in  $C$ ,  $z_1 = 0$  and  $z_2 = 1$ . To find  $\text{Res}[f(z_1)]$ , we expand around the origin:

$$\frac{2z - 3}{z(z - 1)} = \frac{3}{z} - \frac{1}{z - 1} = \frac{3}{z} + \frac{1}{1 - z} = \frac{3}{z} + 1 + z + \dots \quad \text{for } |z| < 1.$$

This gives  $\text{Res}[f(z_1)] = 3$ . Similarly, expanding around  $z = 1$  gives

$$\frac{2z - 3}{z(z - 1)} = \frac{3}{z - 1 + 1} - \frac{1}{z - 1} = -\frac{1}{z - 1} + 3 \sum_{k=0}^{\infty} (-1)^k (z - 1)^k,$$

which yields  $\text{Res}[f(z_2)] = -1$ . Thus,

$$\oint_C \frac{2z - 3}{z(z - 1)} dz = 2\pi i \{\text{Res}[f(z_1)] + \text{Res}[f(z_2)]\} = 2\pi i(3 - 1) = 4\pi i. \quad \blacksquare$$

## 10.2 Classification of Isolated Singularities

Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  have an isolated singularity at  $z_0$ . Then there exist a real number  $r > 0$  and an annular region  $0 < |z - z_0| < r$  such that  $f$  can be represented by the Laurent series

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n}. \quad (10.3)$$

The second sum in Equation (10.3), involving negative powers of  $(z - z_0)$ , is called the **principal part** of  $f$  at  $z_0$ . We can use the principal part to distinguish three types of isolated singularities. The behavior of the function near the isolated singularity is fundamentally different in each case.

principal part of a function

removable singular point

1. If  $b_n = 0$  for all  $n \geq 1$ ,  $z_0$  is called a **removable singular point** of  $f$ . In this case, the Laurent series contains only nonnegative powers of  $(z - z_0)$ , and setting  $f(z_0) = a_0$  makes the function analytic at  $z_0$ . For example, the function  $f(z) = (e^z - 1 - z)/z^2$ , which is indeterminate at  $z = 0$ , becomes entire if we set  $f(0) = \frac{1}{2}$ , because its Laurent series  $f(z) = \frac{1}{2} + \frac{z}{3!} + \frac{z^2}{4!} + \dots$  has no negative power.

poles defined

2. If  $b_n = 0$  for all  $n > m$  and  $b_m \neq 0$ ,  $z_0$  is called a **pole of order  $m$** . In this case, the expansion takes the form

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \frac{b_1}{z - z_0} + \dots + \frac{b_m}{(z - z_0)^m}$$

simple pole

for  $0 < |z - z_0| < r$ . In particular, if  $m = 1$ ,  $z_0$  is called a **simple pole**.

essential singularity

3. If the principal part of  $f$  at  $z_0$  has an infinite number of nonzero terms, the point  $z_0$  is called an **essential singularity**. A prototype of functions that have essential singularities is

$$\exp\left(\frac{1}{z}\right) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{1}{z}\right)^n,$$

which has an essential singularity at  $z = 0$  and a residue of 1 there. To see how strange such functions are, we let  $a$  be any real number, and consider  $z = 1/(\ln a + 2n\pi i)$  for  $n = 0, \pm 1, \pm 2, \dots$ . For such a  $z$  we have  $e^{1/z} = e^{\ln a + 2n\pi i} = ae^{2n\pi i} = a$ . In particular, as  $n \rightarrow \infty$ ,  $z$  gets arbitrarily close to the origin. Thus, in an arbitrarily small neighborhood of the origin, there are infinitely many points at which the function  $\exp(1/z)$  takes on an arbitrary value  $a$ . In other words, as  $z \rightarrow 0$ , the function gets arbitrarily close to any real number! This result holds for all functions with essential singularities.

### 10.2.1. Example. ORDER OF POLES

(a) The function  $(z^2 - 3z + 5)/(z - 1)$  has a Laurent series around  $z = 1$  containing only three terms:  $(z^2 - 3z + 5)/(z - 1) = -1 + (z - 1) + 3/(z - 1)$ . Thus, it has a simple pole at  $z = 1$ , with a residue of 3.

(b) The function  $\sin z/z^6$  has a Laurent series

$$\frac{\sin z}{z^6} = \frac{1}{z^6} \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} = \frac{1}{z^5} - \frac{1}{6z^3} + \frac{1}{(5!)z} - \frac{z}{7!} + \dots$$

about  $z = 0$ . The principal part has three terms. The pole, at  $z = 0$ , is of order 5, and the function has a residue of  $1/120$  at  $z = 0$ .

(c) The function  $(z^2 - 5z + 6)/(z - 2)$  has a removable singularity at  $z = 2$ , because

$$\frac{z^2 - 5z + 6}{z - 2} = \frac{(z - 2)(z - 3)}{z - 2} = z - 3 = -1 + (z - 2)$$

and  $b_n = 0$  for all  $n$ . ■

### 10.2.2. Example. SINGULARITIES OF A RATIONAL FUNCTION

rational function

In this example we show that a function whose only singularities in the entire complex plane are poles must be a **rational function**,<sup>1</sup> i.e., the ratio of two polynomials. Let  $f$  be such a function and let  $\{z_j\}_{j=1}^n$  be its poles such that  $z_j$  is of order  $m_j$ . Expand the function about  $z_1$  in a Laurent series

$$f(z) = \frac{b_1}{z - z_1} + \dots + \frac{b_{m_1}}{(z - z_1)^{m_1}} + \sum_{k=0}^{\infty} a_k (z - z_1)^k = \frac{P_1(z)}{(z - z_1)^{m_1}} + g_1(z),$$

A function whose only singularities in the entire complex plane are poles is a rational function.

where  $P_1(z)$  is a polynomial of degree  $m_1 - 1$  in  $z$  and  $g_1$  is analytic at  $z_1$ . It should be clear that the remaining poles of  $f$  are in  $g_1$ . So, expand  $g_1$  about  $z_2$  in a Laurent series. A similar argument as above yields  $g_1(z) = P_2(z)/(z - z_2)^{m_2} + g_2(z)$  where  $P_2(z)$  is a polynomial of degree  $m_2 - 1$  in  $z$  and  $g_2$  is analytic at  $z_1$  and  $z_2$ . Continuing in this manner, we get

$$f(z) = \frac{P_1(z)}{(z - z_1)^{m_1}} + \frac{P_2(z)}{(z - z_2)^{m_2}} + \dots + \frac{P_n(z)}{(z - z_n)^{m_n}} + g(z),$$

where  $g$  has no poles. Since all poles of  $f$  have been isolated in the sum,  $g$  must be analytic everywhere in  $\mathbb{C}$ , i.e., an entire function. Now substitute  $1/t$  for  $z$ , take the limit  $t \rightarrow 0$ ,

<sup>1</sup>We assume that the point at infinity is not a pole of the function, i.e., that  $f(1/z)$  does not have a pole at the origin.

and note that, since the degree of  $P_i$  is  $m_i - 1$ , all the terms in the preceding equation go to zero except possibly  $g(1/t)$ . Moreover,

$$\lim_{t \rightarrow 0} g(1/t) \neq \infty,$$

because, by assumption, the point at infinity is not a pole of  $f$ . Thus,  $g$  is a bounded entire function. By Proposition 9.5.5,  $g$  must be a constant. Taking a common denominator for all the terms yields a ratio of two polynomials. ■

The type of isolated singularity that is most important in applications is of the second type—poles. For a function that has a pole of order  $m$  at  $z_0$ , the calculation of residues is routine. Such a calculation, in turn, enables us to evaluate many integrals effortlessly. How do we calculate the residue of a function  $f$  having a pole of order  $m$  at  $z_0$ ?

It is clear that if  $f$  has a pole of order  $m$ , then  $g : \mathbb{C} \rightarrow \mathbb{C}$  defined by  $g(z) \equiv (z - z_0)^m f(z)$  is analytic at  $z_0$ . Thus, for any simple closed contour  $C$  that contains  $z_0$  but no other singular point of  $f$ , we have

$$\text{Res}[f(z_0)] = \frac{1}{2\pi i} \oint_C f(z) dz = \frac{1}{2\pi i} \oint_C \frac{g(z) dz}{(z - z_0)^m} = \frac{g^{(m-1)}(z_0)}{(m - 1)!}.$$

In terms of  $f$  this yields<sup>2</sup>

$$\text{Res}[f(z_0)] = \frac{1}{(m - 1)!} \lim_{z \rightarrow z_0} \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)]. \tag{10.4}$$

For the special, but important, case of a simple pole, we obtain

$$\text{Res}[f(z_0)] = \lim_{z \rightarrow z_0} [(z - z_0) f(z)]. \tag{10.5}$$

### 10.3 Evaluation of Definite Integrals

The most widespread application of residues occurs in the evaluation of real definite integrals. It is possible to “complexify” certain real definite integrals and relate them to contour integrations in the complex plane. We will discuss this method shortly; however, we first need a lemma.

**10.3.1. Lemma.** (Jordan’s lemma) *Let  $C_R$  be a semicircle of radius  $R$  in the upper half of the complex plane (UHP) and centered at the origin. Let  $f$  be a function that tends uniformly to zero faster than  $1/|z|$  for  $\arg(z) \in [0, \pi]$  as  $|z| \rightarrow \infty$ . Let  $\alpha$  be a nonnegative real number. Then*

$$\lim_{R \rightarrow \infty} I_R \equiv \lim_{R \rightarrow \infty} \int_{C_R} e^{i\alpha z} f(z) dz = 0.$$

<sup>2</sup>The limit is taken because in many cases the mere substitution of  $z_0$  may result in an indeterminate form.

*Proof.* For  $z \in C_R$  we write  $z = Re^{i\theta}$ ,  $dz = iRe^{i\theta} d\theta$ , and

$$i\alpha z = i\alpha(R \cos \theta + iR \sin \theta) = i\alpha R \cos \theta - \alpha R \sin \theta$$

and substitute in the absolute value of the integral to show that

$$|I_R| \leq \int_0^\pi e^{-\alpha R \sin \theta} R |f(Re^{i\theta})| d\theta.$$

By assumption,  $R |f(Re^{i\theta})| < \epsilon(R)$  independent of  $\theta$ , where  $\epsilon(R)$  is an arbitrary positive number that tends to zero as  $R \rightarrow \infty$ . By breaking up the interval of integration into two equal pieces and changing  $\theta$  to  $\pi - \theta$  in the second integral, one can show that

$$|I_R| < 2\epsilon(R) \int_0^{\pi/2} e^{-\alpha R \sin \theta} d\theta.$$

Furthermore,  $\sin \theta \geq 2\theta/\pi$  for  $0 \leq \theta \leq \pi/2$  (see Figure 10.2 for a “proof”). Thus,

$$|I_R| < 2\epsilon(R) \int_0^{\pi/2} e^{-(2\alpha R/\pi)\theta} d\theta = \frac{\pi\epsilon(R)}{\alpha R} (1 - e^{-\alpha R}),$$

which goes to zero as  $R$  gets larger and larger.  $\square$

Note that Jordan’s lemma applies for  $\alpha = 0$  as well, because  $(1 - e^{-\alpha R}) \rightarrow \alpha R$  as  $\alpha \rightarrow 0$ . If  $\alpha < 0$ , the lemma is still valid if the semicircle  $C_R$  is taken in the lower half of the complex plane (LHP) and  $f(z)$  goes to zero uniformly for  $\pi \leq \arg(z) \leq 2\pi$ .

We are now in a position to apply the residue theorem to the evaluation of definite integrals. The three types of integrals most commonly encountered are discussed separately below. In all cases we assume that Jordan’s lemma holds.

### 10.3.1 Integrals of Rational Functions

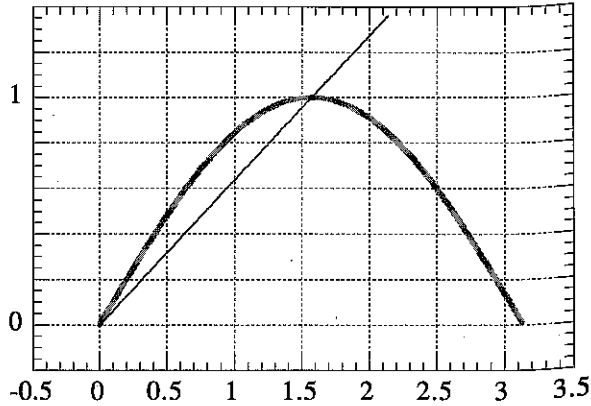
The first type of integral we can evaluate using the residue theorem is of the form

$$I_1 = \int_{-\infty}^{\infty} \frac{p(x)}{q(x)} dx,$$

where  $p(x)$  and  $q(x)$  are real polynomials, and  $q(x) \neq 0$  for any real  $x$ . We can then write

$$I_1 = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{p(x)}{q(x)} dx = \lim_{R \rightarrow \infty} \int_{C_x} \frac{p(z)}{q(z)} dz,$$

where  $C_x$  is the (open) contour lying on the real axis from  $-R$  to  $+R$ . Assuming that Jordan’s lemma holds, we can close that contour by adding to it the semicircle



**Figure 10.2** The “proof” of  $\sin \theta \geq 2\theta/\pi$  for  $0 \leq \theta \leq \pi/2$ . The line is the graph of  $y = 2\theta/\pi$ ; the curve is that of  $y = \sin \theta$ .

of radius  $R$  [see Figure 10.3(a)]. This will not affect the value of the integral, because in the limit  $R \rightarrow \infty$ , the contribution of the integral of the semicircle tends to zero. We close the contour in the UHP if  $q(z)$  has at least one zero there. We then get

$$I_1 = \lim_{R \rightarrow \infty} \oint_C \frac{p(z)}{q(z)} dz = 2\pi i \sum_{j=1}^k \operatorname{Res} \left[ \frac{p(z_j)}{q(z_j)} \right],$$

where  $C$  is the closed contour composed of the interval  $(-R, R)$  and the semicircle  $C_R$ , and  $\{z_j\}_{j=1}^k$  are the zeros of  $q(z)$  in the UHP. We may instead close the contour in the LHP,<sup>3</sup> in which case

$$I_1 = -2\pi i \sum_{j=1}^k \operatorname{Res} \left[ \frac{p(z_j)}{q(z_j)} \right],$$

where  $\{z_j\}_{j=1}^k$  are the zeros of  $q(z)$  in the LHP. The minus sign indicates that in the LHP we (are forced to) integrate clockwise.

**10.3.2. Example.** Let us evaluate the integral  $I = \int_0^\infty x^2 dx / [(x^2 + 1)(x^2 + 9)]$ . Since the integrand is even, we can extend the interval of integration to all real numbers (and divide the result by 2). It is shown below that Jordan’s lemma holds. Therefore, we write the contour integral corresponding to  $I$ :

$$I = \frac{1}{2} \oint_C \frac{z^2 dz}{(z^2 + 1)(z^2 + 9)},$$

<sup>3</sup>Provided that Jordan’s lemma holds there.

where  $C$  is as shown in Figure 10.3(a). Note that the contour is traversed in the positive sense. This is always true for the UHP. The singularities of the function in the UHP are the *simple* poles  $i$  and  $3i$  corresponding to the simple zeros of the denominator. The residues at these poles are

$$\begin{aligned}\operatorname{Res}[f(i)] &= \lim_{z \rightarrow i} (z - i) \frac{z^2}{(z - i)(z + i)(z^2 + 9)} = -\frac{1}{16i}, \\ \operatorname{Res}[f(3i)] &= \lim_{z \rightarrow 3i} (z - 3i) \frac{z^2}{(z^2 + 1)(z - 3i)(z + 3i)} = \frac{3}{16i}.\end{aligned}$$

Thus, we obtain

$$I = \int_0^{\infty} \frac{x^2 dx}{(x^2 + 1)(x^2 + 9)} = \frac{1}{2} \oint_C \frac{z^2 dz}{(z^2 + 1)(z^2 + 9)} = \pi i \left( -\frac{1}{16i} + \frac{3}{16i} \right) = \frac{\pi}{8}.$$

It is instructive to obtain the same results using the LHP. In this case, the contour is as shown in Figure 10.3 (b) and is taken clockwise, so we have to introduce a minus sign. The singular points are at  $z = -i$  and  $z = -3i$ . These are simple poles at which the residues of the function are

$$\begin{aligned}\operatorname{Res}[f(-i)] &= \lim_{z \rightarrow -i} (z + i) \frac{z^2}{(z - i)(z + i)(z^2 + 9)} = \frac{1}{16i}, \\ \operatorname{Res}[f(-3i)] &= \lim_{z \rightarrow -3i} (z + 3i) \frac{z^2}{(z^2 + 1)(z - 3i)(z + 3i)} = -\frac{3}{16i}.\end{aligned}$$

Therefore,

$$I = \int_0^{\infty} \frac{x^2 dx}{(x^2 + 1)(x^2 + 9)} = \frac{1}{2} \oint_C \frac{z^2 dz}{(z^2 + 1)(z^2 + 9)} = -\pi i \left( \frac{1}{16i} - \frac{3}{16i} \right) = \frac{\pi}{8}.$$

To show that Jordan's lemma applies to this integral, we have only to establish that  $\lim_{R \rightarrow \infty} R|f(Re^{i\theta})| = 0$ . In the case at hand,  $\alpha = 0$  because there is no exponential function in the integrand. Thus,

$$R|f(Re^{i\theta})| = R \left| \frac{R^2 e^{2i\theta}}{(R^2 e^{2i\theta} + 1)(R^2 e^{2i\theta} + 9)} \right| = \frac{R^3}{|R^2 e^{2i\theta} + 1| |R^2 e^{2i\theta} + 9|},$$

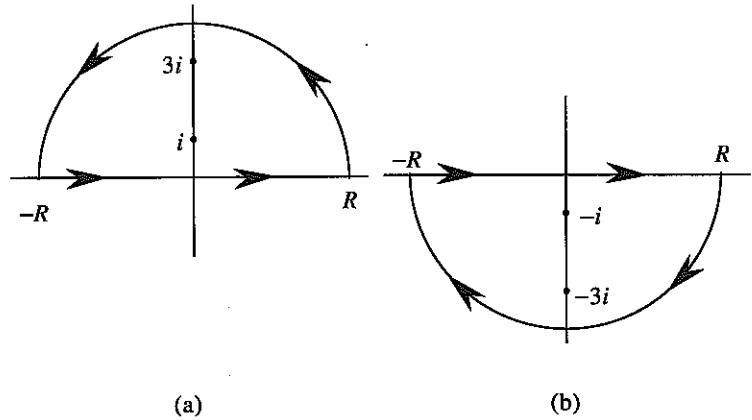
which clearly goes to zero as  $R \rightarrow \infty$ . ■

**10.3.3. Example.** Let us now consider a slightly more complicated integral:

$$\int_{-\infty}^{\infty} \frac{x^2 dx}{(x^2 + 1)(x^2 + 4)^2},$$

which in the UHP turns into  $\oint_C z^2 dz / [(z^2 + 1)(z^2 + 4)^2]$ . The poles in the UHP are at  $z = i$  and  $z = 2i$ . The former is a simple pole, and the latter is a pole of order 2. Thus,





**Figure 10.3** (a) The large semicircle is chosen in the UHP. (b) Note how the direction of contour integration is forced to be clockwise when the semicircle is chosen in the LHP.

using Equations (10.5) and (10.4), we obtain

$$\begin{aligned}\operatorname{Res}[f(i)] &= \lim_{z \rightarrow i} (z - i) \frac{z^2}{(z - i)(z + i)(z^2 + 4)^2} = -\frac{1}{18i}, \\ \operatorname{Res}[f(2i)] &= \frac{1}{(2 - 1)!} \lim_{z \rightarrow 2i} \frac{d}{dz} \left[ (z - 2i)^2 \frac{z^2}{(z^2 + 1)(z + 2i)^2(z - 2i)^2} \right] \\ &= \lim_{z \rightarrow 2i} \frac{d}{dz} \left[ \frac{z^2}{(z^2 + 1)(z + 2i)^2} \right] = \frac{5}{72i},\end{aligned}$$

and

$$\int_{-\infty}^{\infty} \frac{x^2 dx}{(x^2 + 1)(x^2 + 4)^2} = 2\pi i \left( -\frac{1}{18i} + \frac{5}{72i} \right) = \frac{\pi}{36}.$$

Closing the contour in the LHP would yield the same result. ■

### 10.3.2 Products of Rational and Trigonometric Functions

The second type of integral we can evaluate using the residue theorem is of the form

$$\int_{-\infty}^{\infty} \frac{p(x)}{q(x)} \cos ax \, dx \quad \text{or} \quad \int_{-\infty}^{\infty} \frac{p(x)}{q(x)} \sin ax \, dx,$$

where  $a$  is a real number,  $p(x)$  and  $q(x)$  are real polynomials in  $x$ , and  $q(x)$  has no real zeros. These integrals are the real and imaginary parts of

$$I_2 = \int_{-\infty}^{\infty} \frac{p(x)}{q(x)} e^{iax} \, dx.$$

The presence of  $e^{iax}$  dictates the choice of the half-plane: If  $a \geq 0$ , we choose the UHP; otherwise, we choose the LHP. We must, of course, have enough powers of  $x$  in the denominator to render  $R|p(Re^{i\theta})/q(Re^{i\theta})|$  uniformly convergent to zero.

**10.3.4. Example.** Let us evaluate  $\int_{-\infty}^{\infty} [\cos ax/(x^2+1)^2] dx$  where  $a \neq 0$ . This integral is the real part of the integral  $I_2 = \int_{-\infty}^{\infty} e^{iax} dx/(x^2+1)^2$ . When  $a > 0$ , we close in the UHP as advised by Jordan's lemma. Then we proceed as for integrals of rational functions. Thus, we have

$$I_2 = \oint_C \frac{e^{iaz}}{(z^2+1)^2} dz = 2\pi i \operatorname{Res}[f(i)] \quad \text{for } a > 0$$

because there is only one pole (of order 2) in the UHP at  $z = i$ . We next calculate the residue:

$$\begin{aligned} \operatorname{Res}[f(i)] &= \lim_{z \rightarrow i} \frac{d}{dz} \left[ (z-i)^2 \frac{e^{iaz}}{(z-i)^2(z+i)^2} \right] \\ &= \lim_{z \rightarrow i} \frac{d}{dz} \left[ \frac{e^{iaz}}{(z+i)^2} \right] = \lim_{z \rightarrow i} \left[ \frac{(z+i)iae^{iaz} - 2e^{iaz}}{(z+i)^3} \right] = \frac{e^{-a}}{4i}(1+a). \end{aligned}$$

Substituting this in the expression for  $I_2$ , we obtain  $I_2 = \frac{\pi}{2} e^{-a}(1+a)$  for  $a > 0$ .

When  $a < 0$ , we have to close the contour in the LHP, where the pole of order 2 is at  $z = -i$  and the contour is taken clockwise. Thus, we get

$$I_2 = \oint_C \frac{e^{iaz}}{(z^2+1)^2} dz = -2\pi i \operatorname{Res}[f(-i)] \quad \text{for } a < 0.$$

For the residue we obtain

$$\operatorname{Res}[f(-i)] = \lim_{z \rightarrow -i} \frac{d}{dz} \left[ (z+i)^2 \frac{e^{iaz}}{(z-i)^2(z+i)^2} \right] = -\frac{e^a}{4i}(1-a),$$

and the expression for  $I_2$  becomes  $I_2 = \frac{\pi}{2} e^a(1-a)$  for  $a < 0$ . We can combine the two results and write

$$\int_{-\infty}^{\infty} \frac{\cos ax}{(x^2+1)^2} dx = \operatorname{Re}(I_2) = I_2 = \frac{\pi}{2} (1+|a|)e^{-|a|}. \quad \blacksquare$$

**10.3.5. Example.** As another example, let us evaluate

$$\int_{-\infty}^{\infty} \frac{x \sin ax}{x^4+4} dx \quad \text{where } a \neq 0.$$

This is the imaginary part of the integral  $I_2 = \int_{-\infty}^{\infty} x e^{iax} dx/(x^4+4)$ , which, in terms of  $z$  and for the closed contour in the UHP (when  $a > 0$ ), becomes

$$I_2 = \oint_C \frac{z e^{iaz}}{z^4+4} dz = 2\pi i \sum_{j=1}^m \operatorname{Res}[f(z_j)] \quad \text{for } a > 0. \quad (10.6)$$

The singularities are determined by the zeros of the denominator:  $z^4 + 4 = 0$ , or  $z = 1 \pm i, -1 \pm i$ . Of these four simple poles only two,  $1 + i$  and  $-1 + i$ , are in the UHP. We now calculate the residues:

$$\begin{aligned} \operatorname{Res}[f(1+i)] &= \lim_{z \rightarrow 1+i} (z-1-i) \frac{ze^{iaz}}{(z-1-i)(z-1+i)(z+1-i)(z+1+i)} \\ &= \frac{(1+i)e^{ia(1+i)}}{(2i)(2)(2+2i)} = \frac{e^{ia}e^{-a}}{8i}, \end{aligned}$$

$$\begin{aligned} \operatorname{Res}[f(-1+i)] &= \lim_{z \rightarrow -1+i} (z+1-i) \frac{ze^{iaz}}{(z+1-i)(z+1+i)(z-1-i)(z-1+i)} \\ &= \frac{(-1+i)e^{ia(-1+i)}}{(2i)(-2)(-2+2i)} = -\frac{e^{-ia}e^{-a}}{8i}. \end{aligned}$$

Substituting in Equation (10.6), we obtain

$$I_2 = 2\pi i \frac{e^{-a}}{8i} (e^{ia} - e^{-ia}) = i \frac{\pi}{2} e^{-a} \sin a.$$

Thus,

$$\int_{-\infty}^{\infty} \frac{x \sin ax}{x^4 + 4} dx = \operatorname{Im}(I_2) = \frac{\pi}{2} e^{-a} \sin a \quad \text{for } a > 0. \quad (10.7)$$

For  $a < 0$ , we could close the contour in the LHP. But there is an easier way of getting to the answer. We note that  $-a > 0$ , and Equation (10.7) yields

$$\int_{-\infty}^{\infty} \frac{x \sin ax}{x^4 + 4} dx = - \int_{-\infty}^{\infty} \frac{x \sin[(-a)x]}{x^4 + 4} dx = -\frac{\pi}{2} e^{-(-a)} \sin(-a) = \frac{\pi}{2} e^a \sin a.$$

We can collect the two cases in

$$\int_{-\infty}^{\infty} \frac{x \sin ax}{x^4 + 4} dx = \frac{\pi}{2} e^{-|a|} \sin a. \quad \blacksquare$$

### 10.3.3 Functions of Trigonometric Functions

The third type of integral we can evaluate using the residue theorem involves only trigonometric functions and is of the form

$$\int_0^{2\pi} F(\sin \theta, \cos \theta) d\theta,$$

where  $F$  is some (typically rational) function of its arguments. Since  $\theta$  varies from 0 to  $2\pi$ , we can consider it an argument of a point  $z$  on the unit circle centered at the origin. Then  $z = e^{i\theta}$  and  $e^{-i\theta} = 1/z$ , and we can substitute  $\cos \theta = (z + 1/z)/2$ ,  $\sin \theta = (z - 1/z)/(2i)$ , and  $d\theta = dz/(iz)$  in the original integral, to obtain

$$\oint_C F\left(\frac{z-1/z}{2i}, \frac{z+1/z}{2}\right) \frac{dz}{iz}.$$

This integral can often be evaluated using the method of residues.

**10.3.6. Example.** Let us evaluate the integral  $\int_0^{2\pi} d\theta/(1 + a \cos \theta)$  where  $|a| < 1$ . Substituting for  $\cos \theta$  and  $d\theta$  in terms of  $z$ , we obtain

$$\oint_C \frac{dz/iz}{1 + a[(z^2 + 1)/(2z)]} = \frac{2}{i} \oint_C \frac{dz}{2z + az^2 + a},$$

where  $C$  is the unit circle centered at the origin. The singularities of the integrand are the zeros of its denominator:

$$z_1 = \frac{-1 + \sqrt{1 - a^2}}{a} \quad \text{and} \quad z_2 = \frac{-1 - \sqrt{1 - a^2}}{a}.$$

For  $|a| < 1$  it is clear that  $z_2$  will lie outside the unit circle  $C$ ; therefore, it does not contribute to the integral. But  $z_1$  lies inside, and we obtain

$$\oint_C \frac{dz}{2z + az^2 + a} = 2\pi i \operatorname{Res}[f(z_1)].$$

The residue of the simple pole at  $z_1$  can be calculated:

$$\begin{aligned} \operatorname{Res}[f(z_1)] &= \lim_{z \rightarrow z_1} (z - z_1) \frac{1}{a(z - z_1)(z - z_2)} = \frac{1}{a} \left( \frac{1}{z_1 - z_2} \right) \\ &= \frac{1}{a} \left( \frac{a}{2\sqrt{1 - a^2}} \right) = \frac{1}{2\sqrt{1 - a^2}}. \end{aligned}$$

It follows that

$$\int_0^{2\pi} \frac{d\theta}{1 + a \cos \theta} = \frac{2}{i} \oint_C \frac{dz}{2z + az^2 + a} = \frac{2}{i} 2\pi i \left( \frac{1}{2\sqrt{1 - a^2}} \right) = \frac{2\pi}{\sqrt{1 - a^2}}. \quad \blacksquare$$

**10.3.7. Example.** As another example, let us consider the integral

$$I = \int_0^\pi \frac{d\theta}{(a + \cos \theta)^2} \quad \text{where } a > 1.$$

Since  $\cos \theta$  is an even function of  $\theta$ , we may write

$$I = \frac{1}{2} \int_{-\pi}^\pi \frac{d\theta}{(a + \cos \theta)^2} \quad \text{where } a > 1.$$

This integration is over a complete cycle around the origin, and we can make the usual substitution:

$$I = \frac{1}{2} \oint_C \frac{dz/iz}{[a + (z^2 + 1)/2z]^2} = \frac{2}{i} \oint_C \frac{z dz}{(z^2 + 2az + 1)^2}.$$

The denominator has the roots  $z_1 = -a + \sqrt{a^2 - 1}$  and  $z_2 = -a - \sqrt{a^2 - 1}$ , which are both of order 2. The second root is outside the unit circle because  $a > 1$ . Also, it is easily verified that for all  $a > 1$ ,  $z_1$  is inside the unit circle. Since  $z_1$  is a pole of order 2, we have

$$\begin{aligned} \operatorname{Res}[f(z_1)] &= \lim_{z \rightarrow z_1} \frac{d}{dz} \left[ (z - z_1)^2 \frac{z}{(z - z_1)^2 (z - z_2)^2} \right] \\ &= \lim_{z \rightarrow z_1} \frac{d}{dz} \left[ \frac{z}{(z - z_2)^2} \right] = \frac{1}{(z_1 - z_2)^2} - \frac{2z_1}{(z_1 - z_2)^3} = \frac{a}{4(a^2 - 1)^{3/2}}. \end{aligned}$$

$$\text{We thus obtain } I = \frac{2}{i} 2\pi i \operatorname{Res}[f(z_1)] = \frac{\pi a}{(a^2 - 1)^{3/2}}. \quad \blacksquare$$

### 10.3.4 Some Other Integrals

The three types of definite integrals discussed above do not exhaust all possible applications of the residue theorem. There are other integrals that do not fit into any of the foregoing three categories but are still manageable. As the next two examples demonstrate, an ingenious choice of contours allows evaluation of other types of integrals.

**10.3.8. Example.** Let us evaluate the Gaussian integral

$$I = \int_{-\infty}^{\infty} e^{iax-bx^2} dx \quad \text{where } a, b \in \mathbb{R}, b > 0.$$

Completing squares in the exponent, we have

$$I = \int_{-\infty}^{\infty} e^{-b[x-ia/(2b)]^2 - a^2/4b} dx = e^{-a^2/4b} \lim_{R \rightarrow \infty} \int_{-R}^R e^{-b[x-ia/(2b)]^2} dx.$$

If we change the variable of integration to  $z = x - ia/(2b)$ , we obtain

$$I = e^{-a^2/(4b)} \lim_{R \rightarrow \infty} \int_{-R-ia/(2b)}^{R-ia/(2b)} e^{-bz^2} dz.$$

Let us now define  $I_R$ :

$$I_R \equiv \int_{-R-ia/(2b)}^{R-ia/(2b)} e^{-bz^2} dz.$$

This is an integral along a straight line  $C_1$  that is parallel to the  $x$ -axis (see Figure 10.4). We close the contour as shown and note that  $e^{-bz^2}$  is analytic throughout the interior of the closed contour (it is an entire function!). Thus, the contour integral must vanish by the Cauchy-Goursat theorem. So we obtain

$$I_R + \int_{C_3} e^{-bz^2} dz + \int_R^{-R} e^{-bx^2} dx + \int_{C_4} e^{-bz^2} dz = 0.$$

Along  $C_3$ ,  $z = R + iy$  and

$$\int_{C_3} e^{-bz^2} dz = \int_{-ia/(2b)}^0 e^{-b(R+iy)^2} i dy = ie^{-bR^2} \int_{-ia/(2b)}^0 e^{by^2 - 2ibRy} dy$$

which clearly tends to zero as  $R \rightarrow \infty$ . We get a similar result for the integral along  $C_4$ . Therefore, we have

$$I_R = \int_{-R}^R e^{-bx^2} dx \Rightarrow \lim_{R \rightarrow \infty} I_R = \int_{-\infty}^{\infty} e^{-bx^2} dx = \sqrt{\frac{\pi}{b}}.$$

Finally, we get

$$\int_{-\infty}^{\infty} e^{iax-bx^2} dx = \sqrt{\frac{\pi}{b}} e^{-a^2/(4b)}.$$

■

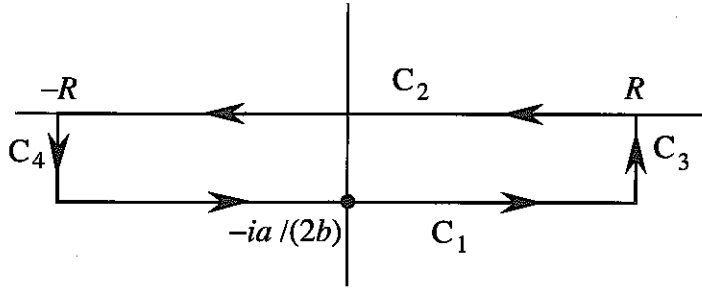


Figure 10.4 The contour for the evaluation of the Gaussian integral.

**10.3.9. Example.** Let us evaluate  $I = \int_0^\infty dx/(x^3 + 1)$ . If the integrand were even, we could extend the lower limit of integration to  $-\infty$  and close the contour in the UHP. Since this is not the case, we need to use a different trick. To get a hint as to how to close the contour, we study the singularities of the integrand. These are simply the roots of the denominator:  $z^3 = -1$  or  $z_n = e^{i(2n+1)\pi/3}$  with  $n = 0, 1, 2$ . These, as well as a contour that has only  $z_0$  as an interior point, are shown in Figure 10.5. We thus have

$$I + \int_{C_R} \frac{dz}{z^3 + 1} + \int_{C_2} \frac{dz}{z^3 + 1} = 2\pi i \operatorname{Res}[f(z_0)]. \tag{10.8}$$

The  $C_R$  integral vanishes, as usual. Along  $C_2$ ,  $z = re^{i\alpha}$ , with constant  $\alpha$ , so that  $dz = e^{i\alpha} dr$  and

$$\int_{C_2} \frac{dz}{z^3 + 1} = \int_\infty^0 \frac{e^{i\alpha} dr}{(re^{i\alpha})^3 + 1} = -e^{i\alpha} \int_0^\infty \frac{dr}{r^3 e^{3i\alpha} + 1}.$$

In particular, if we choose  $3\alpha = 2\pi$ , we obtain

$$\int_{C_2} \frac{dz}{z^3 + 1} = -e^{i2\pi/3} \int_0^\infty \frac{dr}{r^3 + 1} = -e^{i2\pi/3} I.$$

Substituting this in Equation (10.8) gives

$$(1 - e^{i2\pi/3})I = 2\pi i \operatorname{Res}[f(z_0)] \Rightarrow I = \frac{2\pi i}{1 - e^{i2\pi/3}} \operatorname{Res}[f(z_0)].$$

On the other hand,

$$\begin{aligned} \operatorname{Res}[f(z_0)] &= \lim_{z \rightarrow z_0} (z - z_0) \frac{1}{(z - z_0)(z - z_1)(z - z_2)} \\ &= \frac{1}{(z_0 - z_1)(z_0 - z_2)} = \frac{1}{(e^{i\pi/3} - e^{i\pi})(e^{i\pi/3} - e^{i5\pi/3})}. \end{aligned}$$

These last two equations yield

$$I = \frac{2\pi i}{1 - e^{i2\pi/3}} \frac{1}{(e^{i\pi/3} - e^{i\pi})(e^{i\pi/3} - e^{i5\pi/3})} = \frac{2\pi}{3\sqrt{3}}. \quad \blacksquare$$

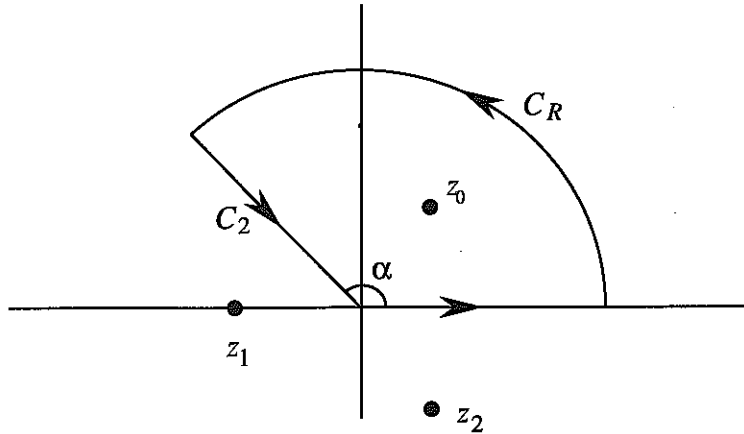


Figure 10.5 The contour is chosen so that only one of the poles lies inside.

### 10.3.5 Principal Value of an Integral

So far we have discussed only integrals of functions that have no singularities on the contour. Let us now investigate the consequences of the presence of singular points on the contour. Consider the integral

$$\int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx, \tag{10.9}$$

where  $x_0$  is a real number and  $f$  is analytic at  $x_0$ . To avoid  $x_0$ —which causes the integrand to diverge—we bypass it by indenting the contour as shown in Figure 10.6 and denoting the new contour by  $C_u$ . The contour  $C_0$  is simply a semicircle of radius  $\epsilon$ . For the contour  $C_u$ , we have

$$\int_{C_u} \frac{f(z)}{z - x_0} dz = \int_{-\infty}^{x_0 - \epsilon} \frac{f(x)}{x - x_0} dx + \int_{x_0 + \epsilon}^{\infty} \frac{f(x)}{x - x_0} dx + \int_{C_0} \frac{f(z)}{z - x_0} dz.$$

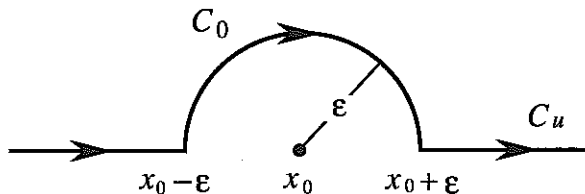
In the limit  $\epsilon \rightarrow 0$ , the sum of the first two terms on the RHS—when it exists—defines the **principal value** of the integral in Equation (10.9):

principal value of an integral

$$P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx = \lim_{\epsilon \rightarrow 0} \left[ \int_{-\infty}^{x_0 - \epsilon} \frac{f(x)}{x - x_0} dx + \int_{x_0 + \epsilon}^{\infty} \frac{f(x)}{x - x_0} dx \right].$$

The integral over the semicircle is calculated by noting that  $z - x_0 = \epsilon e^{i\theta}$  and  $dz = i\epsilon e^{i\theta} d\theta$ :  $\int_{C_0} f(z) dz / (z - x_0) = -i\pi f(x_0)$ . Therefore,

$$\int_{C_u} \frac{f(z)}{z - x_0} dz = P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx - i\pi f(x_0). \tag{10.10}$$



**Figure 10.6** The contour  $C_u$  avoids  $x_0$ .

On the other hand, if  $C_0$  is taken below the singularity on a contour  $C_d$ , say, we obtain

$$\int_{C_d} \frac{f(z)}{z - x_0} dz = P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx + i\pi f(x_0).$$

We see that the contour integral depends on how the singular point  $x_0$  is avoided. However, the principal value, if it exists, is unique. To calculate this principal value we close the contour by adding a large semicircle to it as before, assuming that the contribution from this semicircle goes to zero by Jordan's lemma. The contours  $C_u$  and  $C_d$  are replaced by a closed contour, and the value of the integral will be given by the residue theorem. We therefore have

$$P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx = \pm i\pi f(x_0) + 2\pi i \sum_{j=1}^m \operatorname{Res} \left[ \frac{f(z_j)}{z_j - x_0} \right], \tag{10.11}$$

where the plus sign corresponds to placing the infinitesimal semicircle in the UHP, as shown in Figure 10.6, and the minus sign corresponds to the other choice.

**10.3.10. Example.** Let us use the principal-value method to evaluate the integral

$$I = \int_0^{\infty} \frac{\sin x}{x} dx = \frac{1}{2} \int_{-\infty}^{\infty} \frac{\sin x}{x} dx.$$

It appears that  $x = 0$  is a singular point of the integrand; in reality, however, it is only a removable singularity, as can be verified by the Taylor expansion of  $\sin x/x$ . To make use of the principal-value method, we write

$$I = \frac{1}{2} \operatorname{Im} \left( \int_{-\infty}^{\infty} \frac{e^{ix}}{x} dx \right) = \frac{1}{2} \operatorname{Im} \left( P \int_{-\infty}^{\infty} \frac{e^{ix}}{x} dx \right).$$

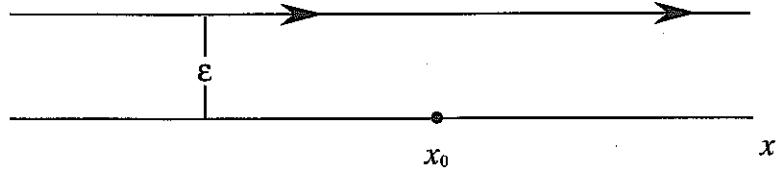
We now use Equation (10.11) with the small circle in the UHP, noting that there are no singularities for  $e^{ix}/x$  there. This yields

$$P \int_{-\infty}^{\infty} \frac{e^{ix}}{x} dx = i\pi e^{(0)} = i\pi.$$

Therefore,

$$\int_0^{\infty} \frac{\sin x}{x} dx = \frac{1}{2} \operatorname{Im}(i\pi) = \frac{\pi}{2}. \quad \blacksquare$$





**Figure 10.7** The equivalent contour obtained by “stretching”  $C_u$ , the contour of Figure 10.6.

The principal value of an integral can be written more compactly if we deform the contour  $C_u$  by stretching it into that shown in Figure 10.7. For small enough  $\epsilon$ , such a deformation will not change the number of singularities within the infinite closed contour. Thus, the LHS of Equation (10.10) will have limits of integration  $-\infty + i\epsilon$  and  $+\infty + i\epsilon$ . If we change the variable of integration to  $\xi = z - i\epsilon$ , this integral becomes

$$\int_{-\infty}^{\infty} \frac{f(\xi + i\epsilon)}{\xi + i\epsilon - x_0} d\xi = \int_{-\infty}^{\infty} \frac{f(\xi) d\xi}{\xi - x_0 + i\epsilon} = \int_{-\infty}^{\infty} \frac{f(z) dz}{z - x_0 + i\epsilon}, \quad (10.12)$$

where in the last step we changed the dummy integration variable back to  $z$ . Note that since  $f$  is assumed to be continuous at all points on the contour,  $f(\xi + i\epsilon) \rightarrow f(\xi)$  for small  $\epsilon$ . The last integral of Equation (10.12) shows that there is no singularity on the new  $x$ -axis; we have pushed the singularity down to  $x_0 - i\epsilon$ . In other words, we have given the singularity on the  $x$ -axis a small negative imaginary part. We can thus rewrite Equation (10.10) as

$$P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx = i\pi f(x_0) + \int_{-\infty}^{\infty} \frac{f(x) dx}{x - x_0 + i\epsilon},$$

where  $x$  is used instead of  $z$  in the last integral because we are indeed integrating along the new  $x$ -axis—assuming that no other singularities are present in the UHP. A similar argument, this time for the LHP, introduces a minus sign for the first term on the RHS and for the  $\epsilon$  term in the denominator. Therefore,

$$P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx = \pm i\pi f(x_0) + \int_{-\infty}^{\infty} \frac{f(x) dx}{x - x_0 \pm i\epsilon}, \quad (10.13)$$

where the plus (minus) sign refers to the UHP (LHP). This result is sometimes abbreviated as

$$\frac{1}{x - x_0 \pm i\epsilon} = P \frac{1}{x - x_0} dx \mp i\pi \delta(x - x_0). \quad (10.14)$$

**10.3.11. Example.** Let us use residues to evaluate the function

$$f(k) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{ikx} dx}{x - i\epsilon}, \quad \epsilon > 0.$$

The integral representation of the  $\theta$  (step) function

We have to close the contour by adding a large semicircle. Whether we do this in the UHP or the LHP is dictated by the sign of  $k$ : If  $k > 0$ , we close in the UHP. Thus,

$$\begin{aligned} f(k) &= \frac{1}{2\pi i} \int_C \frac{e^{ikz} dz}{z - i\epsilon} = \text{Res} \left[ \frac{e^{ikz}}{z - i\epsilon} \right]_{z \rightarrow i\epsilon} \\ &= \lim_{z \rightarrow i\epsilon} \left[ (z - i\epsilon) \frac{e^{ikz}}{z - i\epsilon} \right] = e^{-k\epsilon} \xrightarrow{\epsilon \rightarrow 0} 1. \end{aligned}$$

On the other hand, if  $k < 0$ , we must close in the LHP, in which the integrand is analytic. Thus, by the Cauchy-Goursat theorem, the integral vanishes. Therefore, we have

$$f(k) = \begin{cases} 1 & \text{if } k > 0, \\ 0 & \text{if } k < 0. \end{cases}$$

theta (or step) function

This is precisely the definition of the **theta function** (or step function). Thus, we have obtained an integral representation of that function:

$$\theta(x) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{ixt}}{t - i\epsilon} dt. \quad \blacksquare$$

Now suppose that there are two singular points on the real axis, at  $x_1$  and  $x_2$ . Let us avoid  $x_1$  and  $x_2$  by making little semicircles, as before, letting both semicircles be in the UHP (see Figure 10.8). Without writing the integrands, we can represent the contour integral by

$$\int_{-\infty}^{x_1-\epsilon} + \int_{C_1} + \int_{x_1+\epsilon}^{x_2-\epsilon} + \int_{C_2} + \int_{x_2+\epsilon}^{\infty} + \int_{C_R} = 2\pi i \sum \text{Res}.$$

The principal value of the integral is naturally defined to be the sum of all integrals having  $\epsilon$  in their limits. The contribution from the small semicircle  $C_1$  can be calculated by substituting  $z - x_1 = \epsilon e^{i\theta}$  in the integral:

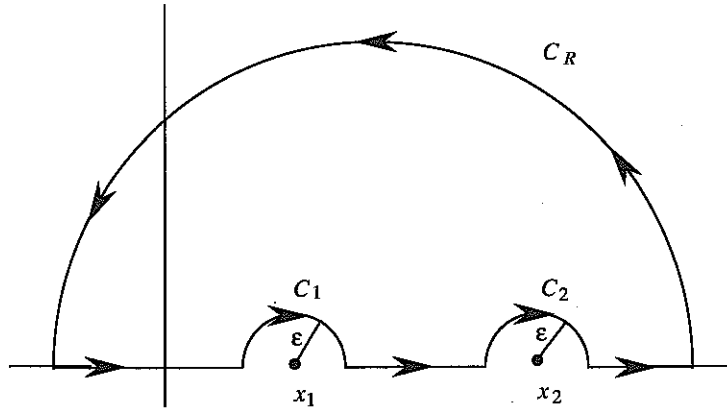
$$\int_{C_1} \frac{f(z) dz}{(z - x_1)(z - x_2)} = \int_{\pi}^0 \frac{f(x_1 + \epsilon e^{i\theta}) i \epsilon e^{i\theta} d\theta}{\epsilon e^{i\theta} (x_1 + \epsilon e^{i\theta} - x_2)} = -i\pi \frac{f(x_1)}{x_1 - x_2},$$

with a similar result for  $C_2$ . Putting everything together, we get

$$P \int_{-\infty}^{\infty} \frac{f(x)}{(x - x_1)(x - x_2)} dx - i\pi \frac{f(x_2) - f(x_1)}{x_2 - x_1} = 2\pi i \sum \text{Res}.$$

If we include the case where both  $C_1$  and  $C_2$  are in the LHP, we get

$$P \int_{-\infty}^{\infty} \frac{f(x)}{(x - x_1)(x - x_2)} dx = \pm i\pi \frac{f(x_2) - f(x_1)}{x_2 - x_1} + 2\pi i \sum \text{Res}, \quad (10.15)$$



**Figure 10.8** One of the four choices of contours for evaluating the principal value of the integral when there are two poles on the real axis.

where the plus sign is for the case where  $C_1$  and  $C_2$  are in the UHP and the minus sign for the case where both are in the LHP. We can also obtain the result for the case where the two singularities coincide by taking the limit  $x_1 \rightarrow x_2$ . Then the RHS of the last equation becomes a derivative, and we obtain

$$P \int_{-\infty}^{\infty} \frac{f(x)}{(x - x_0)^2} dx = \pm i\pi f'(x_0) + 2\pi i \sum \text{Res}.$$

**10.3.12. Example.** An expression encountered in the study of Green’s functions or propagators (which we shall discuss later in the book) is

$$\int_{-\infty}^{\infty} \frac{e^{itx} dx}{x^2 - k^2},$$

where  $k$  and  $t$  are real constants. We want to calculate the principal value of this integral. We use Equation (10.15) and note that for  $t > 0$ , we need to close the contour in the UHP, where there are no poles:

$$P \int_{-\infty}^{\infty} \frac{e^{itx} dx}{x^2 - k^2} = P \int_{-\infty}^{\infty} \frac{e^{itx} dx}{(x - k)(x + k)} = i\pi \frac{e^{ikt} - e^{-ikt}}{2k} = -\pi \frac{\sin kt}{k}.$$

When  $t < 0$ , we have to close the contour in the LHP, where again there are no poles:

$$P \int_{-\infty}^{\infty} \frac{e^{itx} dx}{x^2 - k^2} = P \int_{-\infty}^{\infty} \frac{e^{itx} dx}{(x - k)(x + k)} = -i\pi \frac{e^{ikt} - e^{-ikt}}{2k} = \pi \frac{\sin kt}{k}.$$

The two results above can be combined into a single relation:

$$P \int_{-\infty}^{\infty} \frac{e^{itx} dx}{x^2 - k^2} = -\pi \frac{\sin k|t|}{k}.$$



## 10.4 Problems

10.1. Evaluate each of the following integrals, for all of which  $C$  is the circle  $|z| = 3$ .

$$\begin{array}{lll}
 (a) \oint_C \frac{4z-3}{z(z-2)} dz. & (b) \oint_C \frac{e^z}{z(z-i\pi)} dz. & (c) \oint_C \frac{\cos z}{z(z-\pi)} dz. \\
 (d) \oint_C \frac{z^2+1}{z(z-1)} dz. & (e) \oint_C \frac{\cosh z}{z^2+\pi^2} dz. & (f) \oint_C \frac{1-\cos z}{z^2} dz. \\
 (g) \oint_C \frac{\sinh z}{z^4} dz. & (h) \oint_C z \cos\left(\frac{1}{z}\right) dz. & (i) \oint_C \frac{dz}{z^3(z+5)}. \\
 (j) \oint_C \tan z dz. & (k) \oint_C \frac{dz}{\sinh 2z}. & (l) \oint_C \frac{e^z}{z^2} dz. \\
 (m) \oint_C \frac{dz}{z^2 \sin z}. & (n) \oint_C \frac{e^z dz}{(z-1)(z-2)}. & 
 \end{array}$$

10.2. Let  $h(z)$  be analytic and have a simple zero at  $z = z_0$ , and let  $g(z)$  be analytic there. Let  $f(z) = g(z)/h(z)$ , and show that

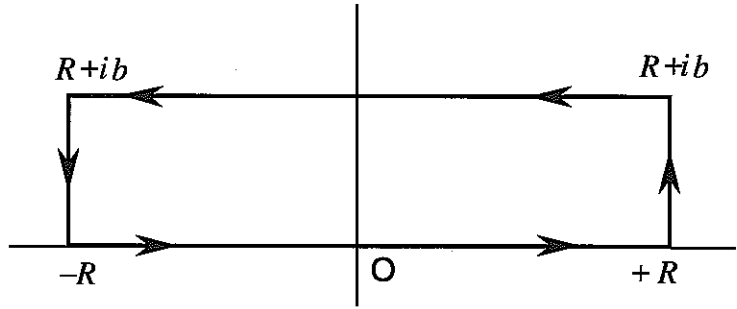
$$\operatorname{Res}[f(z_0)] = \frac{g(z_0)}{h'(z_0)}.$$

10.3. Find the residue of  $f(z) = 1/\cos z$  at each of its poles.

10.4. Evaluate the integral  $\int_0^\infty dx/[(x^2+1)(x^2+4)]$  by closing the contour (a) in the UHP and (b) in the LHP.

10.5. Evaluate the following integrals, in which  $a$  and  $b$  are nonzero real constants.

$$\begin{array}{lll}
 (a) \int_0^\infty \frac{2x^2+1}{x^4+5x^2+6} dx. & (b) \int_0^\infty \frac{dx}{6x^4+5x^2+1}. & (c) \int_0^\infty \frac{dx}{x^4+1}. \\
 (d) \int_0^\infty \frac{\cos x dx}{(x^2+a^2)^2(x^2+b^2)}. & (e) \int_0^\infty \frac{\cos ax}{(x^2+b^2)^2} dx. & (f) \int_0^\infty \frac{dx}{(x^2+1)^2}. \\
 (g) \int_0^\infty \frac{dx}{(x^2+1)^2(x^2+2)}. & (h) \int_0^\infty \frac{2x^2-1}{x^6+1} dx. & (i) \int_0^\infty \frac{x^2 dx}{(x^2+a^2)^2}. \\
 (j) \int_{-\infty}^\infty \frac{x dx}{(x^2+4x+13)^2}. & (k) \int_0^\infty \frac{x^3 \sin ax}{x^6+1} dx. & (l) \int_0^\infty \frac{x^2+1}{x^2+4} dx. \\
 (m) \int_{-\infty}^\infty \frac{x \cos x dx}{x^2-2x+10}. & (n) \int_{-\infty}^\infty \frac{x \sin x dx}{x^2-2x+10}. & (o) \int_0^\infty \frac{dx}{x^2+1}. \\
 (p) \int_0^\infty \frac{x^2 dx}{(x^2+4)^2(x^2+25)}. & (q) \int_0^\infty \frac{\cos ax}{x^2+b^2} dx. & (r) \int_0^\infty \frac{dx}{(x^2+4)^2}.
 \end{array}$$



**Figure 10.9** The contour used in Problem 10.8.

**10.6.** Evaluate each of the following integrals by turning it into a contour integral around a unit circle.

$$(a) \int_0^{2\pi} \frac{d\theta}{5 + 4 \sin \theta}.$$

$$(b) \int_0^{2\pi} \frac{d\theta}{a + \cos \theta} \quad \text{where } a > 1.$$

$$(c) \int_0^{2\pi} \frac{d\theta}{1 + \sin^2 \theta}.$$

$$(d) \int_0^{2\pi} \frac{d\theta}{(a + b \cos^2 \theta)^2} \quad \text{where } a, b > 0.$$

$$(e) \int_0^{2\pi} \frac{\cos^2 3\theta}{5 - 4 \cos 2\theta} d\theta.$$

$$(f) \int_0^\pi \frac{d\phi}{1 - 2a \cos \phi + a^2} \quad \text{where } a \neq \pm 1.$$

$$(g) \int_0^\pi \frac{\cos^2 3\phi d\phi}{1 - 2a \cos \phi + a^2} \quad \text{where } a \neq \pm 1.$$

$$(h) \int_0^\pi \frac{\cos 2\phi d\phi}{1 - 2a \cos \phi + a^2} \quad \text{where } a \neq \pm 1.$$

$$(i) \int_0^\pi \tan(x + ia) dx \quad \text{where } a \in \mathbb{R}.$$

$$(j) \int_0^\pi e^{\cos \phi} \cos(n\phi - \sin \phi) d\phi \quad \text{where } n \in \mathbb{Z}.$$

**10.7.** Evaluate the integral  $I = \int_{-\infty}^{\infty} e^{\alpha x} dx / (1 + e^x)$  for  $0 < \alpha < 1$ . Hint: Choose a closed (long) rectangle that encloses only one of the zeros of the denominator. Show that the contributions of the short sides of the rectangle are zero.

**10.8.** Derive the integration formula  $\int_0^\infty e^{-x^2} \cos(2bx) dx = \frac{\sqrt{\pi}}{2} e^{-b^2}$  where  $b \neq 0$  by integrating the function  $e^{-z^2}$  around the rectangular path shown in Figure 10.9.

**10.9.** Use the result of Example 10.3.11 to show that  $\theta'(k) = \delta(k)$ .

10.10. Find the principal values of the following integrals.

$$(a) \int_{-\infty}^{\infty} \frac{\sin x \, dx}{(x^2 + 4)(x - 1)}, \quad (b) \int_{-\infty}^{\infty} \frac{\cos ax}{1 + x^3} \, dx \quad \text{where } a \geq 0.$$

$$(c) \int_{-\infty}^{\infty} \frac{x \cos x}{x^2 - 5x + 6} \, dx, \quad (d) \int_{-\infty}^{\infty} \frac{1 - \cos x}{x^2} \, dx.$$

10.11. Evaluate the following integrals.

$$(a) \int_0^{\infty} \frac{x^2 - b^2}{x^2 + b^2} \left( \frac{\sin ax}{x} \right) dx, \quad (b) \int_0^{\infty} \frac{\sin ax}{x(x^2 + b^2)} dx.$$

$$(c) \int_0^{\infty} \frac{\sin ax}{x(x^2 + b^2)^2} dx, \quad (d) \int_0^{\infty} \frac{\cos 2ax - \cos 2bx}{x^2} dx.$$

$$(e) \int_0^{\infty} \frac{\sin^2 x \, dx}{x^2}, \quad (f) \int_0^{\infty} \frac{\sin^3 x \, dx}{x^3}.$$

## Additional Reading

1. Denery, P. and Krzywicki, A. *Mathematics for Physicists*, Harper and Row, 1967. Includes a detailed discussion of complex analysis encompassing applications of conformal mappings and the residue theorem to physical problems.
2. Mathews, J. and Walker, R. *Mathematical Methods of Physics*, 2nd ed., Benjamin, 1970. A "practical" guide to mathematical physics, including a long discussion of "how to evaluate integrals" and the use of the residue theorem.

# Complex Analysis: Advanced Topics

The subject of complex analysis is an extremely rich and powerful area of mathematics. We have already seen some of this richness and power in the previous chapter. This chapter concludes our discussion of complex analysis by introducing some other topics with varying degrees of importance.

## 11.1 Meromorphic Functions

Complex functions that have only simple poles as their singularities are numerous in applications and are called **meromorphic functions**. In this section, we derive an important result for such functions.

Assume that  $f(z)$  has simple poles at  $\{z_j\}_{j=1}^N$ , where  $N$  could be infinity. Then, if  $z \neq z_j$  for all  $j$ , the residue theorem yields<sup>1</sup>

$$\frac{1}{2\pi i} \int_{C_n} \frac{f(\xi)}{\xi - z} d\xi = f(z) + \sum_{j=1}^n \operatorname{Res} \left( \frac{f(\xi)}{\xi - z} \right)_{\xi=z_j},$$

where  $C_n$  is a circle containing the first  $n$  poles, and it is assumed that the poles are arranged in order of increasing absolute values. Since the poles of  $f$  are assumed to be simple, we have

$$\begin{aligned} \operatorname{Res} \left( \frac{f(\xi)}{\xi - z} \right)_{\xi=z_j} &= \lim_{\xi \rightarrow z_j} (\xi - z_j) \frac{f(\xi)}{\xi - z} = \frac{1}{z_j - z} \lim_{\xi \rightarrow z_j} [(\xi - z_j) f(\xi)] \\ &= \frac{1}{z_j - z} \operatorname{Res}[f(\xi)]_{\xi=z_j} \equiv \frac{r_j}{z_j - z}, \end{aligned}$$

<sup>1</sup>Note that the residue of  $f(\xi)/(\xi - z)$  at  $\xi = z$  is simply  $f(z)$ .

where  $r_j$  is, by definition, the residue of  $f(\xi)$  at  $\xi = z_j$ . Substituting in the preceding equation gives

$$f(z) = \frac{1}{2\pi i} \int_{C_n} \frac{f(\xi)}{\xi - z} d\xi - \sum_{j=1}^n \frac{r_j}{z_j - z}.$$

Taking the difference between this and the same equation evaluated at  $z = 0$  (assumed to be none of the poles),<sup>2</sup> we can write

$$f(z) - f(0) = \frac{z}{2\pi i} \int_{C_n} \frac{f(\xi)}{\xi(\xi - z)} d\xi + \sum_{j=1}^n r_j \left( \frac{1}{z - z_j} + \frac{1}{z_j} \right).$$

If  $|f(\xi)|$  approaches a finite value as  $|\xi| \rightarrow \infty$ , the integral vanishes for an infinite circle (which includes all poles now), and we obtain what is called the **Mittag-Leffler expansion** of the meromorphic function  $f$ :

Mittag-Leffler expansion

$$f(z) = f(0) + \sum_{j=1}^N r_j \left( \frac{1}{z - z_j} + \frac{1}{z_j} \right). \tag{11.1}$$

Now we let  $g$  be an entire function with *simple zeros*. We claim that (a)  $(dg/dz)/g(z)$  is a meromorphic function that is bounded for all values of  $z$ , and (b) its residues are all unity. To see this, note that  $g$  is of the form<sup>3</sup>

$$g(z) = (z - z_1)(z - z_2) \cdots (z - z_N)f(z),$$

where  $z_1, \dots, z_N$  are all the zeros of  $g$ , and  $f$  is an analytic function that does not vanish anywhere in the complex plane. It is now easy to see that

$$\frac{g'(z)}{g(z)} = \sum_{j=1}^N \frac{1}{z - z_j} + \frac{f'(z)}{f(z)}.$$

This expression has both properties (a) and (b) mentioned above. Furthermore, the last term is an entire function that is bounded for all  $\mathbb{C}$ . Therefore, it must be a constant by Proposition 9.5.5. This derivation also verifies Equation (11.1), which in the case at hand can be written as

$$\frac{d}{dz} \ln g(z) = \frac{g'(z)}{g(z)} = \frac{d}{dz} \ln g(0) + \sum_{j=1}^N \left( \frac{1}{z - z_j} + \frac{1}{z_j} \right),$$

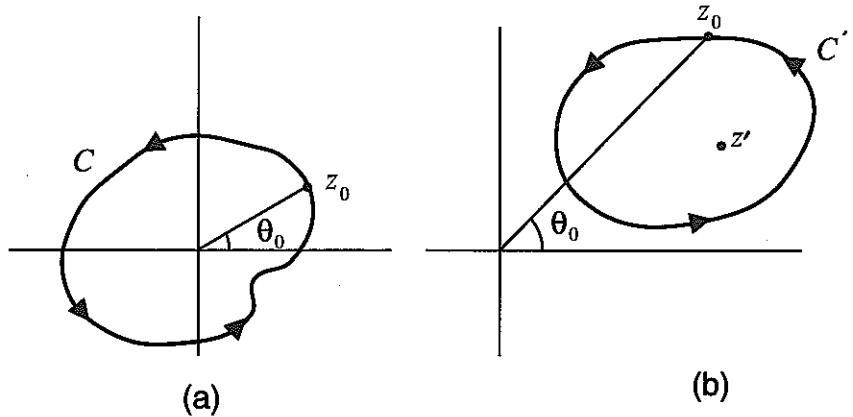
whose solution is readily found to be

$$g(z) = g(0)e^{cz} \prod_{j=1}^N \left( 1 - \frac{z}{z_j} \right) e^{z/z_j} \quad \text{where } c = \frac{(dg/dz)|_{z=0}}{g(0)} \tag{11.2}$$

and it is assumed that  $z_j \neq 0$  for all  $j$ .

<sup>2</sup>This is not a restrictive assumption because we can always move our coordinate system so that the origin avoids all poles.  
<sup>3</sup>One can "prove" this by factoring the simple zeros one by one, writing  $g(z) = (z - z_1)f_1(z)$  and noting that  $g(z_2) = 0$ , with  $z_2 \neq z_1$ , implies that  $f_1(z) = (z - z_2)f_2(z)$ , etc.





**Figure 11.1** (a) The angle  $\theta_0$  changes by  $2\pi$  as  $z_0$  makes a complete circuit around  $C$ . (b) The angle  $\theta_0$  returns to its original value when  $z_0$  completes the circuit.

## 11.2 Multivalued Functions

The arbitrariness, up to a multiple of  $2\pi$ , of the angle  $\theta = \arg(z)$  in  $z = re^{i\theta}$  leads to functions that can take different values at the same point. Consider, for example, the function  $f(z) = \sqrt{z}$ . Writing  $z$  in polar coordinates, we obtain  $f(z) = f(r, \theta) = (re^{i\theta})^{1/2} = \sqrt{r}e^{i\theta/2}$ . This shows that for the same  $z = (r, \theta) = (r, \theta + 2\pi)$ , we get two different values,  $f(r, \theta)$  and  $f(r, \theta + 2\pi) = -f(r, \theta)$ .

This may be disturbing at first. After all, the definition of a function (mapping) ensures that for any point in the domain a *unique* image is obtained. Here two different images are obtained for the same  $z$ . Riemann found a cure for this complex “double vision” by introducing what is now called Riemann sheets. We will discuss these briefly below, but first let us take a closer look at a prototype of multivalued functions. Consider the natural log function,  $\ln z$ . For  $z = re^{i\theta}$  this is defined as  $\ln z = \ln r + i\theta = \ln |z| + i \arg(z)$  where  $\arg(z)$  is defined only to within a multiple of  $2\pi$ ; that is,  $\arg(z) = \theta + 2n\pi$ , for  $n = 0, \pm 1, \pm 2, \dots$

We can see the peculiar nature of the logarithmic function by considering a closed path around the point  $z = 0$ , as shown in Figure 11.1(a). Starting at  $z_0$ , we move counterclockwise, noticing the constant increase in the angle  $\theta_0$ , until we reach the initial point in the  $z$ -plane. However, the angle is then  $\theta_0 + 2\pi$ . Thus, the process of moving around the origin has changed the value of the log function by  $2\pi i$ . Thus,  $(\ln z_0)_{\text{final}} - (\ln z_0)_{\text{initial}} = 2\pi i$ . Note that in this process  $z_0$  does not change, because

$$(z_0)_{\text{final}} = z = re^{i(\theta+2\pi)} = re^{i\theta} e^{2\pi i} = re^{i\theta} = (z_0)_{\text{initial}}.$$

branch point **11.2.1. Definition.** A *branch point* of a function  $f : \mathbb{C} \rightarrow \mathbb{C}$  is a complex number

$z_0$  with the property that  $f(r_0, \theta_0) \neq f(r_0, \theta_0 + 2\pi)$  for any closed curve  $C$  encircling  $z_0$ . Here  $(r_0, \theta_0)$  are the polar coordinates of  $z_0$ .

---

**Victor-Alexandre Puiseux** (1820–1883) was the first to take up the subject of multivalued functions. In 1850 Puiseux published a celebrated paper on complex algebraic functions given by  $f(u, z) = 0$ ,  $f$  a polynomial in  $u$  and  $z$ . He first made clear the distinction between **poles** and **branch points** that Cauchy had barely perceived, and introduced the notion of an **essential singular point**, to which Weierstrass independently had called attention. Though Cauchy, in the 1846 paper, did consider the variation of simple multivalued functions along paths that enclosed branch points, Puiseux clarified this subject too.

Puiseux also showed that the development of a function of  $z$  about a branch point  $z = a$  must involve *fractional powers* of  $z - a$ . He then improved on Cauchy's theorem on the expansion of a function in a Maclaurin series. By his significant investigations of many-valued functions and their branch points in the complex plane, and by his initial work on integrals of such functions, Puiseux brought Cauchy's pioneering work in function theory to the end of what might be called the first stage. The difficulties in the theory of multiple-valued functions and integrals of such functions were still to be overcome. Cauchy did write other papers on the integrals of multiplevalued functions in which he attempted to follow up on Puiseux's work; and though he introduced the notion of branch cuts (*lignes d'arrêt*), he was still confused about the distinction between poles and branch points. This subject of algebraic functions and their integrals was to be pursued by Riemann.

Puiseux was a keen mountaineer and was the first to scale the Alpine peak that is now named after him.

---

Thus,  $z = 0$  is a branch point of the logarithmic function. Studying the behavior of  $\ln(1/z) = -\ln z$  around  $z = 0$  will reveal that the point "at infinity" is also a branch point of  $\ln z$ . Figure 11.1(b) shows that any other point of the complex plane, such as  $z'$ , cannot be a branch point because  $\theta_0$  does not change when  $C'$  is traversed completely.

### 11.2.1 Riemann Surfaces

The idea of a Riemann surface begins with the removal of all points that lie on the line (or any other curve) joining two branch points. For  $\ln z$  this means the removal of all points lying on a curve that starts at  $z = 0$  and extends all the way to infinity. Such a curve is called a **branch cut**, or simply a *cut*.

branch cut or simply  
"cut"

Let us concentrate on  $\ln z$  and take the cut to be along the negative half of the real axis. Let us also define the functions

$$\begin{aligned} f_n(z) &= f_n(r, \theta) \\ &= \ln r + i(\theta + 2n\pi) \quad \text{for } -\pi < \theta < \pi; r > 0; n = 0, \pm 1, \dots, \end{aligned}$$

so  $f_n(z)$  takes on the same values for  $-\pi < \theta < \pi$  that  $\ln z$  takes in the range  $(2n-1)\pi < \theta < (2n+1)\pi$ . We have replaced the multivalued logarithmic function by a series of different functions that are analytic in the cut  $z$ -plane.

This process of cutting the  $z$ -plane and then defining a sequence of functions eliminates the contradiction caused by the existence of branch points, since we are no longer allowed to completely encircle a branch point. A complete circulation involves crossing the cut, which, in turn, violates the domain of definition of  $f_n(z)$ .

We have made good progress. We have replaced the (nonanalytic) multivalued function  $\ln z$  with a series of analytic (in their domain of definition) functions  $f_n(z)$ . However, there is a problem left:  $f_n(z)$  has a discontinuity at the cut. In fact, just above the cut  $f_n(r, \pi - \epsilon) = \ln r + i(\pi - \epsilon + 2n\pi)$  with  $\epsilon > 0$ , and just below it  $f_n(r, -\pi + \epsilon) = \ln r + i(-\pi + \epsilon + 2n\pi)$ , so that

$$\lim_{\epsilon \rightarrow 0} [f_n(r, \pi - \epsilon) - f_n(r, -\pi + \epsilon)] = 2\pi i.$$

To cure this we make the observation that the value of  $f_n(z)$  just above the cut is the same as the value of  $f_{n+1}(z)$  just below the cut. This suggests the following geometrical construction, due to Riemann: Superpose an infinite series of cut complex planes one on top of the other, each plane corresponding to a different value of  $n$ . The adjacent planes are connected along the cut such that the upper lip of the cut in the  $(n-1)$ th plane is connected to the lower lip of the cut in the  $n$ th plane. All planes contain the two branch points. That is, the branch points appear as "hinges" at which all the planes are joined. With this geometrical construction, if we cross the cut, we end up on a different plane adjacent to the previous one (Figure 11.2).

Riemann surfaces  
and sheets

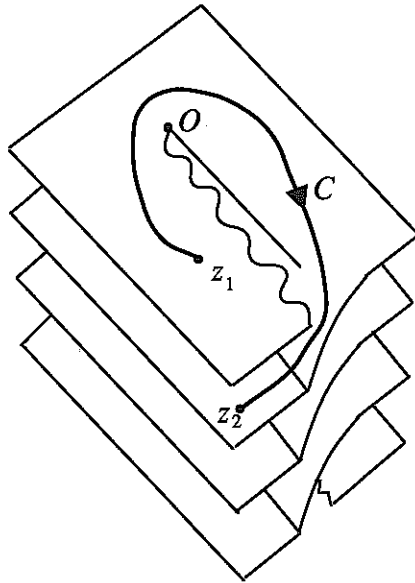
The geometric surface thus constructed is called a **Riemann surface**; each plane is called a **Riemann sheet** and is denoted by  $R_j$ , for  $j = 0, \pm 1, \pm 2, \dots$ . A single-valued function defined on a Riemann sheet is called a **branch** of the original multivalued function.

We have achieved the following: From a multivalued function we have constructed a sequence of single-valued functions, each defined in a single complex plane; from this sequence of functions we have constructed a single complex function defined on a single Riemann surface. Thus, the logarithmic function is analytic throughout the Riemann surface except at the branch points, which are simply the function's singular points.

It is now easy to see the geometrical significance of branch points. A complete cycle around a branch point takes us to another Riemann sheet, where the function takes on a different form. On the other hand, a complete cycle around an ordinary point either never crosses the cut, or if it does, it will cross it back to the original sheet.

Let us now briefly consider two of the more common multivalued functions and their Riemann surfaces.

**11.2.2. Example.** THE FUNCTION  $f(z) = z^{1/n}$   
The only branch points for the function  $f(z) = z^{1/n}$  are  $z = 0$  and the point at infinity.



**Figure 11.2** A few sheets of the Riemann surface of the logarithmic function. The path  $C$  encircling the origin  $O$  ends up on the lower sheet.

Defining  $f_k(z) \equiv r^{1/n} e^{i(\theta+2k\pi/n)}$  for  $k = 0, 1, \dots, n - 1$  and  $0 < \theta < 2\pi$  and following the same procedure as for the logarithmic function, we see that there must be  $n$  Riemann sheets, labeled  $R_0, R_1, \dots, R_{n-1}$ , in the Riemann surface. The lower edge of  $R_{n-1}$  is pasted to the upper edge of  $R_0$  along the cut, which is taken to be along the positive real axis. The Riemann surface for  $n = 2$  is shown in Figure 11.3.

It is clear that for any noninteger value of  $\alpha$  the function  $f(z) = z^\alpha$  has a branch point at  $z = 0$  and another at the point at infinity. For irrational  $\alpha$  the number of Riemann sheets is infinite. ■

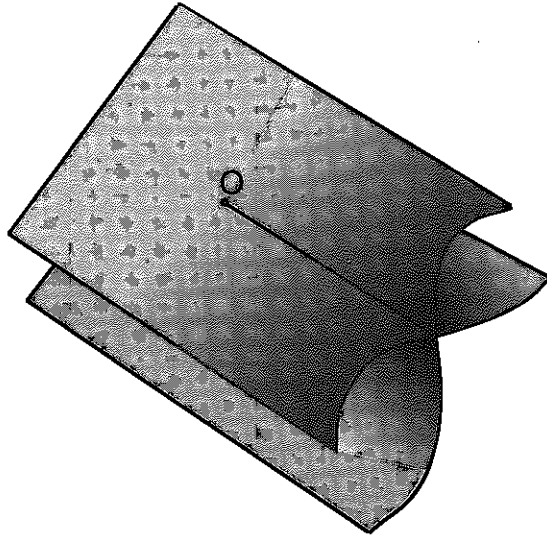
**11.2.3. Example.** THE FUNCTION  $f(z) = (z^2 - 1)^{1/2}$   
 The branch points for the function  $f(z) = (z^2 - 1)^{1/2}$  are at  $z_1 = +1$  and  $z_2 = -1$  (see Figure 11.4). Writing  $z - 1 = r_1 e^{i\theta_1}$  and  $z + 1 = r_2 e^{i\theta_2}$ , we have

$$f(z) = (r_1 e^{i\theta_1})^{1/2} (r_2 e^{i\theta_2})^{1/2} = \sqrt{r_1 r_2} e^{i(\theta_1 + \theta_2)/2}.$$

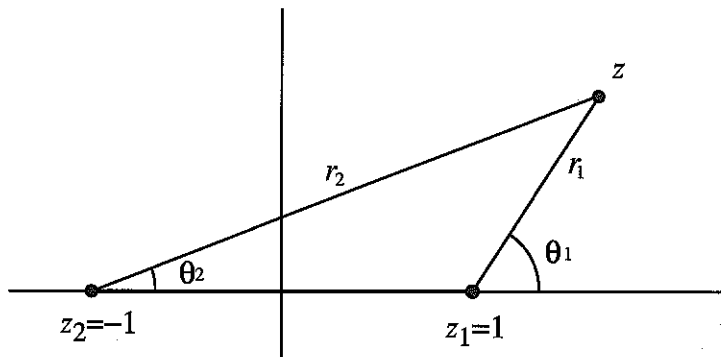
The cut is along the real axis from  $z = -1$  to  $z = +1$ . There are two Riemann sheets in the Riemann surface. Clearly, only cycles of  $2\pi$  involving *one* branch point will cross the cut and therefore end up on a different sheet. Any closed curve that has both  $z_1$  and  $z_2$  as interior points will remain entirely on the original sheet. ■

evaluation of  
 integrals involving  
 cuts

The notion of branch cuts can be used to evaluate certain integrals that do not fit into the three categories discussed in Chapter 10. The basic idea is to circumvent the cut by constructing a contour that is infinitesimally close to the cut and circles around branch points.



**Figure 11.3** The Riemann surface for  $f(z) = z^{1/2}$ .

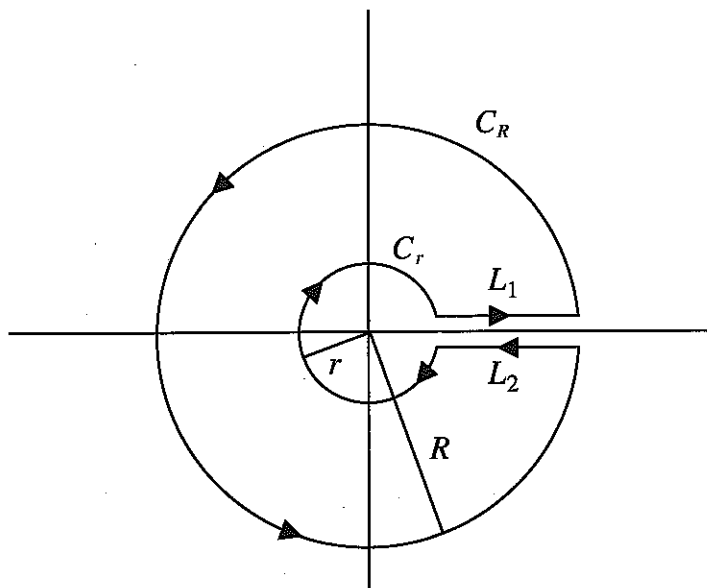


**Figure 11.4** The cut for the function  $f(z) = (z^2 - 1)^{1/2}$  is from  $z_1$  to  $z_2$ . Paths that circle only one of the points cross the cut and end up on the other sheet.

**11.2.4. Example.** To evaluate the integral  $I = \int_0^\infty x^\alpha dx / (x^2 + 1)$  for  $|\alpha| < 1$ , consider the complex integral  $I' = \oint_C z^\alpha dz / (z^2 + 1)$  where  $C$  is as shown in Figure 11.5 and the cut is taken along the positive real axis. To evaluate the contribution from  $C_R$  and  $C_r$ , we let  $\rho$  stand for either  $r$  or  $R$ . Then we have

$$I_\rho = \int_{C_\rho} \frac{(\rho e^{i\theta})^\alpha}{(\rho e^{i\theta})^2 + 1} i\rho e^{i\theta} d\theta = i \int_0^{2\pi} \frac{\rho^{\alpha+1} e^{i(\alpha+1)\theta}}{\rho^2 e^{2i\theta} + 1} d\theta.$$

It is clear that since  $|\alpha| < 1$ ,  $I_\rho \rightarrow 0$  as  $\rho \rightarrow 0$  or  $\rho \rightarrow \infty$ .



**Figure 11.5** The contour for the evaluation of the integrals of Examples 11.2.4 and 11.2.5.

The contributions from  $L_1$  and  $L_2$  do not cancel one another because the value of the function changes above and below the cut. To evaluate these two integrals we have to choose a branch of the function. Let us choose that branch on which  $z^\alpha = |z|^\alpha e^{i\alpha\theta}$  for  $0 < \theta < 2\pi$ . Along  $L_1$ ,  $\theta \approx 0$  or  $z^\alpha = x^\alpha$ , and along  $L_2$ ,  $\theta \approx 2\pi$  or  $z^\alpha = (xe^{2\pi i})^\alpha$ . Thus,

$$\begin{aligned} \oint_C \frac{z^\alpha}{z^2+1} dz &= \int_0^\infty \frac{x^\alpha}{x^2+1} dx + \int_\infty^0 \frac{x^\alpha e^{2\pi i\alpha}}{(xe^{2\pi i})^2+1} dx \\ &= (1 - e^{2\pi i\alpha}) \int_0^\infty \frac{x^\alpha}{x^2+1} dx. \end{aligned} \quad (11.3)$$

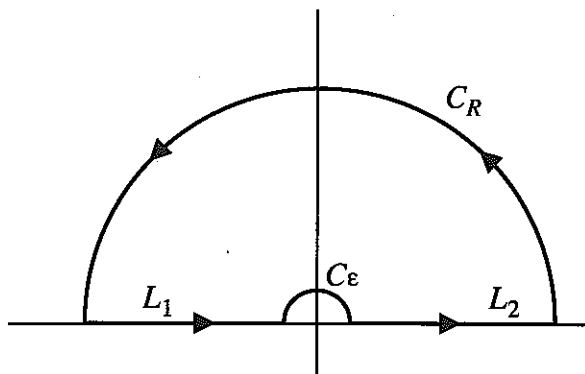
The LHS of this equation can be obtained using the residue theorem. There are two simple poles, at  $z = +i$  and  $z = -i$  with residues  $\text{Res}[f(i)] = (e^{i\pi/2})^\alpha / 2i$  and  $\text{Res}[f(-i)] = -(e^{i3\pi/2})^\alpha / 2i$ . Thus,

$$\oint_C \frac{z^\alpha}{z^2+1} dz = 2\pi i \left( \frac{e^{i\alpha\pi/2}}{2i} - \frac{e^{i3\alpha\pi/2}}{2i} \right) = \pi (e^{i\alpha\pi/2} - e^{i3\alpha\pi/2}).$$

Combining this with Equation (11.3), we obtain

$$\int_0^\infty \frac{x^\alpha}{x^2+1} dx = \frac{\pi(e^{i\alpha\pi/2} - e^{i3\alpha\pi/2})}{1 - e^{2\pi i\alpha}} = \frac{\pi}{2} \sec \frac{\alpha\pi}{2}.$$

If we had chosen a different branch of the function, both the LHS and the RHS of Equation (11.3) would have been different, but the final result would still have been the same. ■



**Figure 11.6** The contour for the evaluation of the integral of Example 11.2.6.

**11.2.5. Example.** Here is another integral involving a branch cut:

$$I = \int_0^{\infty} \frac{x^{-a}}{x+1} dx \quad \text{for } 0 < a < 1.$$

To evaluate this integral we use the zeroth branch of the function and the contour of the previous example (Figure 11.5). Thus, writing  $z = \rho e^{i\theta}$ , we have

$$\begin{aligned} 2\pi i \operatorname{Res}[f(-1)] &= \oint_C \frac{z^{-a}}{z+1} dz = \int_0^{\infty} \frac{\rho^{-a}}{\rho+1} d\rho + \oint_{C_R} \frac{z^{-a}}{z+1} dz \\ &\quad + \int_{\infty}^0 \frac{(\rho e^{2i\pi})^{-a}}{\rho e^{2i\pi} + 1} e^{2i\pi} d\rho + \oint_{C_\epsilon} \frac{z^{-a}}{z+1} dz. \end{aligned} \quad (11.4)$$

The contributions from both circles vanish by the same argument used in the previous example. On the other hand,  $\operatorname{Res}[f(-1)] = (-1)^{-a}$ . For the branch we are using,  $-1 = e^{i\pi}$ . Thus,  $\operatorname{Res}[f(-1)] = e^{-ia\pi}$ . The RHS of Equation (11.4) yields

$$\int_0^{\infty} \frac{\rho^{-a}}{\rho+1} d\rho - e^{-2i\pi a} \int_0^{\infty} \frac{\rho^{-a}}{\rho+1} d\rho = (1 - e^{-2i\pi a})I.$$

It follows from (11.4) that  $(1 - e^{-2i\pi a})I = 2\pi i e^{-i\pi a}$ , or

$$\int_0^{\infty} \frac{x^{-a}}{x+1} dx = \frac{\pi}{\sin a\pi} \quad \text{for } 0 < a < 1. \quad \blacksquare$$

**11.2.6. Example.** Let us evaluate  $I = \int_0^{\infty} \ln x dx / (x^2 + a^2)$  with  $a > 0$ . We choose the zeroth branch of the logarithmic function, in which  $-\pi < \theta < \pi$ , and use the contour of Figure 11.6. For  $L_1$ ,  $z = \rho e^{i\pi}$  (note that  $\rho > 0$ ), and for  $L_2$ ,  $z = \rho$ . Thus, we have

$$\begin{aligned} 2\pi i \operatorname{Res}[f(ia)] &= \oint_C \frac{\ln z}{z^2 + a^2} dz = \int_{\infty}^{\epsilon} \frac{\ln(\rho e^{i\pi})}{(\rho e^{i\pi})^2 + a^2} e^{i\pi} d\rho + \int_{C_\epsilon} \frac{\ln z}{z^2 + a^2} dz \\ &\quad + \int_{\epsilon}^{\infty} \frac{\ln \rho}{\rho^2 + a^2} d\rho + \int_{C_R} \frac{\ln z}{z^2 + a^2} dz, \end{aligned} \quad (11.5)$$

where  $z = ia$  is the only singularity—a simple pole—in the UHP. Now we note that

$$\int_{-\infty}^{\epsilon} \frac{\ln(\rho e^{i\pi})}{(\rho e^{i\pi})^2 + a^2} e^{i\pi} d\rho = \int_{\epsilon}^{\infty} \frac{\ln \rho + i\pi}{\rho^2 + a^2} d\rho = \int_{\epsilon}^{\infty} \frac{\ln \rho}{\rho^2 + a^2} d\rho + i\pi \int_{\epsilon}^{\infty} \frac{d\rho}{\rho^2 + a^2}.$$

The contributions from the circles tend to zero. On the other hand,

$$\text{Res}[f(ia)] = \lim_{z \rightarrow ia} (z - ia) \frac{\ln z}{(z - ia)(z + ia)} = \frac{\ln(ia)}{2ia} = \frac{1}{2ia} \left( \ln a + i \frac{\pi}{2} \right).$$

Substituting the last two results in Equation (11.5), we obtain

$$\frac{\pi}{a} \left( \ln a + i \frac{\pi}{2} \right) = 2 \int_{\epsilon}^{\infty} \frac{\ln \rho}{\rho^2 + a^2} d\rho + i\pi \int_{\epsilon}^{\infty} \frac{d\rho}{\rho^2 + a^2}.$$

It can also easily be shown that  $\int_0^{\infty} d\rho/(\rho^2 + a^2) = \pi/(2a)$ . Thus, in the limit  $\epsilon \rightarrow 0$ , we get  $I = \frac{\pi}{2a} \ln a$ . The sign of  $a$  is irrelevant because it appears as a square in the integral.

Thus, we can write

$$\int_0^{\infty} \frac{\ln x}{x^2 + a^2} dx = \frac{\pi}{2|a|} \ln |a|, \quad a \neq 0. \quad \blacksquare$$

### 11.3 Analytic Continuation

Analytic functions have certain unique properties, some of which we have already noted. For instance, the Cauchy integral formula gives the value of an analytic function inside a simple closed contour once its value on the contour is known. We have also seen that we can deform the contours of integration as long as we do not encounter any singularities of the function.

Combining these two properties and assuming that  $f : \mathbb{C} \rightarrow \mathbb{C}$  is analytic within a region  $S \subset \mathbb{C}$ , we can ask the following question: Is it possible to extend  $f$  beyond  $S$ ? We shall see in this section that the answer is yes in many cases of interest.<sup>4</sup> First consider

equal on a piece,  
equal all over

**11.3.1. Theorem.** *Let  $f_1, f_2 : \mathbb{C} \rightarrow \mathbb{C}$  be analytic in a region  $S$ . If  $f_1 = f_2$  in a neighborhood of a point  $z \in S$ , or for a segment of a curve in  $S$ , then  $f_1 = f_2$  for all  $z \in S$ .*

*Proof.* Let  $g = f_1 - f_2$ , and  $U = \{z \in S \mid g(z) = 0\}$ . Then  $U$  is a subset of  $S$  that includes the neighborhood of  $z$  (or the line segment) in which  $f_1 = f_2$ . If  $U$  is the entire region  $S$ , we are done. Otherwise,  $U$  has a boundary beyond which  $g(z) \neq 0$ . Since all points within the boundary satisfy  $g(z) = 0$ , and since  $g$  is continuous (more than that, it is analytic) on  $S$ ,  $g$  must vanish also on the boundary. But the boundary points are not isolated: Any small circle around any one of them includes points of  $U$  as well as points outside  $U$ . Thus,  $g$  must vanish on a neighborhood of any boundary point, implying that  $g$  vanishes for some points outside  $U$ . This contradicts our assumption. Thus,  $U$  must include the entire region  $S$ .  $\square$

<sup>4</sup>Provided that  $S$  is not discrete (countable). (See [Lang 85, p. 91].)



A consequence of this theorem is the following corollary.

**11.3.2. Corollary.** *The behavior of a function that is analytic in a region  $S \subset \mathbb{C}$  is completely determined by its behavior in a (small) neighborhood of an arbitrary point in that region.*

analytic continuation

This process of determining the behavior of an analytic function outside the region in which it was originally defined is called **analytic continuation**. Although there are infinitely many ways of analytically continuing beyond regions of definition, the values of all functions obtained as a result of diverse continuations are the same at any given point. This follows from Theorem 11.3.1.

Let  $f_1, f_2 : \mathbb{C} \rightarrow \mathbb{C}$  be analytic in regions  $S_1$  and  $S_2$ , respectively. Suppose that  $f_1$  and  $f_2$  have different functional forms in their respective regions of analyticity. If there is an overlap between  $S_1$  and  $S_2$  and if  $f_1 = f_2$  within that overlap, then the (unique) analytic continuation of  $f_1$  into  $S_2$  must be  $f_2$ , and vice versa. In fact, we may regard  $f_1$  and  $f_2$  as a single function  $f : \mathbb{C} \rightarrow \mathbb{C}$  such that

$$f(z) = \begin{cases} f_1(z) & \text{when } z \in S_1, \\ f_2(z) & \text{when } z \in S_2. \end{cases}$$

Clearly,  $f$  is analytic for the combined region  $S = S_1 \cup S_2$ . We then say that  $f_1$  and  $f_2$  are analytic continuations of one another.

**11.3.3. Example.** Let us consider the function  $f_1(z) = \sum_{n=0}^{\infty} z^n$ , which is analytic for  $|z| < 1$ . We have seen that it converges to  $1/(1-z)$  for  $|z| < 1$ . Thus, we have  $f_1(z) = 1/(1-z)$  when  $|z| < 1$ , and  $f_1$  is not defined for  $|z| > 1$ .

Now let us consider a second function,  $f_2(z) = \sum_{n=0}^{\infty} (\frac{3}{5})^{n+1} (z + \frac{2}{3})^n$ , which converges for  $|z + \frac{2}{3}| < \frac{5}{3}$ . To see what it converges to, we note that  $f_2(z) = \frac{3}{5} \sum_{n=0}^{\infty} [\frac{3}{5}(z + \frac{2}{3})]^n$ . Thus,

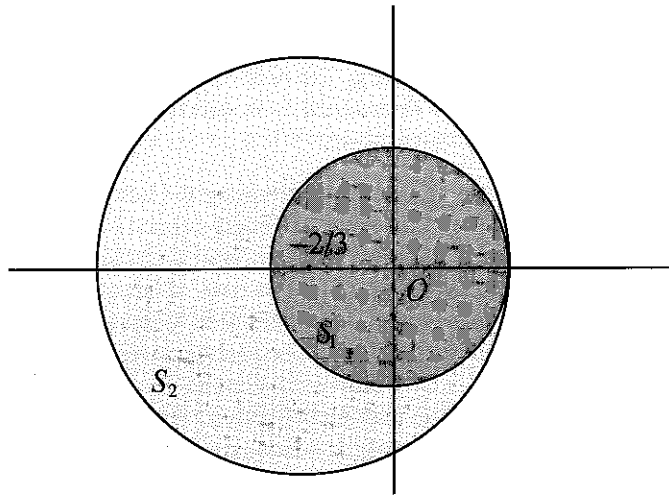
$$f_2(z) = \frac{\frac{3}{5}}{1 - \frac{3}{5}(z + \frac{2}{3})} = \frac{1}{1-z} \quad \text{when } |z + \frac{2}{3}| < \frac{5}{3}.$$

We observe that although  $f_1(z)$  and  $f_2(z)$  have different series representations in the two overlapping regions (see Figure 11.7), they represent the same function,  $f(z) = 1/(1-z)$ . We can therefore write

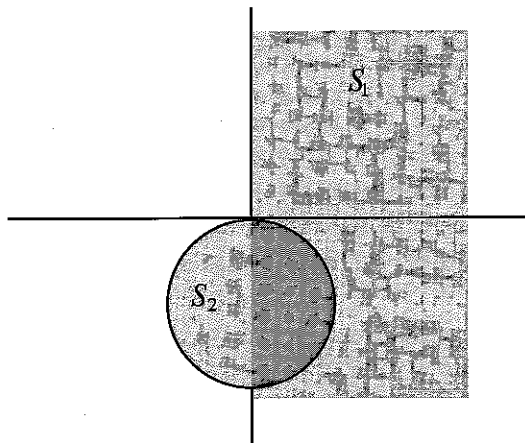
$$f(z) = \begin{cases} f_1(z) & \text{when } |z| < 1, \\ f_2(z) & \text{when } |z + \frac{2}{3}| < \frac{5}{3}, \end{cases}$$

and  $f_1$  and  $f_2$  are analytic continuations of one another. In fact,  $f(z) = 1/(1-z)$  is the analytic continuation of both  $f_1$  and  $f_2$  for all of  $\mathbb{C}$  except  $z = 1$ . Figure 11.7 shows  $S_i$ , the region of definition of  $f_i$ , for  $i = 1, 2$ . ■

**11.3.4. Example.** The function  $f_1(z) = \int_0^{\infty} e^{-zt} dt$  exists only if  $\operatorname{Re}(z) > 0$ , in which case  $f_1(z) = 1/z$ . Its region of definition  $S_1$  is shown in Figure 11.8 and is simply the right half-plane.



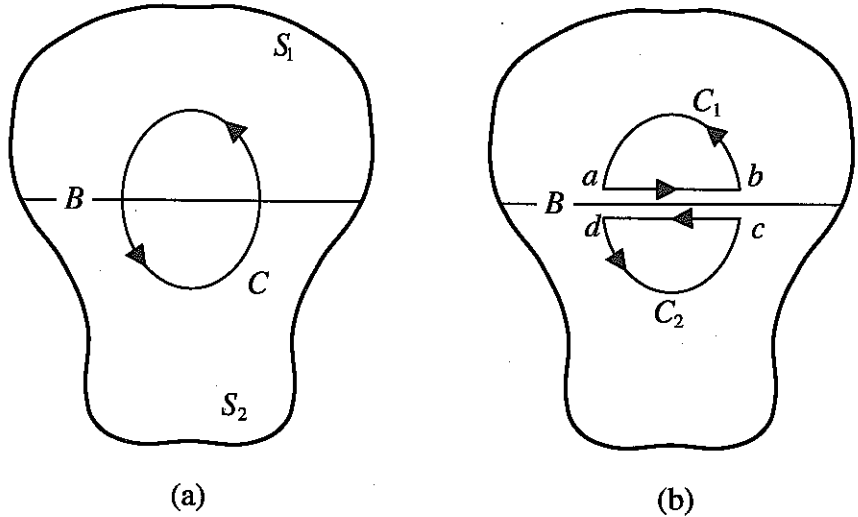
**Figure 11.7** The function defined in the smaller circle is continued analytically into the larger circle.



**Figure 11.8** The functions  $f_1$  and  $f_2$  are analytic continuations of each other:  $f_1$  analytically continues  $f_2$  into the right half-plane, and  $f_2$  analytically continues  $f_1$  into the semicircle in the left half-plane.

Now we define  $f_2$  by a geometric series:  $f_2(z) = i \sum_{n=0}^{\infty} [(z+i)/i]^n$  where  $|z+i| < 1$ . This series converges, within its circle of convergence  $S_2$ , to

$$i \frac{1}{1 - (z+i)/i} = \frac{1}{z}.$$



**Figure 11.9** (a) Regions  $S_1$  and  $S_2$  separated by the boundary  $B$  and the contour  $C$ . (b) The contour  $C$  splits up into  $C_1$  and  $C_2$ .

Thus, we have

$$\frac{1}{z} = \begin{cases} f_1(z) & \text{when } z \in S_1, \\ f_2(z) & \text{when } z \in S_2. \end{cases}$$

The two functions are analytic continuations of one another, and  $f(z) = 1/z$  is the analytic continuation of both  $f_1$  and  $f_2$  for all  $z \in \mathbb{C}$  except  $z = 0$ . ■

### 11.3.1 The Schwarz Reflection Principle

A result that is useful in some physical applications is referred to as a dispersion relation. To derive such a relation we need to know the behavior of analytic functions on either side of the real axis. This is found using the Schwarz reflection principle, for which we need the following result.

**11.3.5. Proposition.** *Let  $f_i$  be analytic throughout  $S_i$ , where  $i = 1, 2$ . Let  $B$  be the boundary between  $S_1$  and  $S_2$  (Figure 11.9) and assume that  $f_1$  and  $f_2$  are continuous on  $B$  and coincide there. Then the two functions are analytic continuations of one another and together they define a (unique) function*

$$f(z) = \begin{cases} f_1(z) & \text{when } z \in S_1 \cup B, \\ f_2(z) & \text{when } z \in S_2 \cup B, \end{cases}$$

which is analytic throughout the entire region  $S_1 \cup S_2 \cup B$ .

*Proof.* The proof consists in showing that the function integrates to zero along any closed curve in  $S_1 \cup S_2 \cup B$ . Once this is done, one can use Morera's theorem to conclude analyticity. The case when the closed curve is entirely in either  $S_1$  or  $S_2$  is trivial. When the curve is partially in  $S_1$  and partially in  $S_2$  the proof becomes only slightly more complicated, because one has to split up the contour  $C$  into  $C_1$  and  $C_2$  of Figure 11.9(b). The details are left as an exercise.  $\square$

Schwarz reflection principle

**11.3.6. Theorem.** (Schwarz reflection principle) *Let  $f$  be a function that is analytic in a region  $S$  that has a segment of the real axis as part of its boundary  $B$ . If  $f(z)$  is real whenever  $z$  is real, then the analytic continuation  $g$  of  $f$  into  $S^*$  (the mirror image of  $S$  with respect to the real axis) exists and is given by*

$$g(z) = f^*(z^*) \quad \text{where } z \in S^*.$$

*Proof.* First, we show that  $g$  is analytic in  $S^*$ . Let

$$f(z) \equiv u(x, y) + iv(x, y), \quad g(z) \equiv U(x, y) + iV(x, y).$$

Then  $f(z^*) = f(x, -y) = u(x, -y) + iv(x, -y)$  and  $g(z) = f^*(z^*)$  imply that  $U(x, y) = u(x, -y)$  and  $V(x, y) = -v(x, -y)$ . Therefore,

$$\begin{aligned} \frac{\partial U}{\partial x} &= \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = -\frac{\partial v}{\partial(-y)} = \frac{\partial V}{\partial y}, \\ \frac{\partial U}{\partial y} &= -\frac{\partial u}{\partial y} = \frac{\partial v}{\partial x} = -\frac{\partial V}{\partial x}. \end{aligned}$$

These are the Cauchy-Riemann conditions for  $g(z)$ . Thus,  $g$  is analytic.

Next, we note that  $f(x, 0) = g(x, 0)$ , implying that  $f$  and  $g$  agree on the real axis. Proposition 11.3.5 then implies that  $f$  and  $g$  are analytic continuations of one another.  $\square$

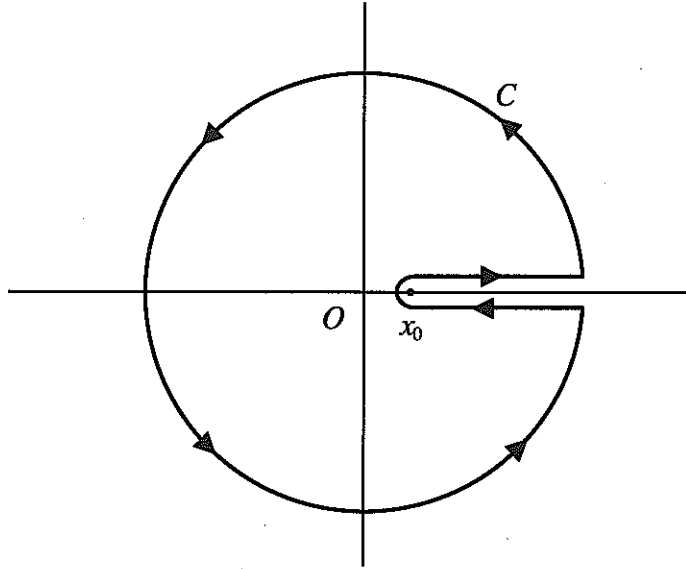
It follows from this theorem that there exists an analytic function  $h$  such that

$$h(z) = \begin{cases} f(z) & \text{when } z \in S, \\ g(z) & \text{when } z \in S^*. \end{cases}$$

We note that  $h(z^*) = g(z^*) = f^*(z) = h^*(z)$ .

### 11.3.2 Dispersion Relations

Let  $f$  be analytic throughout the complex plane except at a cut along the real axis extending from  $x_0$  to infinity. For a point  $z$  not on the  $x$ -axis, the Cauchy integral formula gives  $f(z) = (2\pi i)^{-1} \int_C f(\xi) d\xi / (\xi - z)$  where  $C$  is the contour shown in Figure 11.10. We assume that  $f$  drops to zero fast enough that the contribution



**Figure 11.10** The contour used for dispersion relations.

from the large circle tends to zero. The reader may show that the contribution from the small half-circle around  $x_0$  also vanishes. Then

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \left[ \int_{x_0+i\epsilon}^{\infty+i\epsilon} \frac{f(\xi)}{\xi-z} d\xi - \int_{x_0-i\epsilon}^{\infty-i\epsilon} \frac{f(\xi)}{\xi-z} d\xi \right] \\ &= \frac{1}{2\pi i} \left[ \int_{x_0}^{\infty} \frac{f(x+i\epsilon)}{x-z+i\epsilon} dx - \int_{x_0}^{\infty} \frac{f(x-i\epsilon)}{x-z-i\epsilon} dx \right]. \end{aligned}$$

Since  $z$  is not on the real axis, we can ignore the  $i\epsilon$  terms in the denominators, so that  $f(z) = (2\pi i)^{-1} \int_{x_0}^{\infty} [f(x+i\epsilon) - f(x-i\epsilon)] dx / (x-z)$ . The Schwarz reflection principle in the form  $f^*(z) = f(z^*)$  can now be used to yield

$$f(x+i\epsilon) - f(x-i\epsilon) = f(x+i\epsilon) - f^*(x+i\epsilon) = 2i \operatorname{Im}[f(x+i\epsilon)].$$

dispersion relation      The final result is

$$f(z) = \frac{1}{\pi} \int_{x_0}^{\infty} \frac{\operatorname{Im}[f(x+i\epsilon)]}{x-z} dx.$$

This is one form of a **dispersion relation**. It expresses the value of a function at any point of the cut complex plane in terms of an integral of the imaginary part of the function on the upper edge of the cut.

When there are no residues in the UHP, we can obtain other forms of dispersion relations by equating the real and imaginary parts of Equation (10.11). The result

is

$$\begin{aligned} \operatorname{Re}[f(x_0)] &= \pm \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{\operatorname{Im}[f(x)]}{x - x_0} dx, \\ \operatorname{Im}[f(x_0)] &= \mp \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{\operatorname{Re}[f(x)]}{x - x_0} dx, \end{aligned} \tag{11.6}$$

Hilbert transform where the upper (lower) sign corresponds to placing the small semicircle around  $x_0$  in the UHP (LHP). The real and imaginary parts of  $f$ , as related by Equation (11.6), are sometimes sometimes said to be the **Hilbert transform** of one another.

In some applications the imaginary part of  $f$  is an odd function of its argument. Then the first equation in (11.6) can be written as

$$\operatorname{Re}[f(x_0)] = \pm \frac{2}{\pi} P \int_0^{\infty} \frac{x \operatorname{Im}[f(x)]}{x^2 - x_0^2} dx.$$

To arrive at dispersion relations, the following condition must hold:

$$\lim_{R \rightarrow \infty} R |f(Re^{i\theta})| = 0,$$

dispersion relation with one subtraction where  $R$  is the radius of the large semicircle in the UHP (or LHP). If  $f$  does not satisfy this prerequisite, it is still possible to obtain a dispersion relation called a **dispersion relation with one subtraction**. This can be done by introducing an extra factor of  $x$  in the denominator of the integrand. We start with Equation (10.15), confining ourselves to the UHP and assuming that there are no poles there, so that the sum over residues is dropped:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{1}{i\pi} P \int_{-\infty}^{\infty} \frac{f(x)}{(x - x_1)(x - x_2)} dx.$$

The reader may check that by equating the real and imaginary parts on both sides, letting  $x_1 = 0$  and  $x_2 = x_0$ , and changing  $x$  to  $-x$  in the first half of the interval of integration, we obtain

$$\frac{\operatorname{Re}[f(x_0)]}{x_0} = \frac{\operatorname{Re}[f(0)]}{x_0} + \frac{1}{\pi} \left[ P \int_0^{\infty} \frac{\operatorname{Im}[f(-x)]}{x(x + x_0)} dx + P \int_0^{\infty} \frac{\operatorname{Im}[f(x)]}{x(x - x_0)} dx \right].$$

For the case where  $\operatorname{Im}[f(-x)] = -\operatorname{Im}[f(x)]$ , this equation yields

$$\operatorname{Re}[f(x_0)] = \operatorname{Re}[f(0)] + \frac{2x_0^2}{\pi} P \int_0^{\infty} \frac{\operatorname{Im}[f(x)]}{x(x^2 - x_0^2)} dx. \tag{11.7}$$

optical theorem **11.3.7. Example.** In optics, it has been shown that the imaginary part of the forward-scattering light amplitude with frequency  $\omega$  is related, by the so-called **optical theorem**, to the total cross section for the absorption of light of that frequency:

$$\operatorname{Im}[f(\omega)] = \frac{\omega}{4\pi} \sigma_{\text{tot}}(\omega).$$

Kramers-Kronig relation Substituting this in Equation (11.7) yields

$$\operatorname{Re}[f(\omega_0)] = \operatorname{Re}[f(0)] + \frac{\omega_0^2}{2\pi^2} P \int_0^\infty \frac{\sigma_{\text{tot}}(\omega)}{\omega^2 - \omega_0^2} d\omega. \quad (11.8)$$

Thus, the real part of the (coherent) forward scattering of light, that is, the real part of the *index of refraction*, can be computed from Equation (11.8) by either measuring or calculating  $\sigma_{\text{tot}}(\omega)$ , the simpler quantity describing the absorption of light in the medium. Equation (11.8) is the original **Kramers-Kronig relation**. ■

## 11.4 The Gamma and Beta Functions

We have already encountered the gamma function. In this section, we derive some useful relations involving the gamma function and the closely related beta function. The gamma function is a generalization of the factorial function—which is defined only for positive integers—to the system of complex numbers. By differentiating the integral  $I(\alpha) \equiv \int_0^\infty e^{-\alpha t} dt = 1/\alpha$  with respect to  $\alpha$  repeatedly and setting  $\alpha = 1$  at the end, we get  $\int_0^\infty t^n e^{-t} dt = n!$ . This fact motivates the generalization

gamma function defined

$$\Gamma(z) \equiv \int_0^\infty t^{z-1} e^{-t} dt \quad \text{for } \operatorname{Re}(z) > 0, \quad (11.9)$$

where  $\Gamma$  is called the **gamma** (or **factorial**) **function**. It is also called *Euler's integral of the second kind*. It is clear from its definition that

$$\Gamma(n+1) = n! \quad (11.10)$$

if  $n$  is a positive integer. The restriction  $\operatorname{Re}(z) > 0$  assures the convergence of the integral.

An immediate consequence of Equation (11.9) is obtained by integrating it by parts:

$$\Gamma(z+1) = z\Gamma(z). \quad (11.11)$$

This also leads to Equation (11.10) by iteration.

Another consequence is the analyticity of  $\Gamma(z)$ . Differentiating Equation (11.11) with respect to  $z$ , we obtain

$$\frac{d\Gamma(z+1)}{dz} = \Gamma(z) + z \frac{d\Gamma(z)}{dz}.$$

Thus,  $d\Gamma(z)/dz$  exists and is finite if and only if  $d\Gamma(z+1)/dz$  is finite (recall that  $z \neq 0$ ). The procedure of showing the latter is outlined in Problem 11.16. Therefore,  $\Gamma(z)$  is analytic whenever  $\Gamma(z+1)$  is. To see the singularities of  $\Gamma(z)$ , we note that

$$\Gamma(z+n) = z(z+1)(z+2)\cdots(z+n-1)\Gamma(z),$$

or

$$\Gamma(z) = \frac{\Gamma(z+n)}{z(z+1)(z+2)\cdots(z+n-1)}. \tag{11.12}$$

The numerator is analytic as long as  $\operatorname{Re}(z+n) > 0$ , or  $\operatorname{Re}(z) > -n$ . Thus, for  $\operatorname{Re}(z) > -n$ , the singularities of  $\Gamma(z)$  are the poles at  $z = 0, -1, -2, \dots, -n+1$ . Since  $n$  is arbitrary, we conclude that

**11.4.1. Box.**  $\Gamma(z)$  is analytic at all  $z \in \mathbb{C}$  except at  $z = 0, -1, -2, \dots$ , where  $\Gamma(z)$  has simple poles.

A useful result is obtained by setting  $z = \frac{1}{2}$  in Equation (11.9):

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \tag{11.13}$$

This can be obtained by making the substitution  $u = \sqrt{t}$  in the integral.

We can derive an expression for the logarithmic derivative of the gamma function that involves an infinite series. To do so, we use Equation (11.2) noting that  $1/\Gamma(z+1)$  is an entire function with simple zeros at  $\{-k\}_{k=1}^{\infty}$ . Equation (11.2) gives

$$\frac{1}{\Gamma(z+1)} = e^{\gamma z} \prod_{k=1}^{\infty} \left(1 + \frac{z}{k}\right) e^{-z/k},$$

where  $\gamma$  is a constant to be determined. Using Equation (11.11), we obtain

$$\frac{1}{\Gamma(z)} = ze^{\gamma z} \prod_{k=1}^{\infty} \left(1 + \frac{z}{k}\right) e^{-z/k}. \tag{11.14}$$

Euler–Mascheroni constant

To determine  $\gamma$ , let  $z = 1$  in Equation (11.14) and evaluate the resulting product numerically. The result is  $\gamma = 0.57721566\dots$ , the so-called **Euler–Mascheroni constant**.

Differentiating the logarithm of both sides of Equation (11.14), we obtain

$$\frac{d}{dz} \ln[\Gamma(z)] = -\frac{1}{z} - \gamma + \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{z+k}\right). \tag{11.15}$$

beta function defined

Other properties of the gamma function are derivable from the results presented here. Those derivations are left as problems. The **beta function**, or **Euler’s integral of the first kind**, is defined for complex numbers  $a$  and  $b$  as follows:

$$B(a, b) \equiv \int_0^1 t^{a-1}(1-t)^{b-1} dt \quad \text{where } \operatorname{Re}(a), \operatorname{Re}(b) > 0. \tag{11.16}$$



By changing  $t$  to  $1/t$ , we can also write

$$B(a, b) \equiv \int_1^\infty t^{-a-b}(t-1)^{b-1} dt. \tag{11.17}$$

Since  $0 \leq t \leq 1$  in Equation (11.16), we can define  $\theta$  by  $t = \sin^2 \theta$ . This gives

$$B(a, b) = 2 \int_0^{\pi/2} \sin^{2a-1} \theta \cos^{2b-1} \theta d\theta. \tag{11.18}$$

This relation can be used to establish a connection between the gamma and beta functions. We note that

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt = 2 \int_0^\infty x^{2a-1} e^{-x^2} dx,$$

where in the last step we changed the variable to  $x = \sqrt{t}$ . Multiply  $\Gamma(a)$  by  $\Gamma(b)$  and express the resulting double integral in terms of polar coordinates to obtain  $\Gamma(a)\Gamma(b) = \Gamma(a+b)B(a, b)$ , or

gamma function and  
beta function are  
related

$$B(a, b) = B(b, a) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \tag{11.19}$$

Let us now establish the following useful relation:

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}. \tag{11.20}$$

With  $a = z$  and  $b = 1 - z$ , and using  $u = \tan \theta$ , Equations (11.18) and (11.19) give

$$\Gamma(z)\Gamma(1-z) = B(z, 1-z) = 2 \int_0^\infty \frac{u^{2z-1}}{u^2+1} du \quad \text{for } 0 < \text{Re}(z) < 1.$$

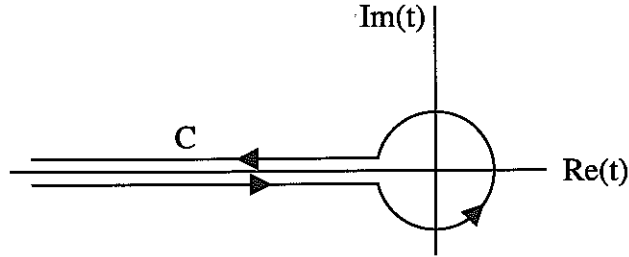
Using the result obtained in Example 11.2.4, we immediately get Equation (11.20), valid for  $0 < \text{Re}(z) < 1$ . By analytic continuation we then generalize Equation (11.20) to values of  $z$  for which both sides are analytic.

**11.4.2. Example.** As an illustration of the use of Equation (11.20), let us show that  $\Gamma(z)$  can also be written as

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i} \int_C \frac{e^t}{t^z} dt, \tag{11.21}$$

where  $C$  is the contour shown in Figure 11.11. From Equations (11.9) and (11.20) it follows that

$$\frac{1}{\Gamma(z)} = \frac{\sin \pi z}{\pi} \Gamma(1-z) = \frac{\sin \pi z}{\pi} \int_0^\infty e^{-r} r^{-z} dr = \frac{e^{i\pi z} - e^{-i\pi z}}{2\pi i} \int_0^\infty \frac{e^{-r}}{r^z} dr.$$



**Figure 11.11** The contour  $C$  used in evaluating the reciprocal gamma function.

The contour integral of Equation (11.21) can be evaluated by noting that above the real axis,  $t = re^{i\pi} = -r$ , below it  $t = re^{-i\pi} = -r$ , and, as the reader may check, that the contribution from the small circle at the origin is zero; so

$$\begin{aligned} \int_C \frac{e^t}{t^z} dt &= \int_0^\infty \frac{e^{-r}}{(re^{i\pi})^z} (-dr) + \int_\infty^0 \frac{e^{-r}}{(re^{-i\pi})^z} (-dr) \\ &= -e^{-i\pi z} \int_0^\infty \frac{e^{-r}}{r^z} dr + e^{i\pi z} \int_0^\infty \frac{e^{-r}}{r^z} dr. \end{aligned}$$

Comparison with the last equation above yields the desired result. ■

Another useful relation can be obtained by combining Equations (11.11) and (11.20):  $\Gamma(z)\Gamma(1-z) = \Gamma(z)(-z)\Gamma(-z) = \pi/\sin \pi z$ . Thus,

$$\Gamma(z)\Gamma(-z) = -\frac{\pi}{z \sin \pi z}. \tag{11.22}$$

Once we know  $\Gamma(x)$  for positive values of real  $x$ , we can use Equation (11.22) to find  $\Gamma(x)$  for  $x < 0$ . Thus, for instance,  $\Gamma(\frac{1}{2}) = \sqrt{\pi}$  gives  $\Gamma(-\frac{1}{2}) = -2\sqrt{\pi}$ . Equation (11.22) also shows that the gamma function has simple poles wherever  $z$  is a negative integer.

## 11.5 Method of Steepest Descent

It is shown in statistical mechanics ([Hill 87, pp. 150–152]) that the partition function, which generates all the thermodynamical quantities, can be written as a contour integral. Debye found a very elegant technique of approximating this contour integral, which we investigate in this section. Consider the integral

$$I(\alpha) \equiv \int_C e^{\alpha f(z)} g(z) dz \tag{11.23}$$

where  $|\alpha|$  is large and  $f$  and  $g$  are analytic in some region of  $\mathbb{C}$  containing the contour  $C$ . Since this integral occurs frequently in physical applications, it would

be helpful if we could find a general approximation for it that is applicable for all  $f$  and  $g$ . The fact that  $|\alpha|$  is large will be of great help. By redefining  $f(z)$ , if necessary, we can assume that  $\alpha = |\alpha|e^{i \arg(\alpha)}$  is real and positive [absorb  $e^{i \arg(\alpha)}$  into the function  $f(z)$  if need be].

The exponent of the integrand can be written as

$$\alpha f(z) = \alpha u(x, y) + i\alpha v(x, y).$$

Since  $\alpha$  is large and positive, we expect the exponential to be the largest at the maximum of  $u(x, y)$ . Thus, if we deform the contour so that it passes through a point  $z_0$  at which  $u(x, y)$  is maximum, the contribution to the integral may come mostly from the neighborhood of  $z_0$ . This opens up the possibility of expanding the exponent about  $z_0$  and keeping the lowest terms in the expansion, which is what we are after. There is one catch, however. Because of the largeness of  $\alpha$ , the imaginary part of  $\alpha f$  in the exponent will oscillate violently as  $v(x, y)$  changes even by a small amount. This oscillation can make the contribution of the real part of  $f(z_0)$  negligibly small and render the whole procedure useless. Thus, we want to tame the variation of  $\exp[i\alpha v(x, y)]$  by making  $v(x, y)$  vary as slowly as possible. A necessary condition is for the derivative of  $v$  to vanish at  $z_0$ . This and the fact that the real part is to have a maximum at  $z_0$  lead to

$$\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \left. \frac{df}{dz} \right|_{z_0} = 0. \quad (11.24)$$

However, we do not stop here but demand that the imaginary part of  $f$  be constant along the deformed contour:  $\text{Im}[f(z)] = \text{Im}[f(z_0)]$  or  $v(x, y) = v(x_0, y_0)$ .

Equation (11.24) and the Cauchy-Riemann conditions imply that  $\partial u/\partial x = 0 = \partial u/\partial y$  at  $z_0$ . Thus, it might appear that  $z_0$  is a maximum (or minimum) of the surface described by the function  $u(x, y)$ . This is not true: For the surface to have a maximum (minimum), both second derivatives,  $\partial^2 u/\partial x^2$  and  $\partial^2 u/\partial y^2$ , must be negative (positive). But that is impossible because  $u(x, y)$  is harmonic—the sum of these two derivatives is zero. Recall that a point at which the derivatives vanish but that is neither a maximum nor a minimum is called a *saddle point*. That is why the procedure described below is sometimes called the **saddle point approximation**.

We are interested in values of  $z$  close to  $z_0$ . So let us expand  $f(z)$  in a Taylor series about  $z_0$ , use Equation (11.24), and keep terms only up to the second, to obtain

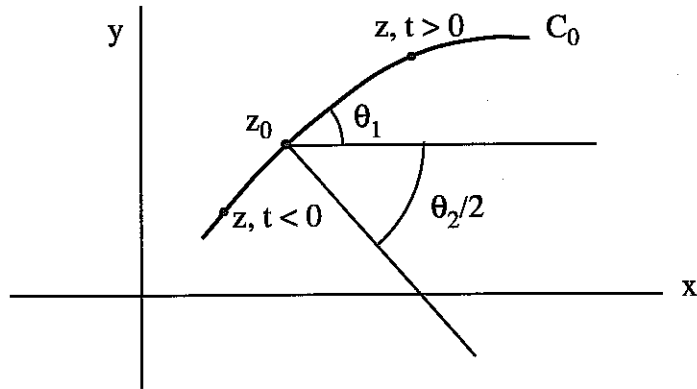
$$f(z) = f(z_0) + \frac{1}{2}(z - z_0)^2 f''(z_0). \quad (11.25)$$

Let us assume that  $f''(z_0) \neq 0$ , and define

$$z - z_0 = r_1 e^{i\theta_1} \quad \text{and} \quad \frac{1}{2} f''(z_0) = r_2 e^{i\theta_2} \quad (11.26)$$

and substitute in the above expansion to obtain

$$f(z) - f(z_0) = r_1^2 r_2 e^{i(2\theta_1 + \theta_2)}, \quad (11.27)$$



**Figure 11.12** A segment of the contour  $C_0$  in the vicinity of  $z_0$ . The lines mentioned in the text are small segments of the contour  $C_0$  centered at  $z_0$ .

or

$$\begin{aligned} \operatorname{Re}[f(z) - f(z_0)] &= r_1^2 r_2 \cos(2\theta_1 + \theta_2), \\ \operatorname{Im}[f(z) - f(z_0)] &= r_1^2 r_2 \sin(2\theta_1 + \theta_2). \end{aligned} \tag{11.28}$$

The constancy of  $\operatorname{Im}[f(z)]$  implies that  $\sin(2\theta_1 + \theta_2) = 0$ , or  $2\theta_1 + \theta_2 = n\pi$ . Thus, for  $\theta_1 = -\theta_2/2 + n\pi/2$  where  $n = 0, 1, 2, 3$ , the imaginary part of  $f$  is constant. The angle  $\theta_2$  is determined by the second equation in (11.26). Once we determine  $n$ , the path of saddle point integration will be specified.

To get insight into this specification, consider  $z - z_0 = r_1 e^{i(-\theta_2/2 + n\pi/2)}$ , and eliminate  $r_1$  from its real and imaginary parts to obtain

$$y - y_0 = \left[ \tan\left(\frac{n\pi}{2} - \frac{\theta_2}{2}\right) \right] (x - x_0).$$

This is the equation of a line passing through  $z_0 = (x_0, y_0)$  and making an angle of  $\theta_1 = (n\pi - \theta_2)/2$  with the real axis. For  $n = 0, 2$  we get one line, and for  $n = 1, 3$  we get another that is perpendicular to the first (see Figure 11.12). It is to be emphasized that along both these lines the imaginary part of  $f(z)$  remains constant. To choose the correct line, we need to look at the real part of the function. Also note that these “lines” are small segments of (or tangents to) the deformed contour at  $z_0$ .

We are looking for directions along which  $\operatorname{Re}(f)$  goes through a relative maximum at  $z_0$ . In fact, we are after a path on which the function decreases maximally. This occurs when  $\operatorname{Re}[f(z)] - \operatorname{Re}[f(z_0)]$  take the largest negative value. Equation (11.28) determines such a path: It is that path on which  $\cos(2\theta_1 + \theta_2) = -1$ , or when  $n = 1, 3$ . There is only one such path in the region of interest, and the procedure is uniquely determined.<sup>5</sup> Because the descent from the maximum value at

method of steepest descent

$z_0$  is maximum along such a path, this procedure is called the method of **steepest descent**.

Now that we have determined the contour, let us approximate the integral. Substituting  $2\theta_1 + \theta_2 = \pi, 3\pi$  in Equation (11.27), we get

$$f(z) - f(z_0) = -r_1^2 r_2 \equiv -t^2 = \frac{1}{2}(z - z_0)^2 f''(z_0). \quad (11.29)$$

Using this in Equation (11.23) yields

$$I(\alpha) \approx \int_{C_0} e^{\alpha[f(z_0) - t^2]} g(z) dz = e^{\alpha f(z_0)} \int_{C_0} e^{-\alpha t^2} g(z) dz, \quad (11.30)$$

where  $C_0$  is the deformed contour passing through  $z_0$ .

To proceed, we need to solve for  $z$  in terms of  $t$ . From Equation (11.29) we have

$$(z - z_0)^2 = -\frac{2}{f''(z_0)} t^2 = -\frac{t^2}{r_2} e^{-i\theta_2}.$$

Therefore,  $|z - z_0| = |t|/\sqrt{r_2}$ , or  $z - z_0 = (|t|/\sqrt{r_2})e^{i\theta_1}$ , by the first equation of (11.26). Let us agree that for  $t > 0$ , the point  $z$  on the contour will move in the direction that makes an angle of  $0 \leq \theta_1 < \pi$ , and that  $t < 0$  corresponds to the opposite direction. This convention removes the remaining ambiguity of the angle  $\theta_1$ , and gives

$$z = z_0 + \frac{t}{\sqrt{r_2}} e^{i\theta_1}, \quad 0 \leq \theta_1 < \pi. \quad (11.31)$$

Using the Taylor expansion of  $g(z)$  about  $z_0$ , we can write

$$\begin{aligned} g(z) dz &= \left\{ \sum_{n=0}^{\infty} \frac{t^n}{r_2^{n/2} n!} e^{in\theta_1} g^{(n)}(z_0) \right\} \frac{e^{i\theta_1}}{\sqrt{r_2}} dt \\ &= \sum_{n=0}^{\infty} \frac{t^n}{r_2^{(n+1)/2} n!} e^{i(n+1)\theta_1} g^{(n)}(z_0) dt, \end{aligned}$$

and substituting this in Equation (11.30) yields

$$\begin{aligned} I(\alpha) &\approx e^{\alpha f(z_0)} \int_{C_0} e^{-\alpha t^2} \left\{ \sum_{n=0}^{\infty} \frac{t^n}{r_2^{(n+1)/2} n!} e^{i(n+1)\theta_1} g^{(n)}(z_0) \right\} dt \\ &= e^{\alpha f(z_0)} \sum_{n=0}^{\infty} \frac{e^{i(n+1)\theta_1}}{r_2^{(n+1)/2} n!} g^{(n)}(z_0) \int_{-\infty}^{\infty} e^{-\alpha t^2} t^n dt. \end{aligned} \quad (11.32)$$

<sup>5</sup>The angle  $\theta_1$  is still ambiguous by  $\pi$ , because  $n$  can be 1 or 3. However, by a suitable sign convention described below, we can remove this ambiguity.

asymptotic expansion of  $\Gamma(\alpha)$

The extension of the integral limits to infinity does not alter the result significantly because  $\alpha$  is assumed large and positive. The integral in the sum is zero for odd  $n$ . When  $n$  is even, we make the substitution  $u = \alpha t^2$  and show that  $\int_{-\infty}^{\infty} e^{-\alpha t^2} t^n dt = \alpha^{-(n+1)/2} \Gamma[(n+1)/2]$ . With  $n = 2k$ , and using  $r_2 = |f''(z_0)|/2$ , the sum becomes

$$I(\alpha) \approx e^{\alpha f(z_0)} \sum_{k=0}^{\infty} \frac{2^{k+1/2} e^{i(2k+1)\theta_1}}{|f''(z_0)|^{k+1/2} (2k)!} g^{(2k)}(z_0) \Gamma(k + \frac{1}{2}) \alpha^{-k-1/2}. \tag{11.33}$$

This is called the **asymptotic expansion** of  $I(\alpha)$ . In most applications, only the first term of the above series is retained, giving

$$I(\alpha) \approx e^{\alpha f(z_0)} \sqrt{\frac{2\pi}{\alpha}} \frac{e^{i\theta_1} g(z_0)}{\sqrt{|f''(z_0)|}}. \tag{11.34}$$

**11.5.1. Example.** Let us approximate the integral

$$I(\alpha) \equiv \Gamma(\alpha + 1) = \int_0^{\infty} e^{-z} z^{\alpha} dz,$$

where  $\alpha$  is a positive real number. First, we must rewrite the integral in the form of Equation (11.23). We can do this by noting that  $z^{\alpha} = e^{\alpha \ln z}$ . Thus, we have

$$I(\alpha) = \int_0^{\infty} e^{\alpha \ln z - z} dz = \int_0^{\infty} e^{\alpha(\ln z - z/\alpha)} dz,$$

and we identify  $f(z) = \ln z - z/\alpha$  and  $g(z) = 1$ . The saddle point is found from  $f'(z) = 0$  or  $z_0 = \alpha$ . Furthermore, from

$$\frac{1}{2} f''(z_0) = \frac{1}{2} \left( -\frac{1}{\alpha^2} \right) = \frac{1}{2\alpha^2} e^{i\pi} \Rightarrow \theta_2 = \pi$$

and  $2\theta_1 + \theta_2 = \pi, 3\pi$ , as well as the condition  $0 \leq \theta_1 < \pi$ , we conclude that  $\theta_1 = 0$ .

Stirling approximation

Substitution in Equation (11.34) yields

$$\Gamma(\alpha + 1) \approx e^{\alpha f(z_0)} \sqrt{\frac{2\pi}{\alpha}} \frac{1}{\sqrt{1/\alpha^2}} = \sqrt{2\pi\alpha} e^{\alpha(\ln\alpha - 1)} = \sqrt{2\pi} e^{-\alpha} \alpha^{\alpha+1/2}, \tag{11.35}$$

called the **Stirling approximation**. ■

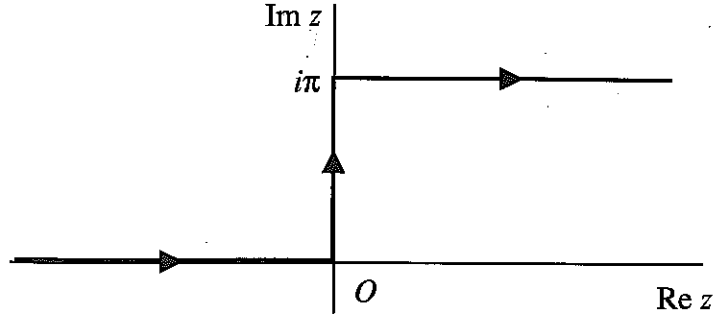
**11.5.2. Example.** The *Hankel function of the first kind* is defined as

$$H_{\nu}^{(1)}(\alpha) \equiv \frac{1}{i\pi} \int_C e^{(\alpha/2)(z-1/z)} \frac{dz}{z^{\nu+1}},$$

where  $C$  is the contour shown in Figure 11.13. We want to find the asymptotic expansion of this function, choosing the branch of the function in which  $-\pi < \theta < \pi$ .

We identify  $f(z) = \frac{1}{2}(z - 1/z)$  and  $g(z) = z^{-\nu-1}$ . Next, the stationary points of  $f$  are calculated:

$$\frac{df}{dz} = \frac{1}{2} + \frac{1}{2z^2} = 0 \Rightarrow z_0 = \pm i.$$



**Figure 11.13** The contour for the evaluation of the Hankel function of the first kind.

The contour of integration suggests the saddle point  $z_0 = +i$ . The second derivative evaluated at the saddle point gives  $f''(z_0) = -1/z_0^3 = -i = e^{-i\pi/2}$ , or  $\theta_2 = -\pi/2$ . This, and the convention  $0 \leq \theta_1 < \pi$ , force us to choose  $\theta_1 = 3\pi/4$ . Substituting this in Equation (11.34) and noting that  $f(i) = i$  and  $|f''(z_0)| = 1$ , we obtain

$$H_v^{(1)}(\alpha) \equiv \frac{1}{i\pi} I(\alpha) \approx \frac{1}{i\pi} e^{\alpha i} \sqrt{\frac{2\pi}{\alpha}} e^{i3\pi/4} i^{-v-1} = \sqrt{\frac{2}{\alpha\pi}} e^{i(\alpha - v\pi/2 - \pi/4)},$$

where we have used  $i^{-v-1} = e^{-i(v+1)\pi/2}$ . ■

Although Equation (11.34) is adequate for most applications, we shall have occasion to demand a better approximation. One may try to keep higher-order terms of Equation (11.33), but that infinite sum is in reality inconsistent. The reason is that in the product  $g(z) dz$ , we kept only the first power of  $t$  in the expansion of  $z$ . To restore consistency, let us expand  $z(t)$  as well. Suppose

$$z - z_0 = \sum_{m=1}^{\infty} b_m t^m \Rightarrow dz = \sum_{m=0}^{\infty} (m+1)b_{m+1} t^m dt,$$

so that

$$\begin{aligned} g(z) dz &= \sum_{n=0}^{\infty} \frac{t^n}{r_2^{n/2} n!} e^{in\theta_1} g^{(n)}(z_0) \sum_{m=0}^{\infty} (m+1)b_{m+1} t^m dt \\ &= \sum_{m,n=0}^{\infty} \frac{e^{in\theta_1}}{r_2^{n/2} n!} (m+1)b_{m+1} g^{(n)}(z_0) t^{m+n} dt. \end{aligned}$$

Now introduce  $l = m + n$  and note that the summation over  $n$  goes up to  $l$ . This gives

$$g(z) dz = \sum_{l=0}^{\infty} \underbrace{\sum_{n=0}^l \frac{e^{in\theta_1}}{r_2^{n/2} n!} (l-n+1)b_{l-n+1} g^{(n)}(z_0)}_{\equiv a_l} t^l dt = \sum_{l=0}^{\infty} a_l t^l dt.$$

Substituting this in Equation (11.30) and changing the contour integration into the integral from  $-\infty$  to  $\infty$  as before yields

$$I(\alpha) \approx e^{\alpha f(z_0)} \sum_{k=0}^{\infty} a_{2k} \alpha^{-k-1/2} \Gamma(k + \frac{1}{2}),$$

$$a_{2k} = \sum_{n=0}^{2k} \frac{e^{in\theta_1}}{r_2^{n/2} n!} (2k - n + 1) b_{2k-n+1} g^{(n)}(z_0). \quad (11.36)$$

The only thing left to do is to evaluate  $b_m$ . We shall not give a general formula for these coefficients. Instead, we shall calculate the first three of them. This should reveal to the reader the general method of approximating them to any order. We have already calculated  $b_1$  in Equation (11.31). To calculate  $b_2$ , keep the next-highest term in the expansion of both  $z$  and  $t^2$ . Thus write

$$z - z_0 = b_1 t + b_2 t^2, \quad t^2 = -\frac{1}{2} f''(z_0)(z - z_0)^2 - \frac{1}{6} f'''(z_0)(z - z_0)^3.$$

Now substitute the first equation in the second and equate the coefficients of equal powers of  $t$  on both sides. The second power of  $t$  gives nothing new: It merely reaffirms the value of  $b_1$ . The coefficient of the third power of  $t$  is  $-b_1 b_2 f''(z_0) - \frac{1}{6} b_1^3 f'''(z_0)$ . Setting this equal to zero gives

$$b_2 = -\frac{b_1^2 f'''(z_0)}{6 f''(z_0)} = \frac{f'''(z_0)}{3 |f''(z_0)|^2} e^{4i\theta_1}, \quad (11.37)$$

where we substituted for  $b_1$  from Equation (11.31) and used  $2\theta_1 + \theta_2 = \pi$ .

To calculate  $b_3$ , keep one more term in the expansion of both  $z$  and  $t^2$  to obtain

$$z - z_0 = b_1 t + b_2 t^2 + b_3 t^3$$

and

$$t^2 = -\frac{1}{2} f''(z_0)(z - z_0)^2 - \frac{1}{6} f'''(z_0)(z - z_0)^3 - \frac{1}{24} f^{(iv)}(z_0)(z - z_0)^4.$$

Once again substitute the first equation in the second and equate the coefficients of equal powers of  $t$  on both sides. The second and third powers of  $t$  give nothing new. Setting the coefficient of the fourth power of  $t$  equal to zero yields

$$b_3 = b_1^3 \left\{ \frac{5[f'''(z_0)]^2}{72[f''(z_0)]^2} - \frac{f^{(iv)}}{24 f''(z_0)} \right\}$$

$$= \frac{\sqrt{2} e^{3i\theta_1}}{12 |f''(z_0)|^{3/2}} \left\{ \frac{5[f'''(z_0)]^2}{3[f''(z_0)]^2} - \frac{f^{(iv)}}{f''(z_0)} \right\}. \quad (11.38)$$



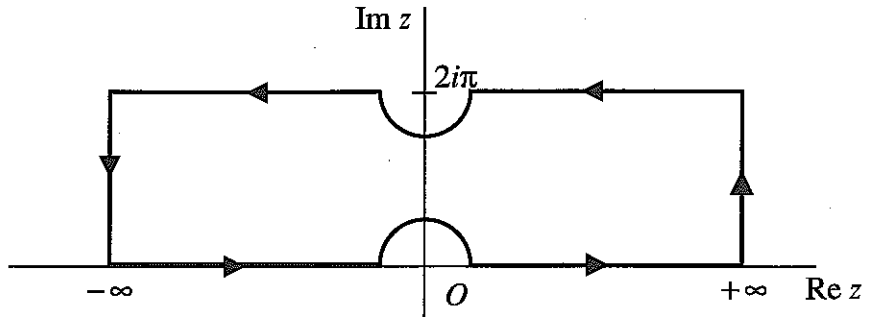


Figure 11.14 Contour used for Problem 11.4.

## 11.6 Problems

11.1. Derive Equation (11.2) from its logarithmic derivative.

11.2. Show that the point at infinity is not a branch point for  $f(z) = (z^2 - 1)^{1/2}$ .

11.3. Find the following integrals, for which  $0 \neq a \in \mathbb{R}$ .

$$(a) \int_0^\infty \frac{\ln x}{(x^2 + a^2)^2} dx. \quad (b) \int_0^\infty \frac{\ln x}{(x^2 + a^2)^2 \sqrt{x}} dx. \quad (c) \int_0^\infty \frac{(\ln x)^2}{x^2 + a^2} dx.$$

11.4. Use the contour in Figure 11.14 to evaluate the following integrals.

$$(a) \int_0^\infty \frac{\sin ax}{\sinh x} dx \qquad (b) \int_0^\infty \frac{x \cos ax}{\sinh x} dx$$

11.5. Show that  $\int_0^\pi f(\sin \theta) d\theta = 2 \int_0^{\pi/2} f(\sin \theta) d\theta$  for an arbitrary function  $f$  defined in the interval  $[-1, +1]$ .

11.6. Find the principal value of the integral  $\int_{-\infty}^\infty x \sin x dx / (x^2 - x_0^2)$  and evaluate

$$I = \int_{-\infty}^\infty \frac{x \sin x}{(x - x_0 \pm i\epsilon)(x + x_0 \pm i\epsilon)} dx$$

for the four possible choices of signs.

11.7. Use analytic continuation, the analyticity of the exponential, hyperbolic, and trigonometric functions, and the analogous identities for real  $z$  to prove the following identities.

$$(a) e^z = \cosh z + \sinh z.$$

$$(b) \cosh^2 z - \sinh^2 z = 1.$$

$$(c) \sin 2z = 2 \sin z \cos z.$$

**11.8.** Show that the function  $1/z^2$  represents the analytic continuation into the domain  $\mathbb{C} - \{0\}$  (all the complex plane minus the origin) of the function defined by  $\sum_{n=0}^{\infty} (n+1)(z+1)^n$  where  $|z+1| < 1$ .

**11.9.** Find the analytic continuation into  $\mathbb{C} - \{i, -i\}$  (all the complex plane except  $i$  and  $-i$ ) of  $f(z) = \int_0^{\infty} e^{-zt} \sin t \, dt$  where  $\operatorname{Re}(z) > 0$ .

**11.10.** Expand  $f(z) = \sum_{n=0}^{\infty} z^n$  (defined in its circle of convergence) in a Taylor series about  $z = a$ . For what values of  $a$  does this expansion permit the function  $f(z)$  to be continued analytically?

**11.11.** The two power series

$$f_1(z) = \sum_{n=1}^{\infty} \frac{z^n}{n} \quad \text{and} \quad f_2(z) = i\pi + \sum_{n=1}^{\infty} (-1)^n \frac{(z-2)^n}{n}$$

have no common domain of convergence. Show that they are nevertheless analytic continuations of one another.

**11.12.** Prove that the functions defined by the two series

$$1 + az + a^2 z^2 + \dots \quad \text{and} \quad \frac{1}{1-z} - \frac{(1-a)z}{(1-z)^2} + \frac{(1-a)^2 z^2}{(1-z)^3} - \dots$$

are analytic continuations of one another.

**11.13.** Show that the function  $f_1(z) = 1/(z^2 + 1)$ , where  $z \neq \pm i$ , is the analytic continuation into  $\mathbb{C} - \{i, -i\}$  of the function  $f_2(z) = \sum_{n=0}^{\infty} (-1)^n z^{2n}$ , where  $|z| < 1$ .

**11.14.** Find the analytic continuation into  $\mathbb{C} - \{0\}$  of the function

$$f(z) = \int_0^{\infty} t e^{-zt} \, dt \quad \text{where } \operatorname{Re}(z) > 0.$$

**11.15.** Show that the integral in Equation (11.9) converges. Hint: First show that  $|\Gamma(z+1)| \leq \int_0^{\infty} t^x e^{-t} \, dt$  where  $x = \operatorname{Re}(z)$ . Now show that

$$\int_0^{\infty} t^x e^{-t} \, dt \leq \int_0^1 t^x e^{-t} \, dt + \int_0^{\infty} t^n e^{-t} \, dt \quad \text{for some integer } n > 0$$

and conclude that  $\Gamma(z)$  is finite.

**11.16.** Show that  $d\Gamma(z+1)/dz$  exists and is finite by establishing the following:  
(a)  $|\ln t| < t + 1/t$  for  $t > 0$ . Hint: For  $t \geq 1$ , show that  $t - \ln t$  is a monotonically increasing function. For  $t < 1$ , make the substitution  $t = 1/s$ .

(b) Use the result from part (a) in the integral for  $d\Gamma(z+1)/dz$  to show that  $|d\Gamma(z+1)/dz|$  is finite. Hint: Differentiate inside the integral.

**11.17.** Derive Equation (11.11) from Equation (11.9).

**11.18.** Show that  $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ , and that

$$(2k-1)!! \equiv (2k-1)(2k-3)\cdots 5\cdot 3\cdot 1 = \frac{2^k}{\sqrt{\pi}} \Gamma\left(\frac{2k+1}{2}\right).$$

**11.19.** Show that  $\Gamma(z) = \int_0^1 [\ln(1/t)]^{z-1} dt$  with  $\operatorname{Re}(z) > 0$ .

**11.20.** Derive the identity  $\int_0^\infty e^{-x^\alpha} dx = \Gamma[(\alpha+1)/\alpha]$ .

**11.21.** Consider the function  $f(z) = (1+z)^\alpha$ .

(a) Show that  $d^n f/dz^n|_{z=0} = \Gamma(\alpha+1)/\Gamma(\alpha-n+1)$ , and use it to derive the relation

$$(1+z)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} z^n, \quad \text{where } \binom{\alpha}{n} \equiv \frac{\alpha!}{n!(\alpha-n)!} \equiv \frac{\Gamma(\alpha+1)}{n!\Gamma(\alpha-n+1)}.$$

(b) Show that for general complex numbers  $a$  and  $b$  we can formally write

$$(a+b)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} a^n b^{\alpha-n}.$$

(c) Show that if  $\alpha$  is a positive integer  $m$ , the series in part (b) truncates at  $n = m$ .

**11.22.** Prove that the residue of  $\Gamma(z)$  at  $z = -k$  is  $r_k = (-1)^k/k!$ . Hint: Use Equation (11.12)

**11.23.** Derive the following relation for  $z = x + iy$ :

$$|\Gamma(z)| = \Gamma(x) \prod_{k=0}^{\infty} \left[ 1 + \frac{y^2}{(x+k)^2} \right]^{-1/2}.$$

**11.24.** Using the definition of  $B(a, b)$ , Equation (11.16), show that  $B(a, b) = B(b, a)$ .

**11.25.** Integrate Equation (11.21) by parts and derive Equation (11.11).

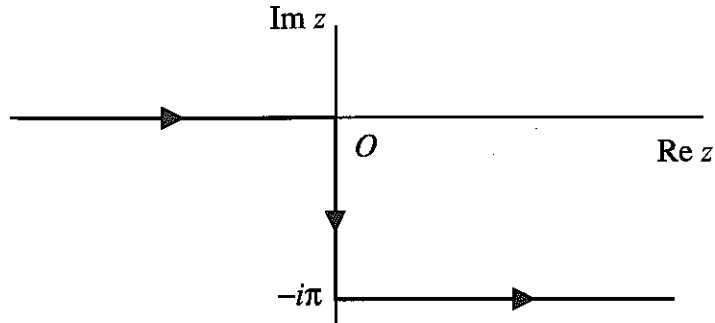
**11.26.** For positive integers  $n$ , show that  $\Gamma(\frac{1}{2} - n)\Gamma(\frac{1}{2} + n) = (-1)^n \pi$ .

**11.27.** Show that

$$(a) B(a, b) = B(a+1, b) + B(a, b+1).$$

$$(b) B(a, b+1) = \left(\frac{b}{a+b}\right) B(a, b).$$

$$(c) B(a, b)B(a+b, c) = B(b, c)B(a, b+c).$$



**Figure 11.15** The contour for the evaluation of the Hankel function of the second kind.

**11.28.** Verify that  $\int_{-1}^1 (1+t)^a (1-t)^b dt = 2^{a+b+1} B(a+1, b+1)$ .

**11.29.** Show that the volume of the solid formed by the surface  $z = x^a y^b$ , the  $xy$ -,  $yz$ -, and  $xz$ -planes, and the plane parallel to the  $z$ -axis and going through the points  $(0, y_0)$  and  $(x_0, 0)$  is

$$\frac{x_0^{a+1} y_0^{b+1}}{a+b+2} B(a+1, b+1).$$

**11.30.** Derive this relation:

$$\int_0^\infty \frac{\sinh^a x}{\cosh^b x} dx = \frac{1}{2} B\left(\frac{a+1}{2}, \frac{b-a}{2}\right) \quad \text{where } -1 < a < b.$$

Hint: Let  $t = \tanh^2 x$  in Equation (11.16).

**11.31.** The Hankel function of the second kind is defined as

$$H_\nu^{(2)}(\alpha) \equiv \frac{1}{i\pi} \int_C e^{(\alpha/2)(z-1/z)} \frac{dz}{z^{\nu+1}},$$

where  $C$  is the contour shown in Figure 11.15. Find the asymptotic expansion of this function.

**11.32.** Find the asymptotic dependence of the modified Bessel function of the first kind, defined as

$$I_\nu(\alpha) \equiv \frac{1}{2\pi i} \oint_C e^{(\alpha/2)(z+1/z)} \frac{dz}{z^{\nu+1}},$$

where  $C$  starts at  $-\infty$ , approaches the origin and circles it, and goes back to  $-\infty$ . Thus the negative real axis is excluded from the domain of analyticity.

11.33. Find the asymptotic dependence of the *modified Bessel function of the second kind*:

$$K_\nu(\alpha) \equiv \frac{1}{2} \int_C e^{-(\alpha/2)(z+1/z)} \frac{dz}{z^{\nu+1}},$$

where  $C$  starts at  $\infty$ , approaches the origin and circles it, and goes back to  $\infty$ . Thus the positive real axis is excluded from the domain of analyticity.

### Additional Reading

1. Dennery, P. and Krzywicki, A. *Mathematics for Physicists*, Harper and Row, 1967.
2. Lang, S. *Complex Analysis*, 2nd ed., Springer-Verlag, 1985. Contains a very lucid discussion of analytic continuation.



Part IV

---

# Differential Equations





# Separation of Variables in Spherical Coordinates

The laws of physics are almost exclusively written in the form of **differential equations** (DEs). In (point) particle mechanics there is only one independent variable, leading to **ordinary differential equations** (ODEs). In other areas of physics in which extended objects such as fields are studied, variations with respect to position are also important. Partial derivatives with respect to coordinate variables show up in the differential equations, which are therefore called **partial differential equations** (PDEs). We list the most common PDEs of mathematical physics in the following.

## 12.1 PDEs of Mathematical Physics

In electrostatics, where time-independent scalar fields such as potentials and vector fields such as electrostatic fields are studied, the law is described by **Poisson's equation**,

$$\nabla^2 \Phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}). \quad (12.1)$$

In vacuum, where  $\rho(\mathbf{r}) = 0$ , Equation (12.1) reduces to **Laplace's equation**,

$$\nabla^2 \Phi(\mathbf{r}) = 0. \quad (12.2)$$

Many electrostatic problems involve conductors held at constant potentials and situated in vacuum. In the space between such conducting surfaces, the electrostatic potential obeys Equation (12.2).

The most simplified version of the **heat equation** is

$$\frac{\partial T}{\partial t} = a^2 \nabla^2 T(\mathbf{r}), \quad (12.3)$$

where  $T$  is the temperature and  $a$  is a constant characterizing the medium in which heat is flowing.

One of the most frequently recurring PDEs encountered in mathematical physics is the **wave equation**,

$$\nabla^2 \Psi - \frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} = 0. \quad (12.4)$$

This equation (or its simplification to lower dimensions) is applied to the vibration of strings and drums; the propagation of sound in gases, solids, and liquids; the propagation of disturbances in plasmas; and the propagation of electromagnetic waves.

The **Schrödinger equation**, describing nonrelativistic quantum phenomena, is

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi + V(\mathbf{r})\Psi = -i\hbar \frac{\partial \Psi}{\partial t}, \quad (12.5)$$

where  $m$  is the mass of a subatomic particle,  $\hbar$  is Planck's constant (divided by  $2\pi$ ),  $V$  is the potential energy of the particle, and  $|\Psi(\mathbf{r}, t)|^2$  is the probability density of finding the particle at  $\mathbf{r}$  at time  $t$ .

A relativistic generalization of the Schrödinger equation for a free particle of mass  $m$  is the **Klein-Gordon equation**, which, in terms of the natural units ( $\hbar = 1 = c$ ), reduces to

$$\nabla^2 \phi - m^2 \phi = \frac{\partial^2 \phi}{\partial t^2}. \quad (12.6)$$

Equations (12.3–12.6) have partial derivatives with respect to time. As a first step toward solving these PDEs and as an introduction to similar techniques used in the solution of PDEs not involving time,<sup>1</sup> let us separate the time variable. We will denote the functions in all four equations by the generic symbol  $\Psi(\mathbf{r}, t)$ . The basic idea is to separate the  $\mathbf{r}$  and  $t$  dependence into factors:  $\Psi(\mathbf{r}, t) \equiv R(\mathbf{r})T(t)$ . This factorization permits us to separate the two operations of space differentiation and time differentiation. Let  $\mathbf{L}$  stand for all spatial derivative operators and write all the relevant equations either as  $\mathbf{L}\Psi = \partial\Psi/\partial t$  or as  $\mathbf{L}\Psi = \partial^2\Psi/\partial t^2$ . With this notation and the above separation, we have

$$\mathbf{L}(RT) = T(\mathbf{L}R) = \begin{cases} RdT/dt, \\ Rd^2T/dt^2. \end{cases}$$

<sup>1</sup>See [Hass 99] for a thorough discussion of separation in Cartesian and cylindrical coordinates. Chapter 19 of this book also contains examples of solutions to some second-order linear DEs resulting from such separation.

time is separated  
from space

Klein-Gordon  
equation

Schrödinger equation

wave equation

Dividing both sides by  $RT$ , we obtain

$$\frac{1}{R}\mathbf{L}R = \begin{cases} \frac{1}{T}\frac{dT}{dt}, \\ \frac{1}{T}\frac{d^2T}{dt^2}. \end{cases} \quad (12.7)$$

Now comes the crucial step in the process of separation of variables. The LHS of Equation (12.7) is a function of *position alone*, and the RHS is a function of *time alone*. Since  $\mathbf{r}$  and  $t$  are independent variables, the only way that (12.7) can hold is for *both sides to be constant*, say  $\alpha$ :

$$\frac{1}{R}\mathbf{L}R = \alpha \Rightarrow \mathbf{L}R = \alpha R$$

and<sup>2</sup>

$$\frac{1}{T}\frac{dT}{dt} = \alpha \Rightarrow \frac{dT}{dt} = \alpha T \quad \text{or} \quad \frac{1}{T}\frac{d^2T}{dt^2} = \alpha \Rightarrow \frac{d^2T}{dt^2} = \alpha T.$$

We have reduced the original time-dependent PDE to an ODE,

$$\frac{dT}{dt} = \alpha T \quad \text{or} \quad \frac{d^2T}{dt^2} = \alpha T, \quad (12.8)$$

and a PDE involving only the position variables,  $(\mathbf{L} - \alpha)R = 0$ . The most general form of  $\mathbf{L} - \alpha$  arising from Equations (12.3–12.6) is  $\mathbf{L} - \alpha \equiv \nabla^2 + f(\mathbf{r})$ . Therefore, Equations (12.3–12.6) are equivalent to (12.8), and

$$\nabla^2 R + f(\mathbf{r})R = 0. \quad (12.9)$$

To include Poisson's equation, we replace the zero on the RHS by  $g(\mathbf{r}) \equiv -4\pi\rho(\mathbf{r})$ , obtaining  $\nabla^2 R + f(\mathbf{r})R = g(\mathbf{r})$ . With the exception of Poisson's equation (an inhomogeneous PDE), in all the foregoing equations the term on the RHS is zero.<sup>3</sup> We will restrict ourselves to this so-called *homogeneous* case and rewrite (12.9) as

$$\nabla^2 \Psi(\mathbf{r}) + f(\mathbf{r})\Psi(\mathbf{r}) = 0. \quad (12.10)$$

Depending on the geometry of the problem, Equation (12.10) is further separated into ODEs each involving a single coordinate of a suitable coordinate system. We shall see examples of all major coordinate systems (Cartesian, cylindrical, and

<sup>2</sup>In most cases,  $\alpha$  is chosen to be real. In the case of the Schrödinger equation, it is more convenient to choose  $\alpha$  to be purely imaginary so that the  $i$  in the definition of  $\mathbf{L}$  can be compensated. In all cases, the precise nature of  $\alpha$  is determined by boundary conditions.

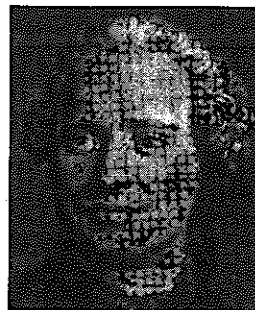
<sup>3</sup>Techniques for solving inhomogeneous PDEs are discussed in Chapters 21 and 22.

spherical) in Chapter 19. For the rest of this chapter, we shall concentrate on some general aspects of the spherical coordinates.

---

**Jean Le Rond d'Alembert** (1717–1783) was the illegitimate son of a famous salon hostess of eighteenth-century Paris and a cavalry officer. Abandoned by his mother, d'Alembert was raised by a foster family and later educated by the arrangement of his father at a nearby church-sponsored school, in which he received instruction in the classics and above-average instruction in mathematics. After studying law and medicine, he finally chose to pursue a career in mathematics. In the 1740s he joined the ranks of the *philosophes*, a growing group of deistic and materialistic thinkers and writers who actively questioned the social and intellectual standards of the day. He traveled little (he left France only once, to visit the court of Frederick the Great), preferring instead the company of his friends in the salons, among whom he was well known for his wit and laughter.

D'Alembert turned his mathematical and philosophical talents to many of the outstanding scientific problems of the day, with mixed success. Perhaps his most famous scientific work, entitled *Traité de dynamique*, shows his appreciation that a revolution was taking place in the science of mechanics—the formalization of the principles stated by Newton into a rigorous mathematical framework. The philosophy to which d'Alembert subscribed, however, refused to acknowledge the primacy of a concept as unclear and arbitrary as “force,” introducing a certain awkwardness to his treatment and perhaps causing him to overlook the important principle of conservation of energy. Later, d'Alembert produced a treatise on fluid mechanics (the priority of which is still debated by historians), a paper dealing with vibrating strings (in which the wave equation makes its first appearance in physics), and a skillful treatment of celestial mechanics. D'Alembert is also credited with use of the first partial differential equation as well as the first solution to such an equation using **separation of variables**. (One should be careful interpreting “first”: many of d'Alembert's predecessors and contemporaries gave similar, though less satisfactory, treatments of these milestones.) Perhaps his most well-known contribution to mathematics (at least among students) is the ratio test for the convergence of infinite series.



Much of the work for which d'Alembert is remembered occurred outside mathematical physics. He was chosen as the science editor of the *Encyclopédie*, and his lengthy *Discours Préliminaire* in that volume is considered one of the defining documents of the Enlightenment. Other works included writings on law, religion, and music.

Since d'Alembert's final years were not especially happy ones, perhaps this account of his life should end with a glimpse at the humanity his philosophy often gave his work. Like many of his contemporaries, he considered the problem of calculating the relative risk associated with the new practice of smallpox inoculation, which in rare cases caused the disease it was designed to prevent. Although not very successful in the mathematical sense, he was careful to point out that the probability of accidental infection, however slight or elegantly derived, would be small consolation to a father whose child died from the

inoculation. It is greatly to his credit that d'Alembert did not believe such considerations irrelevant to the problem.

## 12.2 Separation of the Angular Part of the Laplacian

With Cartesian and cylindrical variables, the boundary conditions are important in determining the nature of the solutions of the ODE obtained from the PDE. In almost all applications, however, the angular part of the spherical variables can be separated and studied very generally. This is because the angular part of the Laplacian in the spherical coordinate system is closely related to the operation of rotation and the angular momentum, which are independent of any particular situation.

The separation of the angular part in spherical coordinates can be done in a fashion exactly analogous to the separation of time by writing  $\Psi$  as a product of three functions, each depending on only one of the variables. However, we will follow an approach that is used in quantum mechanical treatments of angular momentum. This approach, which is based on the operator algebra of Chapter 2 and is extremely powerful and elegant, gives solutions for the angular part in closed form.

Define the vector operator  $\vec{\mathbf{p}}$  as  $\vec{\mathbf{p}} = -i\nabla$  so that its  $j$ th Cartesian component is  $\mathbf{p}_j = -i\partial/\partial x_j$ , for  $j = 1, 2, 3$ . In quantum mechanics  $\vec{\mathbf{p}}$  (multiplied by  $\hbar$ ) is the momentum operator. It is easy to verify that<sup>4</sup>  $[x_j, \mathbf{p}_k] = i\delta_{jk}$  and  $[x_j, x_k] = 0 = [\mathbf{p}_j, \mathbf{p}_k]$ .

angular momentum  
operator  
commutation  
relations between  
components of  
angular momentum  
operator

We can also define the **angular momentum operator** as  $\vec{\mathbf{L}} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}$ . This is expressed in components as  $L_i = (\vec{\mathbf{r}} \times \vec{\mathbf{p}})_i = \epsilon_{ijk}x_j\mathbf{p}_k$  for  $i = 1, 2, 3$ , where Einstein's summation convention (summing over repeated indices) is utilized.<sup>5</sup> Using the commutation relations above, we obtain

$$[L_j, L_k] = i\epsilon_{jki}L_i.$$

We will see shortly that  $\vec{\mathbf{L}}$  can be written solely in terms of the angles  $\theta$  and  $\varphi$ . Moreover, there is one factor of  $\vec{\mathbf{p}}$  in the definition of  $\vec{\mathbf{L}}$ , so if we square  $\vec{\mathbf{L}}$ , we will get two factors of  $\vec{\mathbf{p}}$ , and a Laplacian may emerge in the expression for  $\vec{\mathbf{L}} \cdot \vec{\mathbf{L}}$ . In this manner, we may be able to write  $\nabla^2$  in terms of  $L^2$ , which depends only on

<sup>4</sup>These operators act on the space of functions possessing enough "nice" properties as to render the space suitable. The operator  $x_j$  simply multiplies functions, while  $\mathbf{p}_j$  differentiates them.

<sup>5</sup>It is assumed that the reader is familiar with vector algebra using indices and such objects as  $\delta_{ij}$  and  $\epsilon_{ijk}$ . For an introductory treatment, sufficient for our present discussion, see [Hass 99]. A more advanced treatment of these objects (tensors) can be found in Part VII of this book.

angles. Let us try this:

$$\begin{aligned} \mathbf{L}^2 &= \vec{\mathbf{L}} \cdot \vec{\mathbf{L}} = \sum_{i=1}^3 \mathbf{L}_i \mathbf{L}_i = \epsilon_{ijk} x_j \mathbf{p}_k \epsilon_{imn} x_m \mathbf{p}_n = \epsilon_{ijk} \epsilon_{imn} x_j \mathbf{p}_k x_m \mathbf{p}_n \\ &= (\delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}) x_j \mathbf{p}_k x_m \mathbf{p}_n = x_j \mathbf{p}_k x_j \mathbf{p}_k - x_j \mathbf{p}_k x_k \mathbf{p}_j. \end{aligned}$$

We need to write this expression in such a way that factors with the same index are next to each other, to give a dot product. We must also try, when possible, to keep the  $\vec{\mathbf{p}}$  factors to the right so that they can operate on functions without intervention from the  $x$  factors. We do this using the commutation relations between the  $x$ 's and the  $\mathbf{p}$ 's:

$$\begin{aligned} \mathbf{L}^2 &= x_j (x_j \mathbf{p}_k - i \delta_{kj}) \mathbf{p}_k - (\mathbf{p}_k x_j + i \delta_{kj}) x_k \mathbf{p}_j \\ &= x_j x_j \mathbf{p}_k \mathbf{p}_k - i x_j \mathbf{p}_j - \mathbf{p}_k x_k x_j \mathbf{p}_j - i x_j \mathbf{p}_j \\ &= x_j x_j \mathbf{p}_k \mathbf{p}_k - 2i x_j \mathbf{p}_j - (x_k \mathbf{p}_k - i \delta_{kk}) x_j \mathbf{p}_j. \end{aligned}$$

Recalling that  $\delta_{kk} = \sum_{k=1}^3 \delta_{kk} = 3$  and  $x_j x_j = \sum_{j=1}^3 x_j x_j = \vec{\mathbf{r}} \cdot \vec{\mathbf{r}} = r^2$  etc., we can write  $\mathbf{L}^2 = r^2 \vec{\mathbf{p}} \cdot \vec{\mathbf{p}} + i \vec{\mathbf{r}} \cdot \vec{\mathbf{p}} - (\vec{\mathbf{r}} \cdot \vec{\mathbf{p}})(\vec{\mathbf{r}} \cdot \vec{\mathbf{p}})$ , which, if we make the substitution  $\vec{\mathbf{p}} = -i \nabla$ , yields

$$\nabla^2 = -r^{-2} \mathbf{L}^2 + r^{-2} (\mathbf{r} \cdot \nabla)(\mathbf{r} \cdot \nabla) + r^{-2} \mathbf{r} \cdot \nabla.$$

Letting both sides act on the function  $\Psi(r, \theta, \varphi)$ , we get

$$\nabla^2 \Psi = -\frac{1}{r^2} \mathbf{L}^2 \Psi + \frac{1}{r^2} (\mathbf{r} \cdot \nabla)(\mathbf{r} \cdot \nabla) \Psi + \frac{1}{r^2} \mathbf{r} \cdot \nabla \Psi. \quad (12.11)$$

But we note that  $\mathbf{r} \cdot \nabla = r \hat{\mathbf{e}}_r \cdot \nabla = r \partial / \partial r$ . We thus get the final form of  $\nabla^2 \Psi$  in spherical coordinates:

$$\nabla^2 \Psi = -\frac{1}{r^2} \mathbf{L}^2 \Psi + \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \Psi}{\partial r} \right) + \frac{1}{r} \frac{\partial \Psi}{\partial r}. \quad (12.12)$$

It is important to note that Equation (12.11) is a general relation that holds in all coordinate systems. Although all the manipulations leading to it were done in Cartesian coordinates, since it is written in vector notation, there is no indication in the final form that it was derived using specific coordinates.

Equation (12.12) is the spherical version of (12.11) and is the version we shall use. We will first make the simplifying assumption that in Equation (12.10), the master equation,  $f(\mathbf{r})$  is a function of  $r$  only. Equation (12.10) then becomes

$$-\frac{1}{r^2} \mathbf{L}^2 \Psi + \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \Psi}{\partial r} \right) + \frac{1}{r} \frac{\partial \Psi}{\partial r} + f(r) \Psi = 0.$$

Assuming, for the time being, that  $\mathbf{L}^2$  depends only on  $\theta$  and  $\varphi$ , and separating  $\Psi$  into a product of two functions,  $\Psi(r, \theta, \varphi) = R(r)Y(\theta, \varphi)$ , we can rewrite this

Laplacian separated  
into angular and  
radial parts

equation as

$$-\frac{1}{r^2} \mathbf{L}^2(RY) + \frac{1}{r} \frac{\partial}{\partial r} \left[ r \frac{\partial}{\partial r} (RY) \right] + \frac{1}{r} \frac{\partial}{\partial r} (RY) + f(r)RY = 0.$$

Dividing by  $RY$  and multiplying by  $r^2$  yields

$$\underbrace{-\frac{1}{Y} \mathbf{L}^2(Y)}_{-\alpha} + \underbrace{\frac{r}{R} \frac{d}{dr} \left( r \frac{dR}{dr} \right) + \frac{r}{R} \frac{dR}{dr} + r^2 f(r)}_{+\alpha} = 0,$$

or

$$\mathbf{L}^2 Y(\theta, \varphi) = \alpha Y(\theta, \varphi) \quad (12.13)$$

and

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left[ f(r) - \frac{\alpha}{r^2} \right] R = 0. \quad (12.14)$$

We will concentrate on the angular part, Equation (12.13), leaving the radial part to the general discussion of ODEs. The rest of this subsection will focus on showing that  $\mathbf{L}_1 \equiv \mathbf{L}_x$ ,  $\mathbf{L}_2 \equiv \mathbf{L}_y$ , and  $\mathbf{L}_3 \equiv \mathbf{L}_z$  are independent of  $r$ .

Since  $\mathbf{L}_i$  is an operator, we can study its action on an arbitrary function  $f$ . Thus,  $\mathbf{L}_i f = -i \epsilon_{ijk} x_j \nabla_k f \equiv -i \epsilon_{ijk} x_j \partial f / \partial x_k$ . We can express the Cartesian  $x_j$  in terms of  $r$ ,  $\theta$ , and  $\varphi$ , and use the chain rule to express  $\partial f / \partial x_k$  in terms of spherical coordinates. This will give us  $\mathbf{L}_i f$  expressed in terms of  $r$ ,  $\theta$ , and  $\varphi$ . It will then emerge that  $r$  is absent in the final expression.

Let us start with  $x = r \sin \theta \cos \varphi$ ,  $y = r \sin \theta \sin \varphi$ ,  $z = r \cos \theta$ , and their inverses,  $r = (x^2 + y^2 + z^2)^{1/2}$ ,  $\cos \theta = z/r$ ,  $\tan \varphi = y/x$ , and express the Cartesian derivatives in terms of spherical coordinates using the chain rule. The first such derivative is

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial x} + \frac{\partial f}{\partial \varphi} \frac{\partial \varphi}{\partial x}. \quad (12.15)$$

The derivative of one coordinate system with respect to the other can be easily calculated. For example,  $\partial r / \partial x = x/r = \sin \theta \cos \varphi$ , and differentiating both sides of the equation  $\cos \theta = z/r$ , we obtain

$$-\sin \theta \frac{\partial \theta}{\partial x} = -\frac{z \partial r / \partial x}{r^2} = -\frac{zx}{r^3} = -\frac{\cos \theta \sin \theta \cos \varphi}{r} \Rightarrow \frac{\partial \theta}{\partial x} = \frac{\cos \theta \cos \varphi}{r}.$$

Finally, differentiating both sides of  $\tan \varphi = y/x$  with respect to  $x$  yields  $\partial \varphi / \partial x = -\sin \varphi / (r \sin \theta)$ . Using these expressions in Equation (12.15), we get

$$\frac{\partial f}{\partial x} = \sin \theta \cos \varphi \frac{\partial f}{\partial r} + \frac{\cos \theta \cos \varphi}{r} \frac{\partial f}{\partial \theta} - \frac{\sin \varphi}{r \sin \theta} \frac{\partial f}{\partial \varphi}.$$

In exactly the same way, we obtain

$$\frac{\partial f}{\partial y} = \sin \theta \sin \varphi \frac{\partial f}{\partial r} + \frac{\cos \theta \sin \varphi}{r} \frac{\partial f}{\partial \theta} + \frac{\cos \varphi}{r \sin \theta} \frac{\partial f}{\partial \varphi},$$

$$\frac{\partial f}{\partial z} = \cos \theta \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta}.$$

We can now calculate  $L_x$  by letting it act on an arbitrary function and expressing all Cartesian coordinates and derivatives in terms of spherical coordinates. The result is

$$L_x f = -iy \frac{\partial f}{\partial z} + iz \frac{\partial f}{\partial y} = i \left( \sin \varphi \frac{\partial}{\partial \theta} + \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right) f,$$

Cartesian components of angular momentum operator expressed in spherical coordinates or

$$L_x = i \left( \sin \varphi \frac{\partial}{\partial \theta} + \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right). \tag{12.16}$$

Analogous arguments yield

$$L_y = i \left( -\cos \varphi \frac{\partial}{\partial \theta} + \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right), \quad L_z = -i \frac{\partial}{\partial \varphi}. \tag{12.17}$$

angular momentum squared as differential operator in  $\theta$  and  $\varphi$

It is left as a problem for the reader to show that by adding the squares of the components of the angular momentum operator, one obtains

$$L^2 = -\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}. \tag{12.18}$$

Substitution in Equation (12.12) yields the familiar expression for the Laplacian in spherical coordinates.

### 12.3 Construction of Eigenvalues of $L^2$

Now that we have  $L^2$  in terms of  $\theta$  and  $\varphi$ , we could substitute in Equation (12.13), separate the  $\theta$  and  $\varphi$  dependence, and solve the corresponding ODEs. However, there is a much more elegant way of solving this problem algebraically, because Equation (12.13) is simply an eigenvalue equation for  $L^2$ . In this section, we will find the eigenvalues of  $L^2$ . The next section will evaluate the eigenvectors of  $L^2$ .

Let us consider  $L^2$  as an abstract operator and write (12.13) as

$$L^2 |Y\rangle = \alpha |Y\rangle,$$

where  $|Y\rangle$  is an abstract vector whose  $(\theta, \varphi)$ th component can be calculated later. Since  $L^2$  is a differential operator, it does not have a (finite-dimensional) matrix



representation. Thus, the determinantal procedure for calculating eigenvalues and eigenfunctions will not work here, and we have to find another way.

The equation above specifies an eigenvalue,  $\alpha$ , and an eigenvector,  $|Y\rangle$ . There may be more than one  $|Y\rangle$  corresponding to the same  $\alpha$ . To distinguish among these so-called *degenerate eigenvectors*, we choose a second operator, say  $L_3 \in \{L_i\}$  that commutes with  $L^2$ . This allows us to select a basis in which both  $L^2$  and  $L_3$  are diagonal, or, equivalently, a basis whose vectors are simultaneous eigenvectors of both  $L^2$  and  $L_3$ . This is possible by Theorem 4.4.15 and the fact that both  $L^2$  and  $L_3$  are hermitian operators in the space of square-integrable functions. (The proof is left as a problem.) In general, we would want to continue adding operators until we obtained a maximum set of commuting operators which could label the eigenvectors. In this case,  $L^2$  and  $L_3$  exhaust the set.<sup>6</sup> Using the more common subscripts  $x$ ,  $y$ , and  $z$  instead of 1, 2, 3 and attaching labels to the eigenvectors, we have

$$L^2 |Y_{\alpha,\beta}\rangle = \alpha |Y_{\alpha,\beta}\rangle, \quad L_z |Y_{\alpha,\beta}\rangle = \beta |Y_{\alpha,\beta}\rangle. \quad (12.19)$$

The hermiticity of  $L^2$  and  $L_z$  implies the reality of  $\alpha$  and  $\beta$ . Next we need to determine the possible values for  $\alpha$  and  $\beta$ .

Define two new operators  $L_+ \equiv L_x + iL_y$  and  $L_- \equiv L_x - iL_y$ . It is then easily verified that

$$[L^2, L_{\pm}] = 0, \quad [L_z, L_{\pm}] = \pm L_{\pm}, \quad [L_+, L_-] = 2L_z. \quad (12.20)$$

The first equation implies that  $L_{\pm}$  are invariant operators when acting in the subspace corresponding to the eigenvalue  $\alpha$ ; that is,  $L_{\pm} |Y_{\alpha,\beta}\rangle$  are eigenvectors of  $L^2$  with the same eigenvalue  $\alpha$ :

$$L^2(L_{\pm} |Y_{\alpha,\beta}\rangle) = L_{\pm}(L^2 |Y_{\alpha,\beta}\rangle) = \alpha L_{\pm} |Y_{\alpha,\beta}\rangle.$$

The second equation in (12.20) yields

$$\begin{aligned} L_z(L_+ |Y_{\alpha,\beta}\rangle) &= (L_z L_+) |Y_{\alpha,\beta}\rangle = (L_+ L_z + L_+) |Y_{\alpha,\beta}\rangle \\ &= L_+ L_z |Y_{\alpha,\beta}\rangle + L_+ |Y_{\alpha,\beta}\rangle = \beta L_+ |Y_{\alpha,\beta}\rangle + L_+ |Y_{\alpha,\beta}\rangle \\ &= (\beta + 1)L_+ |Y_{\alpha,\beta}\rangle. \end{aligned}$$

angular momentum  
raising and lowering  
operators

This indicates that  $L_+ |Y_{\alpha,\beta}\rangle$  has one more unit of the  $L_z$  eigenvalue than  $|Y_{\alpha,\beta}\rangle$  does. In other words,  $L_+$  *raises* the eigenvalue of  $L_z$  by one unit. That is why  $L_+$  is called a **raising operator**. Similarly,  $L_-$  is called a **lowering operator** because

$$L_z(L_- |Y_{\alpha,\beta}\rangle) = (\beta - 1)L_- |Y_{\alpha,\beta}\rangle.$$

We can summarize the above discussion as

$$L_{\pm} |Y_{\alpha,\beta}\rangle = C_{\pm} |Y_{\alpha,\beta \pm 1}\rangle,$$

<sup>6</sup>We could just as well have chosen  $L^2$  and any other component as our maximal set. However,  $L^2$  and  $L_3$  is the universally accepted choice.

where  $C_{\pm}$  are constants to be determined by a suitable normalization.

There are restrictions on (and relations between)  $\alpha$  and  $\beta$ . First note that as  $\mathbf{L}^2$  is a sum of squares of hermitian operators, it must be a positive operator; that is,  $\langle a | \mathbf{L}^2 | a \rangle \geq 0$  for all  $|a\rangle$ . In particular,

$$0 \leq \langle Y_{\alpha,\beta} | \mathbf{L}^2 | Y_{\alpha,\beta} \rangle = \alpha \langle Y_{\alpha,\beta} | Y_{\alpha,\beta} \rangle = \alpha \|Y_{\alpha,\beta}\|^2.$$

Therefore,  $\alpha \geq 0$ . Next, one can readily show that

$$\mathbf{L}^2 = \mathbf{L}_+ \mathbf{L}_- + \mathbf{L}_z^2 - \mathbf{L}_z = \mathbf{L}_- \mathbf{L}_+ + \mathbf{L}_z^2 + \mathbf{L}_z. \quad (12.21)$$

Sandwiching both sides of the first equality between  $|Y_{\alpha,\beta}\rangle$  and  $\langle Y_{\alpha,\beta}|$  yields

$$\langle Y_{\alpha,\beta} | \mathbf{L}^2 | Y_{\alpha,\beta} \rangle = \langle Y_{\alpha,\beta} | \mathbf{L}_+ \mathbf{L}_- | Y_{\alpha,\beta} \rangle + \langle Y_{\alpha,\beta} | \mathbf{L}_z^2 | Y_{\alpha,\beta} \rangle - \langle Y_{\alpha,\beta} | \mathbf{L}_z | Y_{\alpha,\beta} \rangle,$$

with an analogous expression involving  $\mathbf{L}_- \mathbf{L}_+$ . Using the fact that  $\mathbf{L}_+ = (\mathbf{L}_-)^{\dagger}$ , we get

$$\begin{aligned} \alpha \|Y_{\alpha,\beta}\|^2 &= \langle Y_{\alpha,\beta} | \mathbf{L}_+ \mathbf{L}_- | Y_{\alpha,\beta} \rangle + \beta^2 \|Y_{\alpha,\beta}\|^2 - \beta \|Y_{\alpha,\beta}\|^2 \\ &= \langle Y_{\alpha,\beta} | \mathbf{L}_- \mathbf{L}_+ | Y_{\alpha,\beta} \rangle + \beta^2 \|Y_{\alpha,\beta}\|^2 + \beta \|Y_{\alpha,\beta}\|^2 \\ &= \| \mathbf{L}_- | Y_{\alpha,\beta} \rangle \|^2 + \beta^2 \|Y_{\alpha,\beta}\|^2 + \beta \|Y_{\alpha,\beta}\|^2 \end{aligned} \quad (12.22)$$

Because of the positivity of norms, this yields  $\alpha \geq \beta^2 - \beta$  and  $\alpha \geq \beta^2 + \beta$ . Adding these two inequalities gives  $2\alpha \geq 2\beta^2 \Rightarrow -\sqrt{\alpha} \leq \beta \leq \sqrt{\alpha}$ . It follows that the values of  $\beta$  are bounded. That is, there exist a maximum  $\beta$ , denoted by  $\beta_+$ , and a minimum  $\beta$ , denoted by  $\beta_-$ , beyond which there are no more values of  $\beta$ . This can happen only if

$$\mathbf{L}_+ | Y_{\alpha,\beta_+} \rangle = 0, \quad \mathbf{L}_- | Y_{\alpha,\beta_-} \rangle = 0,$$

because if  $\mathbf{L}_{\pm} | Y_{\alpha,\beta_{\pm}} \rangle$  are not zero, then they must have values of  $\beta$  corresponding to  $\beta_{\pm} \pm 1$ , which are not allowed.

Using  $\beta_+$  for  $\beta$  in Equation (12.22) yields

$$(\alpha - \beta_+^2 - \beta_+) \|Y_{\alpha,\beta_+}\|^2 = 0.$$

By definition  $|Y_{\alpha,\beta_+}\rangle \neq 0$  (otherwise  $\beta_+ - 1$  would be the maximum). Thus, we obtain  $\alpha = \beta_+^2 + \beta_+$ . An analogous procedure using  $\beta_-$  for  $\beta$  yields  $\alpha = \beta_-^2 - \beta_-$ . We solve these two equations for  $\beta_+$  and  $\beta_-$ :

$$\beta_+ = \frac{1}{2}(-1 \pm \sqrt{1 + 4\alpha}), \quad \beta_- = \frac{1}{2}(1 \pm \sqrt{1 + 4\alpha}).$$

Since  $\beta_+ \geq \beta_-$  and  $\sqrt{1 + 4\alpha} \geq 1$ , we must choose

$$\beta_+ = \frac{1}{2}(-1 + \sqrt{1 + 4\alpha}) = -\beta_-.$$

Starting with  $|Y_{\alpha, \beta_+}\rangle$ , we can apply  $L_-$  to it repeatedly. In each step we decrease the value of  $\beta$  by one unit. There must be a limit to the number of vectors obtained in this way, because  $\beta$  has a minimum. Therefore, there must exist a nonnegative integer  $k$  such that

$$(L_-)^{k+1} |Y_{\alpha, \beta_+}\rangle = L_-(L_-^k |Y_{\alpha, \beta_+}\rangle) = 0.$$

Thus,  $L_-^k |Y_{\alpha, \beta_+}\rangle$  must be proportional to  $|Y_{\alpha, \beta_-}\rangle$ . In particular, since  $L_-^k |Y_{\alpha, \beta_+}\rangle$  has a  $\beta$  value equal to  $\beta_+ - k$ , we have  $\beta_- = \beta_+ - k$ . Now, using  $\beta_- = -\beta_+$  (derived above) yields the important result

$$\beta_+ = \frac{k}{2} \equiv j \quad \text{for } k \in \mathbb{N},$$

or  $\alpha = j(j+1)$ , since  $\alpha = \beta_+^2 + \beta_+$ . This result is important enough to be stated as a theorem.

eigenvalues of  $L^2$  and  $L_z$  given

**12.3.1. Theorem.** *The eigenvectors of  $L^2$ , denoted by  $|Y_{jm}\rangle$ , satisfy the eigenvalue relations*

$$\begin{aligned} L^2 |Y_{jm}\rangle &= j(j+1) |Y_{jm}\rangle, \\ L_z |Y_{jm}\rangle &= m |Y_{jm}\rangle, \end{aligned}$$

where  $j$  is a positive integer or half-integer, and  $m$  can take a value in the set  $\{-j, -j+1, \dots, j-1, j\}$  of  $2j+1$  numbers.

Let us briefly consider the normalization of the eigenvectors. We already know that the  $|Y_{jm}\rangle$ , being eigenvectors of the hermitian operators  $L^2$  and  $L_z$ , are orthogonal. We also demand that they be of unit norm; that is,

$$\langle Y_{jm} | Y_{j'm'} \rangle = \delta_{jj'} \delta_{mm'}. \quad (12.23)$$

This will determine the constants  $C_{\pm}$ , introduced earlier. Let us consider  $C_+$  first, which is defined by  $L_+ |Y_{jm}\rangle = C_+ |Y_{j, m+1}\rangle$ . The hermitian conjugate of this equation is  $\langle Y_{jm} | L_- = C_+^* \langle Y_{j, m+1} |$ . We contract these two equations to get  $\langle Y_{jm} | L_- L_+ |Y_{jm}\rangle = |C_+|^2 \langle Y_{j, m+1} | Y_{j, m+1}\rangle$ . Then we use the second relation in Equation (12.21), Theorem 12.3.1, and (12.23) to obtain

$$j(j+1) - m(m+1) = |C_+|^2 \Rightarrow |C_+| = \sqrt{j(j+1) - m(m+1)}.$$

Adopting the convention that the argument (phase) of the complex number  $C_+$  is zero (and therefore that  $C_+$  is real), we get

$$C_+ = \sqrt{j(j+1) - m(m+1)}$$

Similarly,  $C_- = \sqrt{j(j+1) - m(m-1)}$ . Thus, we get

$$\begin{aligned} L_+ |Y_{jm}\rangle &= \sqrt{j(j+1) - m(m+1)} |Y_{j, m+1}\rangle, \\ L_- |Y_{jm}\rangle &= \sqrt{j(j+1) - m(m-1)} |Y_{j, m-1}\rangle. \end{aligned} \quad (12.24)$$

**12.3.2. Example.** Let us find an expression for  $|Y_{lm}\rangle$  by repeatedly applying  $L_-$  to  $|Y_{ll}\rangle$ . The action for  $L_-$  is completely described by Equation (12.24). For the first power of  $L_-$ , we obtain

$$L_- |Y_{ll}\rangle = \sqrt{l(l+1) - l(l-1)} |Y_{l,l-1}\rangle = \sqrt{2l} |Y_{l,l-1}\rangle.$$

We apply  $L_-$  once more:

$$\begin{aligned} (L_-)^2 |Y_{ll}\rangle &= \sqrt{2l} L_- |Y_{l,l-1}\rangle = \sqrt{2l} \sqrt{l(l+1) - (l-1)(l-2)} |Y_{l,l-2}\rangle \\ &= \sqrt{2l} \sqrt{2(2l-1)} |Y_{l,l-2}\rangle = \sqrt{2(2l)(2l-1)} |Y_{l,l-2}\rangle. \end{aligned}$$

Applying  $L_-$  a third time yields

$$\begin{aligned} (L_-)^3 |Y_{ll}\rangle &= \sqrt{2(2l)(2l-1)} L_- |Y_{l,l-2}\rangle = \sqrt{2(2l)(2l-1)} \sqrt{6(l-1)} |Y_{l,l-3}\rangle \\ &= \sqrt{3!(2l)(2l-1)(2l-2)} |Y_{l,l-3}\rangle. \end{aligned}$$

The pattern suggests the following formula for a general power  $k$ :

$$L_-^k |Y_{ll}\rangle = \sqrt{k!(2l)(2l-1) \dots (2l-k+1)} |Y_{l,l-k}\rangle,$$

or  $L_-^k |Y_{ll}\rangle = \sqrt{k!(2l)!/(2l-k)!} |Y_{l,l-k}\rangle$ . If we set  $l-k = m$  and solve for  $|Y_{l,m}\rangle$ , we get

$$|Y_{l,m}\rangle = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}} L_-^{l-m} |Y_{ll}\rangle. \quad \blacksquare$$

The discussion in this section is the standard treatment of angular momentum in quantum mechanics. In the context of quantum mechanics, Theorem 12.3.1 states the far-reaching physical result that particles can have integer or half-integer spin. Such a conclusion is tied to the rotation group in three dimensions, which, in turn, is an example of a Lie group, or a continuous group of transformations. We shall come back to a study of groups later. It is worth noting that it was the study of differential equations that led the Norwegian mathematician Sophus Lie to the investigation of their symmetries and the development of the beautiful branch of mathematics and theoretical physics that bears his name. Thus, the existence of a connection between group theory (rotation, angular momentum) and the differential equation we are trying to solve should not come as a surprise.

## 12.4 Eigenvectors of $L^2$ : Spherical Harmonics

The treatment in the preceding section took place in an abstract vector space. Let us go back to the function space and represent the operators and vectors in terms of  $\theta$  and  $\varphi$ .

First, let us consider  $L_z$  in the form of a differential operator, as given in Equation (12.17). The eigenvalue equation for  $L_z$  becomes

$$-i \frac{\partial}{\partial \varphi} Y_{jm}(\theta, \varphi) = m Y_{jm}(\theta, \varphi).$$

We write  $Y_{jm}(\theta, \varphi) = P_{jm}(\theta)Q_{jm}(\varphi)$  and substitute in the above equation to obtain the ODE for  $\varphi$ ,  $dQ_{jm}/d\varphi = imQ_{jm}$ , which has a solution of the form  $Q_{jm}(\varphi) = C_{jm}e^{im\varphi}$ , where  $C_{jm}$  is a constant. Absorbing this constant into  $P_{jm}$ , we can write

$$Y_{jm}(\theta, \varphi) = P_{jm}(\theta)e^{im\varphi}.$$

In classical physics the value of functions must be the same at  $\varphi$  as at  $\varphi + 2\pi$ . This condition restricts the values of  $m$  to integers. In quantum mechanics, on the other hand, it is the absolute values of functions that are physically measurable quantities, and therefore  $m$  can also be a half-integer.

**12.4.1. Box.** From now on, we shall assume that  $m$  is an integer and denote the eigenvectors of  $L^2$  by  $Y_{lm}(\theta, \varphi)$ , in which  $l$  is a nonnegative integer.

Our task is to find an analytic expression for  $Y_{lm}(\theta, \varphi)$ . We need differential expressions for  $L_{\pm}$ . These can easily be obtained from the expressions for  $L_x$  and  $L_y$  given in Equations (12.16) and (12.17). (The straightforward manipulations are left as a problem.) We thus have

$$L_{\pm} = e^{\pm i\varphi} \left( \pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right). \quad (12.25)$$

Since  $l$  is the highest value of  $m$ , when  $L_+$  acts on  $Y_{ll}(\theta, \varphi) = P_{ll}(\theta)e^{il\varphi}$  the result must be zero. This leads to the differential equation

$$\left( \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right) [P_{ll}(\theta)e^{il\varphi}] = 0 \Rightarrow \left( \frac{d}{d\theta} - l \cot \theta \right) P_{ll}(\theta) = 0.$$

The solution to this differential equation is readily found to be

$$P_{ll}(\theta) = C_l (\sin \theta)^l.$$

The constant is subscripted because each  $P_{ll}$  may lead to a different constant of integration. We can now write  $Y_{ll}(\theta, \varphi) = C_l (\sin \theta)^l e^{il\varphi}$ .

With  $Y_{ll}(\theta, \varphi)$  at our disposal, we can obtain any  $Y_{lm}(\theta, \varphi)$  by repeated application of  $L_-$ . In principle, the result of Example 12.3.2 gives all the (abstract) eigenvectors. In practice, however, it is helpful to have a closed form (in terms of derivatives) for just the  $\theta$  part of  $Y_{lm}(\theta, \varphi)$ . So, let us apply  $L_-$ , as given in Equation (12.25) to  $Y_{ll}(\theta, \varphi)$ :

$$\begin{aligned} L_- Y_{ll} &= e^{-i\varphi} \left( -\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right) [P_{ll}(\theta)e^{il\varphi}] \\ &= e^{-i\varphi} \left[ -\frac{\partial}{\partial \theta} + i \cot \theta (il) \right] [P_{ll}(\theta)e^{il\varphi}] \\ &= (-1)e^{i(l-1)\varphi} \left( \frac{d}{d\theta} + l \cot \theta \right) P_{ll}(\theta). \end{aligned}$$

It can be shown that for a positive integer,

$$\left(\frac{d}{d\theta} + n \cot \theta\right) f(\theta) = \frac{1}{\sin^n \theta} \frac{d}{d\theta} [\sin^n \theta f(\theta)]. \quad (12.26)$$

Using this result yields

$$\begin{aligned} \mathbf{L}_- Y_{ll} &= \sqrt{2l} Y_{l,l-1} = \sqrt{2l} e^{i(l-1)\varphi} P_{l,l-1}(\theta) \\ &= (-1) e^{i(l-1)\varphi} \frac{1}{\sin^l \theta} \frac{d}{d\theta} [\sin^l \theta (C_l \sin^l \theta)] \\ &= (-1) C_l \frac{e^{i(l-1)\varphi}}{\sin^l \theta} \frac{d}{d\theta} (\sin^{2l} \theta). \end{aligned} \quad (12.27)$$

We apply  $\mathbf{L}_-$  to (12.27), and use Equation (12.26) with  $n = l - 1$  to obtain

$$\begin{aligned} \mathbf{L}_-^2 Y_{ll} &= (-1)^2 C_l e^{i(l-2)\varphi} \frac{1}{\sin^{l-1} \theta} \frac{d}{d\theta} \left[ \sin^{l-1} \theta \frac{1}{\sin^l \theta} \frac{d}{d\theta} (\sin^{2l} \theta) \right] \\ &= (-1)^2 C_l \frac{e^{i(l-2)\varphi}}{\sin^{l-1} \theta} \frac{d}{d\theta} \left[ \frac{1}{\sin \theta} \frac{d}{d\theta} (\sin^{2l} \theta) \right]. \end{aligned}$$

Making the substitution  $u = \cos \theta$  yields

$$\mathbf{L}_-^2 Y_{ll} = C_l \frac{e^{i(l-2)\varphi}}{(1-u^2)^{l/2-1}} \frac{d^2}{du^2} [(1-u^2)^l].$$

With a little more effort one can detect a pattern and obtain

$$\mathbf{L}_-^k Y_{ll} = C_l \frac{e^{i(l-k)\varphi}}{(1-u^2)^{(l-k)/2}} \frac{d^k}{du^k} [(1-u^2)^l].$$

If we let  $k = l - m$  and make use of the result obtained in Example 12.3.2, we obtain

$$Y_{lm}(\theta, \varphi) = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}} C_l \frac{e^{im\varphi}}{(1-u^2)^{m/2}} \frac{d^{l-m}}{du^{l-m}} [(1-u^2)^l].$$

To specify  $Y_{lm}(\theta, \varphi)$  completely, we need to evaluate  $C_l$ . Since  $C_l$  does not depend on  $m$ , we set  $m = 0$  in the above expression, obtaining

$$Y_{l0}(u, \varphi) = \frac{1}{\sqrt{(2l)!}} C_l \frac{d^l}{du^l} [(1-u^2)^l].$$

The RHS looks very much like the Legendre polynomials of Chapter 7. In fact,

$$Y_{l0}(u, \varphi) = \frac{C_l}{\sqrt{(2l)!}} (-1)^l 2^l l! P_l(u) \equiv A_l P_l(u). \quad (12.28)$$

Therefore, the normalization of  $Y_{l0}$  and the Legendre polynomials  $P_l$  determines  $C_l$ .

We now use Equation (6.9) to obtain the integral form of the orthonormality relation for  $Y_{lm}$ :

$$\begin{aligned} \delta_{ll'}\delta_{mm'} &= \langle Y_{l'm'} | Y_{lm} \rangle = \langle Y_{l'm'} | \left( \int_0^{2\pi} d\varphi \int_0^\pi \sin\theta d\theta |\theta, \varphi\rangle \langle \theta, \varphi| \right) | Y_{lm} \rangle \\ &= \int_0^{2\pi} d\varphi \int_0^\pi Y_{l'm'}^*(\theta, \varphi) Y_{lm}(\theta, \varphi) \sin\theta d\theta, \end{aligned} \quad (12.29)$$

which in terms of  $u = \cos\theta$  becomes

$$\int_0^{2\pi} d\varphi \int_{-1}^1 Y_{l'm'}^*(u, \varphi) Y_{lm}(u, \varphi) du = \delta_{ll'}\delta_{mm'}. \quad (12.30)$$

Problem 12.15 shows that using (12.29) one gets  $A_l = \sqrt{(2l+1)/(4\pi)}$ . Therefore, Equation (12.28) yields not only the value of  $C_l$ , but also the useful relation

$$Y_{l0}(u, \varphi) = \sqrt{\frac{2l+1}{4\pi}} P_l(u). \quad (12.31)$$

spherical harmonics

Substituting the value of  $C_l$  thus obtained, we finally get

$$Y_{lm}(\theta, \varphi) = (-1)^l \sqrt{\frac{2l+1}{4\pi}} \frac{e^{im\varphi}}{2^l l!} \sqrt{\frac{(l+m)!}{(l-m)!}} (1-u^2)^{-m/2} \frac{d^{l-m}}{du^{l-m}} [(1-u^2)^l], \quad (12.32)$$

where  $u = \cos\theta$ . These functions, the eigenfunctions of  $L^2$  and  $L_z$ , are called **spherical harmonics**. They occur frequently in those physical applications for which the Laplacian is expressed in terms of spherical coordinates.

One can immediately read off the  $\theta$  part of the spherical harmonics:

$$P_{lm}(u) = (-1)^l \sqrt{\frac{2l+1}{4\pi}} \frac{1}{2^l l!} \sqrt{\frac{(l+m)!}{(l-m)!}} (1-u^2)^{-m/2} \frac{d^{l-m}}{du^{l-m}} [(1-u^2)^l].$$

associated Legendre functions

However, this is not the version used in the literature. For historical reasons the **associated Legendre functions**  $P_l^m(u)$  are used. These are defined by

$$\begin{aligned} P_l^m(u) &= (-1)^m \sqrt{\frac{(l+m)!}{(l-m)!}} \sqrt{\frac{4\pi}{2l+1}} P_{lm}(u) \\ &= (-1)^{l+m} \frac{(l+m)!}{(l-m)!} \frac{(1-u^2)^{-m/2}}{2^l l!} \frac{d^{l-m}}{du^{l-m}} [(1-u^2)^l]. \end{aligned} \quad (12.33)$$

Thus,

$$Y_{lm}(\theta, \varphi) = (-1)^m \left[ \frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{1/2} P_l^m(\cos\theta) e^{im\varphi}. \quad (12.34)$$

We generated the spherical harmonics starting with  $Y_{ll}(\theta, \varphi)$  and applying the lowering operator  $L_-$ . We could have started with  $Y_{l,-l}(\theta, \varphi)$  instead, and applied the raising operator  $L_+$ . The latter procedure is identical to the former; nevertheless, we outline it below because of some important relations that emerge along the way. We first note that

$$|Y_{l,-m}\rangle = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}} L_+^{l-m} |Y_{l,-l}\rangle. \quad (12.35)$$

(This can be obtained following the steps of Example 12.3.2.) Next, we use  $L_- |Y_{l,-l}\rangle = 0$  in differential form to obtain

$$\left(\frac{d}{d\theta} - l \cot \theta\right) P_{l,-l}(\theta) = 0,$$

which has the same form as the differential equation for  $P_{ll}$ . Thus, the solution is  $P_{l,-l}(\theta) = C'_l (\sin \theta)^l$ , and

$$Y_{l,-l}(\theta, \varphi) = P_{l,-l}(\theta) e^{-il\varphi} = C'_l (\sin \theta)^l e^{-il\varphi}.$$

Applying  $L_+$  repeatedly yields

$$L_+^k Y_{l,-l}(u, \varphi) = C'_l \frac{(-1)^k e^{-i(l-k)\varphi}}{(1-u^2)^{(l-k)/2}} \frac{d^k}{du^k} [(1-u^2)^l],$$

where  $u = \cos \theta$ . Substituting  $k = l - m$  and using Equation (12.35) gives

$$Y_{l,-m}(u, \varphi) = \sqrt{\frac{(l+m)!}{(l-m)!(2l)!}} C'_l \frac{(-1)^{l-m} e^{-im\varphi}}{(1-u^2)^{m/2}} \frac{d^{l-m}}{du^{l-m}} [(1-u^2)^l].$$

The constant  $C'_l$  can be determined as before. In fact, for  $m = 0$  we get exactly the same result as before, so we expect  $C'_l$  to be identical to  $C_l$ . Thus,

$$Y_{l,-m}(u, \varphi) = (-1)^{l+m} \sqrt{\frac{2l+1}{4\pi}} \frac{e^{-im\varphi}}{2^l l!} \sqrt{\frac{(l+m)!}{(l-m)!}} \cdot (1-u^2)^{-m/2} \frac{d^{l-m}}{du^{l-m}} [(1-u^2)^l]$$

Comparison with Equation (12.32) yields

$$Y_{l,-m}(\theta, \varphi) = (-1)^m Y_{l,m}^*(\theta, \varphi), \quad (12.36)$$

and using the definition  $Y_{l,-m}(\theta, \varphi) = P_{l,-m}(\theta) e^{-im\varphi}$  and the first part of Equation (12.33), we obtain

$$P_{l,-m}(\theta) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(\theta). \quad (12.37)$$



The first few spherical harmonics with positive  $m$  are given below. Those with negative  $m$  can be obtained using Equation (12.36).

$$\text{For } l = 0, \quad Y_{00} = \frac{1}{\sqrt{4\pi}}.$$

$$\text{For } l = 1, \quad Y_{10} = \sqrt{\frac{3}{4\pi}} \cos \theta, \quad Y_{11} = -\sqrt{\frac{3}{8\pi}} e^{i\varphi} \sin \theta.$$

$$\text{For } l = 2, \quad Y_{20} = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1), \quad Y_{21} = -\sqrt{\frac{15}{8\pi}} e^{i\varphi} \sin \theta \cos \theta,$$

$$Y_{22} = \sqrt{\frac{15}{32\pi}} e^{2i\varphi} \sin^2 \theta.$$

$$\text{For } l = 3, \quad Y_{30} = \sqrt{\frac{7}{16\pi}} (5 \cos^3 \theta - 3 \cos \theta),$$

$$Y_{31} = -\sqrt{\frac{21}{64\pi}} e^{i\varphi} \sin \theta (5 \cos^2 \theta - 1),$$

$$Y_{32} = \sqrt{\frac{105}{32\pi}} e^{2i\varphi} \sin^2 \theta \cos \theta, \quad Y_{33} = -\sqrt{\frac{35}{64\pi}} e^{3i\varphi} \sin^3 \theta.$$

From Equations (12.13), (12.18), and (12.34) and the fact that  $\alpha = l(l+1)$  for some nonnegative integer  $l$ , we obtain

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) [P_l^m e^{im\varphi}] + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} [P_l^m e^{im\varphi}] + l(l+1) P_l^m e^{im\varphi} = 0,$$

which gives

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{dP_l^m}{d\theta} \right) - \frac{m^2}{\sin^2 \theta} P_l^m + l(l+1) P_l^m = 0.$$

As before, we let  $u = \cos \theta$  to obtain

$$\frac{d}{du} \left[ (1-u^2) \frac{dP_l^m}{du} \right] + \left[ l(l+1) - \frac{m^2}{1-u^2} \right] P_l^m = 0. \quad (12.38)$$

associated Legendre  
differential equation

This is called the **associated Legendre differential equation**. Its solutions, the associated Legendre functions, are given in closed form in Equation (12.33). For  $m = 0$ , Equation (12.38) reduces to the Legendre differential equation whose solutions, again given by Equation (12.33) with  $m = 0$ , are the Legendre polynomials encountered in Chapter 7. When  $m = 0$ , the spherical harmonics become  $\varphi$ -independent. This corresponds to a physical situation in which there is an explicit azimuthal symmetry. In such cases (when it is obvious that the physical property in question does not depend on  $\varphi$ ) a Legendre polynomial, depending only on  $\cos \theta$ , will multiply the radial function.

### 12.4.1 Expansion of Angular Functions

The orthonormality of spherical harmonics can be utilized to expand functions of  $\theta$  and  $\varphi$  in terms of them. The fact that these functions are complete will be discussed in a general way in the context of Sturm-Liouville theory. Assuming completeness for now, we write

$$f(\theta, \varphi) = \begin{cases} \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \varphi) & \text{if } l \text{ is not fixed,} \\ \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \varphi) & \text{if } l \text{ is fixed,} \end{cases} \quad (12.39)$$

where we have included the case where it is known a priori that  $f(\theta, \varphi)$  has a given fixed  $l$  value. To find  $a_{lm}$ , we multiply both sides by  $Y_{lm}^*(\theta, \varphi)$  and integrate over the solid angle. The result, obtained by using the orthonormality relation, is

$$a_{lm} = \iint d\Omega f(\theta, \varphi) Y_{lm}^*(\theta, \varphi), \quad (12.40)$$

where  $d\Omega \equiv \sin\theta d\theta d\varphi$  is the element of solid angle. A useful special case of this formula is

$$a_{l0}^{(f)} = \iint d\Omega f(\theta, \varphi) Y_{l0}^*(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi}} \iint d\Omega f(\theta, \varphi) P_l(\cos\theta), \quad (12.41)$$

where we have introduced an extra superscript to emphasize the relation of the expansion coefficients with the function being expanded. Another useful relation is obtained when we let  $\theta = 0$  in Equation (12.39):

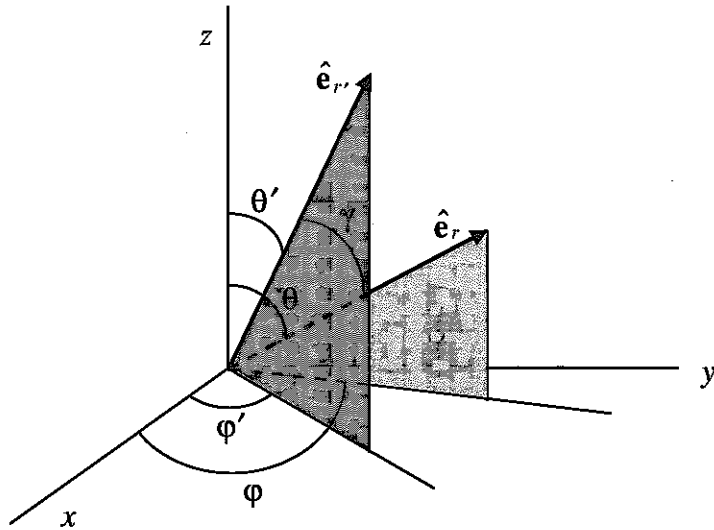
$$f(\theta, \varphi)|_{\theta=0} = \begin{cases} \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \varphi)|_{\theta=0} & \text{if } l \text{ is not fixed,} \\ \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \varphi)|_{\theta=0} & \text{if } l \text{ is fixed.} \end{cases}$$

From Equations (12.33) and (12.34) one can show that

$$Y_{lm}(\theta, \varphi)|_{\theta=0} = \delta_{m0} Y_{l0}(0, \varphi) = \delta_{m0} \sqrt{\frac{2l+1}{4\pi}}.$$

Therefore,

$$f(\theta, \varphi)|_{\theta=0} = \begin{cases} \sum_{l=0}^{\infty} a_{l0}^{(f)} \sqrt{\frac{2l+1}{4\pi}} & \text{if } l \text{ is not fixed,} \\ a_{l0}^{(f)} \sqrt{\frac{2l+1}{4\pi}} & \text{if } l \text{ is fixed.} \end{cases} \quad (12.42)$$



**Figure 12.1** The unit vectors  $\hat{e}_r$  and  $\hat{e}_{r'}$  with their spherical angles and the angle  $\gamma$  between them.

### 12.4.2 Addition Theorem for Spherical Harmonics

An important consequence of the expansion in terms of  $Y_{lm}$  is called the **addition theorem** for spherical harmonics. Consider two unit vectors  $\hat{e}_r$  and  $\hat{e}_{r'}$  making spherical angles  $(\theta, \varphi)$  and  $(\theta', \varphi')$ , respectively, as shown in Figure 12.1. Let  $\gamma$  be the angle between the two vectors. The addition theorem states that

addition theorem for  
spherical harmonics

$$P_l(\cos \gamma) = \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi). \quad (12.43)$$

We shall not give a proof of this theorem here and refer the reader to an elegant proof on page 866 which uses the representation theory of groups. The addition theorem is particularly useful in the expansion of the frequently occurring expression  $1/|\mathbf{r} - \mathbf{r}'|$ . For definiteness we assume  $|\mathbf{r}'| \equiv r' < |\mathbf{r}| \equiv r$ . Then, introducing  $t = r'/r$ , we have

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{(r^2 + r'^2 - 2rr' \cos \gamma)^{1/2}} = \frac{1}{r} (1 + t^2 - 2t \cos \gamma)^{-1/2}.$$

Recalling the generating function for Legendre polynomials from Chapter 7 and using the addition theorem, we get

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{r} \sum_{l=0}^{\infty} t^l P_l(\cos \gamma) = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi)$$

$$= 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{1}{2l+1} \frac{r^l}{r^{l+1}} Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi).$$

It is clear that if  $r < r'$ , we should expand in terms of the ratio  $r/r'$ . It is therefore customary to use  $r_<$  to denote the smaller and  $r_>$  to denote the larger of the two radii  $r$  and  $r'$ . Then the above equation is written as

expansion of  
 $1/|r - r'|$  in spherical  
coordinates

$$\frac{1}{|r - r'|} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{1}{2l+1} \frac{r_<^l}{r_>^{l+1}} Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi). \quad (12.44)$$

This equation is used frequently in the study of Coulomb-like potentials.

## 12.5 Problems

**12.1.** By applying the operator  $[x_j, \mathbf{p}_k]$  to an arbitrary function  $f(\mathbf{r})$ , show that  $[x_j, \mathbf{p}_k] = i\delta_{jk}$ .

**12.2.** Use the defining relation  $\mathbf{L}_i = \epsilon_{ijk} x_j \mathbf{p}_k$  to show that  $x_j \mathbf{p}_k - x_k \mathbf{p}_j = \epsilon_{ijk} \mathbf{L}_i$ . In both of these expressions a sum over the repeated indices is understood.

**12.3.** For the angular momentum operator  $\mathbf{L}_i = \epsilon_{ijk} x_j \mathbf{p}_k$ , show that the commutation relation  $[\mathbf{L}_j, \mathbf{L}_k] = i\epsilon_{jkl} \mathbf{L}_l$  holds.

**12.4.** Evaluate  $\partial f/\partial y$  and  $\partial f/\partial z$  in spherical coordinates and find  $\mathbf{L}_y$  and  $\mathbf{L}_z$  in terms of spherical coordinates.

**12.5.** Obtain an expression for  $\mathbf{L}^2$  in terms of  $\theta$  and  $\varphi$ , and substitute the result in Equation (12.12) to obtain the Laplacian in spherical coordinates.

**12.6.** Show that  $\mathbf{L}^2 = \mathbf{L}_+ \mathbf{L}_- + \mathbf{L}_z^2 - \mathbf{L}_z$  and  $\mathbf{L}^2 = \mathbf{L}_- \mathbf{L}_+ + \mathbf{L}_z^2 + \mathbf{L}_z$ .

**12.7.** Show that  $\mathbf{L}^2$ ,  $\mathbf{L}_x$ ,  $\mathbf{L}_y$ , and  $\mathbf{L}_z$  are hermitian operators in the space of square-integrable functions.

**12.8.** Verify the following commutation relations:

$$[\mathbf{L}^2, \mathbf{L}_{\pm}] = 0, \quad [\mathbf{L}_z, \mathbf{L}_{\pm}] = \pm \mathbf{L}_{\pm}, \quad [\mathbf{L}_+, \mathbf{L}_-] = 2\mathbf{L}_z.$$

**12.9.** Show that  $\mathbf{L}_- |Y_{\alpha\beta}\rangle$  has  $\beta - 1$  as its eigenvalue for  $\mathbf{L}_z$ , and that  $|Y_{\alpha, \beta\pm}\rangle$  cannot be zero.

**12.10.** Show that if the  $|Y_{jm}\rangle$  are normalized to unity, then with proper choice of phase,  $\mathbf{L}_- |Y_{jm}\rangle = \sqrt{j(j+1) - m(m-1)} |Y_{j, m-1}\rangle$ .

**12.11.** Derive Equation (12.35).

12.12. Starting with  $L_x$  and  $L_y$ , derive the following expression for  $L_{\pm}$ :

$$L_{\pm} = e^{\pm i\varphi} \left( \pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right).$$

12.13. Integrate  $dP/d\theta - l \cot \theta P = 0$  to find  $P(\theta)$ .

12.14. Verify the following differential identity:

$$\left( \frac{d}{d\theta} + n \cot \theta \right) f(\theta) = \frac{1}{\sin^n \theta} \frac{d}{d\theta} [\sin^n \theta f(\theta)].$$

12.15. Let  $l = l'$  and  $m = m' = 0$  in Equation (12.30), and substitute for  $Y_{l0}$  from Equation (12.28) to obtain  $A_l = \sqrt{(2l+1)/4\pi}$ .

12.16. Show that

$$L_{+}^k Y_{l,-l}(u, \varphi) = C_l' \frac{(-1)^k e^{-i(l-k)\varphi}}{(1-u^2)^{(l-k)/2}} \frac{d^k}{du^k} [(1-u^2)^l].$$

12.17. Derive the relations  $Y_{l,-m}(\theta, \varphi) = (-1)^m Y_{l,m}^*(\theta, \varphi)$  and

$$P_l^{-m}(\theta) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(\theta).$$

12.18. Show that  $\sum_{m=-l}^l |Y_{lm}(\theta, \varphi)|^2 = (2l+1)/(4\pi)$ . Verify this explicitly for  $l = 1$  and  $l = 2$ .

12.19. Show that the addition theorem for spherical harmonics can be written as

$$P_l(\cos \gamma) = P_l(\cos \theta) P_l(\cos \theta') + 2 \sum_{m=1}^l \frac{(l-m)!}{(l+m)!} P_l^m(\cos \theta) P_l^m(\cos \theta') \cos[m(\varphi - \varphi')].$$

## Additional Reading

1. Morse, P. and Feshbach, M. *Methods of Theoretical Physics*, McGraw-Hill, 1953. A two-volume classic including a long discussion of the separation of variables in many (sometimes exotic) coordinate systems.
2. The angular momentum eigenvalues and eigenfunctions are discussed in most books on quantum mechanics. See, e.g., Messiah, A. *Quantum Mechanics*, volume II, Wiley, 1966.

# Second-Order Linear Differential Equations

The discussion of Chapter 12 has clearly singled out ODEs, especially those of second order, as objects requiring special attention because most common PDEs of mathematical physics can be separated into ODEs (of second order). This is really an oversimplification of the situation. Many PDEs of physics, both at the fundamental theoretical level (as in the general theory of relativity) and from a practical standpoint (weather forecast) are nonlinear, and the method of the separation of variables does not work. Since no general analytic solutions for such nonlinear systems have been found, we shall confine ourselves to the linear systems, especially those that admit a separated solution.

With the exception of the infinite power series, no systematic method of solving DEs existed during the first half of the nineteenth century. The majority of solutions were completely ad hoc and obtained by trial and error, causing frustration and anxiety among mathematicians. It was to overcome this frustration that Sophus Lie, motivated by the newly developed concept of group, took up the systematic study of DEs in the second half of the nineteenth century. This study not only gave a handle on the disarrayed area of DEs, but also gave birth to one of the most beautiful and fundamental branches of mathematical physics, Lie group theory. We shall come back to a thorough treatment of this theory in Parts VII and VIII.

Our main task in this chapter is to study the second-order linear differential equations (SOLDEs). However, to understand SOLDEs, we need some basic understanding of differential equations in general. The next section outlines some essential properties of general DEs. Section 2 is a very brief introduction to first-order DEs, and the remainder of the chapter deals with SOLDEs.

### 13.1 General Properties of ODEs

The most general ODE can be expressed as

$$F\left(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \dots, \frac{d^ny}{dx^n}\right) = 0, \tag{13.1}$$

in which  $F : \mathbb{R}^{n+2} \rightarrow \mathbb{R}$  is a real-valued function of  $n + 2$  real variables. When  $F$  depends explicitly and nontrivially on  $d^ny/dx^n$ , Equation (13.1) is called an *n*th-order ODE. An ODE is said to be **linear** if the part of the function  $F$  that includes  $y$  and all its derivatives is linear in  $y$ . The most general *n*th-order *linear* ODE is

$$p_0(x)y + p_1(x)\frac{dy}{dx} + \dots + p_n(x)\frac{d^ny}{dx^n} = q(x) \quad \text{for } p_n(x) \neq 0, \tag{13.2}$$

where  $\{p_i\}_{i=0}^n$  and  $q$  are functions of the independent variable  $x$ . Equation (13.2) is said to be **homogeneous** if  $q = 0$ ; otherwise, it is said to be **inhomogeneous** and  $q(x)$  is called the *inhomogeneous term*. It is customary, and convenient, to define a linear differential operator  $\mathbf{L}$  by<sup>1</sup>

$$\mathbf{L} \equiv p_0(x) + p_1(x)\frac{d}{dx} + \dots + p_n(x)\frac{d^n}{dx^n}, \quad p_n(x) \neq 0, \tag{13.3}$$

and write Equation (13.2) as

$$\mathbf{L}[y] = q(x). \tag{13.4}$$

A **solution** of Equation (13.2) or (13.4) is a single-variable function  $f : \mathbb{R} \rightarrow \mathbb{R}$  such that  $F(x, f(x), f'(x), \dots, f^{(n)}(x)) = 0$ , or  $\mathbf{L}[f] = q(x)$ , for all  $x$  in the domain of definition of  $f$ . The solution of a differential equation may not exist if we put too many restrictions on it. For instance, if we demand that  $f : \mathbb{R} \rightarrow \mathbb{R}$  be differentiable too many times, we may not be able to find a solution, as the following example shows.

**13.1.1. Example.** The most general solution of  $dy/dx = |x|$  that vanishes at  $x = 0$  is

$$f(x) = \begin{cases} \frac{1}{2}x^2 & \text{if } x \geq 0, \\ -\frac{1}{2}x^2 & \text{if } x \leq 0. \end{cases}$$

This function is continuous and has first derivative  $f'(x) = |x|$ , which is also continuous at  $x = 0$ . However, if we demand that its second derivative also be continuous at  $x = 0$ , we cannot find a solution, because

$$f''(x) = \begin{cases} +1 & \text{if } x > 0, \\ -1 & \text{if } x < 0. \end{cases}$$

<sup>1</sup>Do not confuse this linear differential operator with the angular momentum (vector) operator  $\mathbf{L}$ .

If we want  $f'''(x)$  to exist at  $x = 0$ , then we have to expand the notion of a function to include distributions, or generalized functions. ■

Overrestricting a solution for a differential equation results in its absence, but underrestricting it allows multiple solutions. To strike a balance between these two extremes, we agree to make a solution as many times differentiable as plausible and to satisfy certain **initial conditions**. For an  $n$ th-order DE such initial conditions are commonly equivalent (but not restricted) to a specification of the function and of its first  $n - 1$  derivatives. This sort of specification is made feasible by the following theorem.

implicit function  
theorem

**13.1.2. Theorem.** (implicit function theorem) *Let  $G : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ , given by  $G(x_1, x_2, \dots, x_{n+1}) \in \mathbb{R}$ , have continuous partial derivatives up to the  $k$ th order in some neighborhood of a point  $P_0 = (r_1, r_2, \dots, r_{n+1})$  in  $\mathbb{R}^{n+1}$ . Let  $(\partial G / \partial x_{n+1})|_{P_0} \neq 0$ . Then there exists a unique function  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  that is continuously differentiable  $k$  times at (some smaller) neighborhood of  $P_0$  such that  $x_{n+1} = F(x_1, x_2, \dots, x_n)$  for all points  $P = (x_1, x_2, \dots, x_{n+1})$  in a neighborhood of  $P_0$  and*

$$G(x_1, x_2, \dots, x_n, F(x_1, x_2, \dots, x_n)) = 0.$$

Theorem 13.1.2 simply asserts that under certain (mild) conditions we can “solve” for one of the independent variables in  $G(x_1, x_2, \dots, x_{n+1}) = 0$  in terms of the others. A proof of this theorem is usually given in advanced calculus books.

Application of this theorem to Equation (13.1) leads to

$$\frac{d^n y}{dx^n} = F\left(x, y, \frac{dy}{dx}, \frac{d^2 y}{dx^2}, \dots, \frac{d^{n-1} y}{dx^{n-1}}\right),$$

provided that  $G$  satisfies the conditions of the theorem. If we know the solution  $y = f(x)$  and its derivatives up to order  $n - 1$ , we can evaluate its  $n$ th derivative using this equation. In addition, we can calculate the derivatives of all orders (assuming they exist) by differentiating this equation. This allows us to expand the solution in a Taylor series. Thus—for solutions that have derivatives of all orders—knowledge of the value of a solution and its first  $n - 1$  derivatives at a point  $x_0$  determines that solution at a neighboring point  $x$ .

We shall not study the general ODE of Equation (13.1) or even its simpler linear version (13.2). We will only briefly study ODEs of the first order in the next section, and then concentrate on linear ODEs of the second order for the rest of this chapter.

## 13.2 Existence and Uniqueness for First-Order DEs

A general first-order DE (FODE) is of the form  $G(x, y, y') = 0$ . We can find  $y'$  (the derivative of  $y$ ) in terms of a function of  $x$  and  $y$  if the function  $G(x_1, x_2, x_3)$



the most general FODE in normal form

is differentiable with respect to its third argument and  $\partial G/\partial x_3 \neq 0$ . In that case we have

$$y' \equiv \frac{dy}{dx} = F(x, y), \tag{13.5}$$

which is said to be a **normal FODE**. If  $F(x, y)$  is a *linear* function of  $y$ , then Equation (13.5) becomes a first-order linear DE (FOLDE), which can generally be written as

$$p_1(x) \frac{dy}{dx} + p_0(x)y = q(x). \tag{13.6}$$

It can be shown that the general FOLDE has an explicit solution: (see [Hass 99])

explicit solution to a general first-order linear differential equation

**13.2.1. Theorem.** *Any first order linear DE of the form  $p_1(x)y' + p_0(x)y = q(x)$ , in which  $p_0, p_1$ , and  $q$  are continuous functions in some interval  $(a, b)$ , has a general solution*

$$y = f(x) = \frac{1}{\mu(x)p_1(x)} \left[ C + \int_{x_1}^x \mu(t)q(t) dt \right], \tag{13.7}$$

where  $C$  is an arbitrary constant and

$$\mu(x) = \frac{1}{p_1(x)} \exp \left[ \int_{x_0}^x \frac{p_0(t)}{p_1(t)} dt \right], \tag{13.8}$$

where  $x_0$  and  $x_1$  are arbitrary points in the interval  $(a, b)$ .

No such explicit solution exists for nonlinear first-order DEs. Nevertheless, it is reassuring to know that a solution of such a DE always exists and under some mild conditions, this solution is unique. We summarize some of the ideas involved in the proof of the existence and uniqueness of the solutions to FODEs. (For proofs, see the excellent book by Birkhoff and Rota [Birk 78].) We first state an existence theorem due to Peano:

Peano existence theorem

**13.2.2. Theorem.** (Peano existence theorem) *If the function  $F(x, y)$  is continuous for the points on and within the rectangle defined by  $|y - c| \leq K$  and  $|x - a| \leq N$ , and if  $|F(x, y)| \leq M$  there, then the differential equation  $y' = F(x, y)$  has at least one solution,  $y = f(x)$ , defined for  $|x - a| \leq \min(N, K/M)$  and satisfying the initial condition  $f(a) = c$ .*

This theorem guarantees only the existence of solutions. To ensure uniqueness, the function  $F$  needs to have some additional properties. An important property is stated in the following definition.

Lipschitz condition

**13.2.3. Definition.** *A function  $F(x, y)$  satisfies a **Lipschitz condition** in a domain  $D \subset \mathbb{R}^2$  if for some finite constant  $L$  (Lipschitz constant), it satisfies the inequality*

$$|F(x, y_1) - F(x, y_2)| \leq L|y_1 - y_2|$$

for all points  $(x, y_1)$  and  $(x, y_2)$  in  $D$ .

uniqueness theorem

**13.2.4. Theorem.** (uniqueness theorem) *Let  $f(x)$  and  $g(x)$  be any two solutions of the FODE  $y' = F(x, y)$  in a domain  $D$ , where  $F$  satisfies a Lipschitz condition with Lipschitz constant  $L$ . Then*

$$|f(x) - g(x)| \leq e^{L|x-a|} |f(a) - g(a)|.$$

*In particular, the FODE has at most one solution curve passing through the point  $(a, c) \in D$ .*

The final conclusion of this theorem is an easy consequence of the assumed differentiability of  $F$  and the requirement  $f(a) = g(a) = c$ . The theorem says that if there is a solution  $y = f(x)$  to the DE  $y' = F(x, y)$  satisfying  $f(a) = c$ , then it is *the* solution.

The requirements of the Peano existence theorem are too broad to yield solutions that have some nice properties. For instance, the interval of definition of the solutions may depend on their initial values. The following example illustrates this point.

**13.2.5. Example.** Consider the DE  $dy/dx = e^y$ . The general solution of this DE can be obtained by direct integration:

$$e^{-y} dy = dx \Rightarrow -e^{-y} = x + C.$$

If  $y = b$  when  $x = 0$ , then  $C = -e^{-b}$ , and

$$e^{-y} = -x + e^{-b} \Rightarrow y = -\ln(e^{-b} - x).$$

Thus, the solution is defined for  $-\infty < x < e^{-b}$ , i.e., the interval of definition of a solution changes with its initial value. ■

To avoid situations illustrated in the example above, one demands not just the continuity of  $F$ —as does the Peano existence theorem—but a Lipschitz condition for it. Then one ensures not only the existence, but also the uniqueness:

local existence and uniqueness theorem

**13.2.6. Theorem.** (local existence and uniqueness theorem) *Suppose that the function  $F(x, y)$  is defined and continuous in the rectangle  $|y - c| \leq K, |x - a| \leq N$  and satisfies a Lipschitz condition there. Let  $M = \max |F(x, y)|$  in this rectangle. Then the differential equation  $y' = F(x, y)$  has a unique solution  $y = f(x)$  satisfying  $f(a) = c$  and defined on the interval  $|x - a| \leq \min(N, K/M)$ .*

### 13.3 General Properties of SOLDEs

The most general SOLDE is

$$p_2(x) \frac{d^2 y}{dx^2} + p_1(x) \frac{dy}{dx} + p_0(x)y = p_3(x). \tag{13.9}$$

normal form of a SOLDE Dividing by  $p_2(x)$  and writing  $p$  for  $p_1/p_2$ ,  $q$  for  $p_0/p_2$ , and  $r$  for  $p_3/p_2$  reduces this to the **normal form**

$$\frac{d^2y}{dx^2} + p(x)\frac{dy}{dx} + q(x)y = r(x). \tag{13.10}$$

singular points of a SOLDE Equation (13.10) is equivalent to (13.9) if  $p_2(x) \neq 0$ . The points at which  $p_2(x)$  vanishes are called the **singular points** of the differential equation.

There is a crucial difference between the singular points of linear differential equations and those of nonlinear differential equations. For a nonlinear differential equation such as  $(x^2 - y)y' = x^2 + y^2$ , the curve  $y = x^2$  is the collection of singular points. This makes it impossible to construct solutions  $y = f(x)$  that are defined on an interval  $I = [a, b]$  of the  $x$ -axis because for any  $x \in I$ , there is a  $y$  for which the differential equation is undefined. *Linear* differential equations do not have this problem, because the coefficients of the derivatives are functions of  $x$  only. Therefore, all the singular “curves” are vertical. Thus, we have the following definition.

regular SOLDE **13.3.1. Definition.** *The normal form of a SOLDE, Equation (13.10), is **regular** on an interval  $[a, b]$  of the  $x$ -axis if  $p(x)$ ,  $q(x)$ , and  $r(x)$  are continuous on  $[a, b]$ . A solution of a normal SOLDE is a twice-differentiable function  $y = f(x)$  that satisfies the SOLDE at every point of  $[a, b]$ .*

It is clear that any function that satisfies Equation (13.10)—or Equation (13.9)—must necessarily be twice differentiable, and that is all that is demanded of the solutions. Any higher-order differentiability requirement may be too restrictive, as was pointed out in Example 13.1.1. Most solutions to a normal SOLDE, however, automatically have derivatives of order higher than two.

We write Equation (13.9) in the operator form as

$$\mathbf{L}[y] = p_3, \quad \text{where} \quad \mathbf{L} \equiv p_2 \frac{d^2}{dx^2} + p_1 \frac{d}{dx} + p_0. \tag{13.11}$$

It is clear that  $\mathbf{L}$  is a *linear* operator because  $d/dx$  is linear, as are all powers of it. Thus, for constants  $\alpha$  and  $\beta$ ,  $\mathbf{L}[\alpha y_1 + \beta y_2] = \alpha \mathbf{L}[y_1] + \beta \mathbf{L}[y_2]$ . In particular, if  $y_1$  and  $y_2$  are two solutions of Equation (13.11), then  $\mathbf{L}[y_1 - y_2] = 0$ . That is, the difference between any two solutions of a SOLDE is a solution of the **homogeneous equation** obtained by setting  $p_3 = 0$ .<sup>2</sup>

An immediate consequence of the linearity of  $\mathbf{L}$  is the following:

**13.3.2. Lemma.** *If  $\mathbf{L}[u] = r(x)$ ,  $\mathbf{L}[v] = s(x)$ ,  $\alpha$  and  $\beta$  are constants, and  $w = \alpha u + \beta v$ , then  $\mathbf{L}[w] = \alpha r(x) + \beta s(x)$ .*

superposition principle The proof of this lemma is trivial, but the result describes the fundamental property of linear operators: When  $r = s = 0$ , that is, in dealing with homogeneous

<sup>2</sup>This conclusion is, of course, not limited to the SOLDE; it holds for all linear DEs.

equations, the lemma says that any linear combination of solutions of the homogeneous SOLDE (HSOLDE) is also a solution. This is called the **superposition principle**.

Based on physical intuition, we expect to be able to predict the behavior of a physical system if we know the differential equation obeyed by that system, and, equally importantly, the initial data. A prediction is not a prediction unless it is unique.<sup>3</sup> This expectation for linear equations is borne out in the language of mathematics in the form of an existence theorem and a uniqueness theorem. We consider the latter next. But first, we need a lemma.

**13.3.3. Lemma.** *The only solution  $g(x)$  of the homogeneous equation  $y'' + py' + qy = 0$  defined on the interval  $[a, b]$  that satisfies  $g(a) = 0 = g'(a)$  is the trivial solution  $g = 0$ .*

*Proof.* Introduce the nonnegative function  $u(x) \equiv [g(x)]^2 + [g'(x)]^2$  and differentiate it to get

$$\begin{aligned} u'(x) &= 2g'g + 2g'g'' = 2g'(g + g'') = 2g'(g - pg' - qg) \\ &= -2p(g')^2 + 2(1 - q)gg'. \end{aligned}$$

Since  $(g \pm g')^2 \geq 0$ , it follows that  $2|gg'| \leq g^2 + g'^2$ . Thus,

$$\begin{aligned} 2(1 - q)gg' &\leq 2|(1 - q)gg'| = 2|(1 - q)| |gg'| \\ &\leq |(1 - q)|(g^2 + g'^2) \leq (1 + |q|)(g^2 + g'^2), \end{aligned}$$

and therefore,

$$\begin{aligned} u'(x) &\leq |u'(x)| = |-2pg'^2 + 2(1 - q)gg'| \\ &\leq 2|p|g'^2 + (1 + |q|)(g^2 + g'^2) \\ &= [1 + |q(x)|]g^2 + [1 + |q(x)| + 2|p(x)|]g'^2. \end{aligned}$$

Now let  $K = 1 + \max[|q(x)| + 2|p(x)|]$ , where the maximum is taken over  $[a, b]$ . Then we obtain

$$u'(x) \leq K(g^2 + g'^2) = Ku(x) \quad \forall x \in [a, b].$$

Using the result of Problem 13.1 yields  $u(x) \leq u(a)e^{K(x-a)}$  for all  $x \in [a, b]$ . This equation, plus  $u(a) = 0$ , as well as the fact that  $u(x) \geq 0$  imply that  $u(x) = g^2(x) + g'^2(x) = 0$ . It follows that  $g(x) = 0 = g'(x)$  for all  $x \in [a, b]$ .  $\square$

uniqueness of  
solutions to SOLDE

**13.3.4. Theorem.** (Uniqueness theorem) *If  $p$  and  $q$  are continuous on  $[a, b]$ , then*

<sup>3</sup>Physical intuition also tells us that if the initial conditions are changed by an infinitesimal amount, then the solutions will be changed infinitesimally. Thus, the solutions of linear differential equations are said to be continuous functions of the initial conditions. Nonlinear differential equations can have completely different solutions for two initial conditions that are infinitesimally close. Since initial conditions cannot be specified with mathematical precision in practice, nonlinear differential equations lead to unpredictable solutions, or chaos. This subject has received much attention in recent years. For an elementary discussion of chaos see [Hass 99, Chapter 15].

at most one solution  $y = f(x)$  of Equation (13.10) can satisfy the initial conditions  $f(a) = c_1$  and  $f'(a) = c_2$ , where  $c_1$  and  $c_2$  are arbitrary constants.

*Proof.* Let  $f_1$  and  $f_2$  be two solutions satisfying the given initial conditions. Then their difference,  $g \equiv f_1 - f_2$ , satisfies the homogeneous equation [with  $r(x) = 0$ ]. The initial condition that  $g(x)$  satisfies is clearly  $g(a) = 0 = g'(a)$ . By Lemma 13.3.3,  $g = 0$  or  $f_1 = f_2$ .  $\square$

Theorem 13.3.4 can be applied to any *homogeneous SOLDE* to find the latter's most general solution. In particular, let  $f_1(x)$  and  $f_2(x)$  be any two solutions of

$$y'' + p(x)y' + q(x)y = 0 \quad (13.12)$$

defined on the interval  $[a, b]$ . Assume that the two vectors  $\mathbf{v}_1 = (f_1(a), f_1'(a))$  and  $\mathbf{v}_2 = (f_2(a), f_2'(a))$  in  $\mathbb{R}^2$  are linearly independent.<sup>4</sup> Let  $g(x)$  be another solution. The vector  $(g(a), g'(a))$  can be written as a linear combination of  $\mathbf{v}_1$  and  $\mathbf{v}_2$ , giving the two equations

$$\begin{aligned} g(a) &= c_1 f_1(a) + c_2 f_2(a), \\ g'(a) &= c_1 f_1'(a) + c_2 f_2'(a). \end{aligned}$$

Now consider the function  $u(x) \equiv g(x) - c_1 f_1(x) - c_2 f_2(x)$ , which satisfies Equation (13.12) and the initial conditions  $u(a) = u'(a) = 0$ . By Lemma 13.3.3, we must have  $u(x) = 0$  or  $g(x) = c_1 f_1(x) + c_2 f_2(x)$ . We have proved the following:

**13.3.5. Theorem.** *Let  $f_1$  and  $f_2$  be two solutions of the HSOLDE*

$$y'' + py' + qy = 0,$$

where  $p$  and  $q$  are continuous functions defined on the interval  $[a, b]$ . If

$$(f_1(a), f_1'(a)) \quad \text{and} \quad (f_2(a), f_2'(a))$$

are linearly independent vectors in  $\mathbb{R}^2$ , then every solution  $g(x)$  of this HSOLDE is equal to some linear combination  $g(x) = c_1 f_1(x) + c_2 f_2(x)$  of  $f_1$  and  $f_2$  with constant coefficients  $c_1$  and  $c_2$ .

## 13.4 The Wronskian

The two solutions  $f_1(x)$  and  $f_2(x)$  in Theorem 13.3.5 have the property that any other solution  $g(x)$  can be expressed as a linear combination of them. We call  $f_1$  and  $f_2$  a **basis of solutions** of the HSOLDE. To form a basis of solutions,  $f_1$

<sup>4</sup>If they are not, then one must choose a different initial point for the interval.

and  $f_2$  must be linearly independent. The linear dependence or independence of a number of functions  $\{f_i\}_{i=1}^n : [a, b] \rightarrow \mathbb{R}$  is a concept that must hold for all  $x \in [a, b]$ . Thus, if  $\{\alpha_i\}_{i=1}^n \in \mathbb{R}$  can be found such that

$$\alpha_1 f_1(x_0) + \alpha_2 f_2(x_0) + \cdots + \alpha_n f_n(x_0) = 0$$

for some  $x_0 \in [a, b]$ , it does not mean that the  $f$ 's are linearly dependent. Linear dependence requires that the equality hold for all  $x \in [a, b]$ . In fact, we must write

$$\alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_n f_n = \mathbf{0},$$

where  $\mathbf{0}$  is the zero function.

Wronskian defined

**13.4.1. Definition.** *The Wronskian of any two differentiable functions  $f_1(x)$  and  $f_2(x)$  is*

$$W(f_1, f_2; x) = f_1(x)f_2'(x) - f_2(x)f_1'(x) = \det \begin{pmatrix} f_1(x) & f_1'(x) \\ f_2(x) & f_2'(x) \end{pmatrix}.$$

**13.4.2. Proposition.** *The Wronskian of any two solutions of Equation (13.12) satisfies*

$$W(f_1, f_2; x) = W(f_1, f_2; c)e^{-\int_c^x p(t) dt},$$

where  $c$  is any number in the interval  $[a, b]$ .

*Proof.* Differentiating both sides of the definition of Wronskian and substituting from Equation (13.12) yields a FOLDE for  $W(f_1, f_2; x)$ , which can be easily solved. The details are left as a problem.  $\square$

An important consequence of Proposition 13.4.2 is that the Wronskian of any two solutions of Equation (13.12) does not change sign in  $[a, b]$ . In particular, if the Wronskian vanishes at one point in  $[a, b]$ , it vanishes at all points in  $[a, b]$ .

The real importance of the Wronskian is contained in the following theorem.

**13.4.3. Theorem.** *Two differentiable functions  $f_1$  and  $f_2$ , which are nonzero in the interval  $[a, b]$ , are linearly dependent if and only if their Wronskian vanishes.*

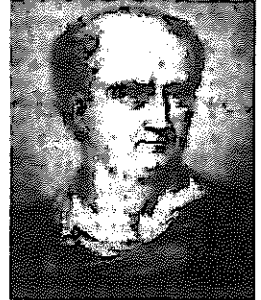
*Proof.* If  $f_1$  and  $f_2$  are linearly dependent, then one is a multiple of the other, and the Wronskian is readily seen to vanish. Conversely, assume that the Wronskian is zero. Then

$$f_1(x)f_2'(x) - f_2(x)f_1'(x) = 0 \Rightarrow f_1 df_2 = f_2 df_1 \Rightarrow f_2 = C f_1$$

and the two functions are linearly dependent.  $\square$

**Josef Hoëné de Wronski** (1778–1853) was born Josef Hoëné, but he adopted the name Wronski around 1810 just after he married. He had moved to France and become a French citizen in 1800 and moved to Paris in 1810, the same year he published his first memoir on the foundations of mathematics, which received less than favorable reviews from Lacroix and Lagrange. His other interests included the design of caterpillar vehicles to compete with the railways. However, they were never manufactured.

Wronski was interested mainly in applying philosophy to mathematics, the philosophy taking precedence over rigorous mathematical proofs. He criticised Lagrange's use of infinite series and introduced his own ideas for series expansions of a function. The coefficients in this series are determinants now known as **Wronskians** [so named by Thomas Muir (1844–1934), a Glasgow High School science master who became an authority on determinants by devoting most of his life to writing a five-volume treatise on the history of determinants].



For many years Wronski's work was dismissed as rubbish. However, a closer examination of the work in more recent times shows that although some is wrong and he has an incredibly high opinion of himself and his ideas, there are also some mathematical insights of great depth and brilliance hidden within the papers.

**13.4.4. Example.** Let  $f_1(x) = x$  and  $f_2(x) = |x|$  for  $x \in [-1, 1]$ . These two functions are linearly independent in the given interval, because  $\alpha_1 x + \alpha_2 |x| = 0$  for all  $x$  if and only if  $\alpha_1 = \alpha_2 = 0$ . The Wronskian, on the other hand, vanishes for all  $x \in [-1, +1]$ :

$$\begin{aligned} W(f_1, f_2; x) &= x \frac{d|x|}{dx} - |x| \frac{dx}{dx} = x \frac{d|x|}{dx} - |x| \\ &= x \frac{d}{dx} \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x \leq 0 \end{cases} - \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x \leq 0 \end{cases} \\ &= \begin{cases} x - x = 0 & \text{if } x > 0 \\ -x - (-x) = 0 & \text{if } x < 0. \end{cases} \end{aligned}$$

Thus, it is possible for two functions to have a vanishing Wronskian without being linearly dependent. However, as we showed in the proof of the theorem above, if the functions are differentiable in their interval of definition, then they are linearly dependent if their Wronskian vanishes. ■

**13.4.5. Example.** The Wronskian can be generalized to  $n$  functions. The Wronskian of the functions  $f_1, f_2, \dots, f_n$  is

$$W(f_1, f_2, \dots, f_n; x) = \det \begin{pmatrix} f_1(x) & f_1'(x) & \dots & f_1^{(n-1)}(x) \\ f_2(x) & f_2'(x) & \dots & f_2^{(n-1)}(x) \\ \vdots & \vdots & \dots & \vdots \\ f_n(x) & f_n'(x) & \dots & f_n^{(n-1)}(x) \end{pmatrix}.$$

If the functions are linearly dependent, then  $W(f_1, f_2, \dots, f_n; x) = 0$ .

For instance, it is clear that  $e^x$ ,  $e^{-x}$ , and  $\sinh x$  are linearly dependent. Thus, we expect

$$W(e^x, e^{-x}, \sinh x; x) = \det \begin{pmatrix} e^x & e^x & e^x \\ e^{-x} & -e^{-x} & e^{-x} \\ \sinh x & \cosh x & \sinh x \end{pmatrix}$$

to vanish, as is easily seen (the first and last columns are the same). ■

### 13.4.1 A Second Solution to the HSOLDE

If we know one solution to Equation (13.12), say  $f_1$ , then by differentiating both sides of

$$f_1(x)f_2'(x) - f_2(x)f_1'(x) = W(x) = W(c)e^{-\int_c^x p(t) dt},$$

dividing the result by  $f_1^2$ , and noting that the LHS will be the derivative of  $f_2/f_1$ , we can solve for  $f_2$  in terms of  $f_1$ . The result is

$$f_2(x) = f_1(x) \left\{ C + K \int_{\alpha}^x \frac{1}{f_1^2(s)} \exp \left[ - \int_c^s p(t) dt \right] ds \right\}, \quad (13.13)$$

where  $K \equiv W(c)$  is another arbitrary (nonzero) constant; we do not have to know  $W(x)$  (this would require knowledge of  $f_2$ , which we are trying to calculate!) to obtain  $W(c)$ . In fact, the reader is urged to check directly that  $f_2(x)$  satisfies the DE of (13.12) for arbitrary  $C$  and  $K$ . Whenever possible—and convenient—it is customary to set  $C = 0$ , because its presence simply gives a term that is proportional to the known solution  $f_1(x)$ .

**13.4.6. Example.** (a) A solution to the SOLDE  $y'' - k^2y = 0$  is  $e^{kx}$ . To find a second solution, we let  $C = 0$  and  $K = 1$  in Equation (13.13). Since  $p(x) = 0$ , we have

$$f_2(x) = e^{kx} \left( 0 + \int_{\alpha}^x \frac{ds}{e^{2ks}} \right) = -\frac{1}{2k} e^{-kx} + \frac{e^{-2k\alpha}}{2k} e^{kx},$$

which, ignoring the second term (which is proportional to the first solution), leads directly to the choice of  $e^{-kx}$  as a second solution.

(b) The differential equation  $y'' + k^2y = 0$  has  $\sin kx$  as a solution. With  $C = 0$ ,  $\alpha = \pi/(2k)$ , and  $K = 1$ , we get

$$f_2(x) = \sin kx \left( 0 + \int_{\pi/2k}^x \frac{ds}{\sin^2 ks} \right) = -\sin kx \cot ks \Big|_{\pi/2k}^x = -\cos kx.$$

(c) For the solutions in part (a),

$$W(x) = \det \begin{pmatrix} e^{kx} & ke^{kx} \\ e^{-kx} & -ke^{-kx} \end{pmatrix} = -2k,$$

and for those in part (b),

$$W(x) = \det \begin{pmatrix} \sin kx & k \cos kx \\ \cos kx & -k \sin kx \end{pmatrix} = -k.$$



Both Wronskians are constant. In general, the Wronskian of any two linearly independent solutions of  $y'' + q(x)y = 0$  is constant. ■

Most special functions used in mathematical physics are solutions of SOLDEs. The behavior of these functions at certain special points is determined by the physics of the particular problem. In most situations physical expectation leads to a preference for one particular solution over the other. For example, although there are two linearly independent solutions to the Legendre DE

$$\frac{d}{dx} \left[ (1-x^2) \frac{dy}{dx} \right] + n(n+1)y = 0,$$

the solution that is most frequently encountered is the Legendre polynomial  $P_n(x)$  discussed in Chapter 7. The other solution can be obtained by solving the Legendre equation or by using Equation (13.13), as done in the following example.

**13.4.7. Example.** The Legendre equation can be reexpressed as

$$\frac{d^2 y}{dx^2} - \frac{2x}{1-x^2} \frac{dy}{dx} + \frac{n(n+1)}{1-x^2} y = 0.$$

This is an HSOLDE with

$$p(x) = -\frac{2x}{1-x^2} \quad \text{and} \quad q(x) = \frac{n(n+1)}{1-x^2}.$$

One solution of this HSOLDE is the well-known Legendre polynomial  $P_n(x)$ . Using this as our input and employing Equation (13.13), we can generate another set of solutions.

Let  $Q_n(x)$  stand for the linearly independent “partner” of  $P_n(x)$ . Then, setting  $C = 0 = c$  in Equation (13.13) yields

$$\begin{aligned} Q_n(x) &= K P_n(x) \int_{\alpha}^x \frac{1}{P_n^2(s)} \exp \left[ \int_0^s \frac{2t}{1-t^2} dt \right] ds \\ &= K P_n(x) \int_{\alpha}^x \frac{1}{P_n^2(s)} \left[ \frac{1}{1-s^2} \right] ds = A_n P_n(x) \int_{\alpha}^x \frac{ds}{(1-s^2)P_n^2(s)}, \end{aligned}$$

where  $A_n$  is an arbitrary constant determined by standardization, and  $\alpha$  is an arbitrary point in the interval  $[-1, +1]$ . For instance, for  $n = 0$ , we have  $P_0 = 1$ , and we obtain

$$Q_0(x) = A_0 \int_{\alpha}^x \frac{ds}{1-s^2} = A_0 \left[ \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right| - \frac{1}{2} \ln \left| \frac{1+\alpha}{1-\alpha} \right| \right].$$

The standard form of  $Q_0(x)$  is obtained by setting  $A_0 = 1$  and  $\alpha = 0$ :

$$Q_0(x) = \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right| \quad \text{for } |x| < 1.$$

Similarly, since  $P_1(x) = x$ ,

$$Q_1(x) = A_1 x \int_{\alpha}^x \frac{ds}{s^2(1-s^2)} = Ax + Bx \ln \left| \frac{1+x}{1-x} \right| + C \quad \text{for } |x| < 1.$$

Here standardization is  $A = 0$ ,  $B = \frac{1}{2}$ , and  $C = -1$ . Thus,

$$Q_1(x) = \frac{1}{2} x \ln \left| \frac{1+x}{1-x} \right| - 1. \quad \blacksquare$$

### 13.4.2 The General Solution to an ISOLDE

Inhomogeneous SOLDEs (ISOLDEs) can be most elegantly discussed in terms of Green's functions, the subject of Chapter 20, which automatically incorporate the boundary conditions. However, the most general solution of an ISOLDE, with no boundary specification, can be discussed at this point.

Let  $g(x)$  be a particular solution of

$$L[y] = y'' + py' + qy = r(x) \tag{13.14}$$

and let  $h(x)$  be any other solution of this equation. Then  $h(x) - g(x)$  satisfies Equation (13.12) and can be written as a linear combination of a basis of solutions  $f_1(x)$  and  $f_2(x)$ , leading to the following equation:

$$h(x) = c_1 f_1(x) + c_2 f_2(x) + g(x). \tag{13.15}$$

Thus, if we have a *particular* solution of the ISOLDE of Equation (13.14) and two basis solutions of the HSOLDE, then the *most general* solution of (13.14) can be expressed as the sum of a linear combination of the two basis solutions and the particular solution.

We know how to find a second solution to the HSOLDE once we know one solution. We now show that knowing one such solution will also allow us to find a particular solution to the ISOLDE. The method we use is called the **variation of constants**. This method can also be used to find a second solution to the HSOLDE.

method of variation  
of constants

Let  $f_1$  and  $f_2$  be the two (known) solutions of the HSOLDE and  $g(x)$  the sought-after solution to Equation (13.14). Write  $g$  as  $g(x) = f_1(x)v(x)$  and substitute it in (13.14) to get a SOLDE for  $v(x)$ :

$$v'' + \left( p + \frac{2f_1'}{f_1} \right) v' = \frac{r}{f_1}.$$

This is a *first* order linear DE in  $v'$ , which has a solution of the form

$$v' = \frac{W(x)}{f_1^2(x)} \left[ C + \int_a^x \frac{f_1(t)r(t)}{W(t)} dt \right],$$

where  $W(x)$  is the (known) Wronskian of Equation (13.14). Substituting

$$\frac{W(x)}{f_1^2(x)} = \frac{f_1(x)f_2'(x) - f_2(x)f_1'(x)}{f_1^2(x)} = \frac{d}{dx} \left( \frac{f_2}{f_1} \right)$$

in the above expression for  $v'$  and setting  $C = 0$  (we are interested in a *particular* solution), we get

$$\begin{aligned} \frac{dv}{dx} &= \frac{d}{dx} \left( \frac{f_2}{f_1} \right) \int_a^x \frac{f_1(t)r(t)}{W(t)} dt \\ &= \frac{d}{dx} \left[ \frac{f_2(x)}{f_1(x)} \int_a^x \frac{f_1(t)r(t)}{W(t)} dt \right] - \underbrace{\frac{f_2(x)}{f_1(x)} \frac{d}{dx} \int_a^x \frac{f_1(t)r(t)}{W(t)} dt}_{=f_1(x)r(x)/W(x)} \end{aligned}$$

and

$$v(x) = \frac{f_2(x)}{f_1(x)} \int_a^x \frac{f_1(t)r(t)}{W(t)} dt - \int_a^x \frac{f_2(t)r(t)}{W(t)} dt.$$

This leads to the particular solution

$$g(x) = f_1(x)v(x) = f_2(x) \int_a^x \frac{f_1(t)r(t)}{W(t)} dt - f_1(x) \int_a^x \frac{f_2(t)r(t)}{W(t)} dt. \quad (13.16)$$

We have just proved the following result.

**13.4.8. Proposition.** *Given a single solution  $f_1(x)$  of the homogeneous equation corresponding to an HSOLDE, one can use Equation (13.13) to find a second solution  $f_2(x)$  of the homogeneous equation and Equation (13.16) to find a particular solution  $g(x)$ . The most general solution  $h$  will then be*

$$h(x) = c_1 f_1(x) + c_2 f_2(x) + g(x).$$

### 13.4.3 Separation and Comparison Theorems

The Wronskian can be used to derive some properties of the graphs of solutions of HSOLDEs. One such property concerns the relative position of the zeros of two linearly independent solutions of an HSOLDE.

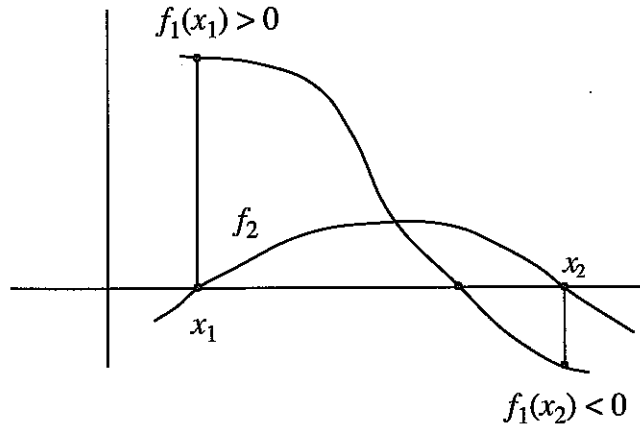
the separation  
theorem

**13.4.9. Theorem.** (the separation theorem) *The zeros of two linearly independent solutions of an HSOLDE occur alternately.*

*Proof.* Let  $f_1(x)$  and  $f_2(x)$  be two independent solutions of Equation (13.12). We have to show that a zero of  $f_1$  exists between any two zeros of  $f_2$ . The linear independence of  $f_1$  and  $f_2$  implies that  $W(f_1, f_2; x) \neq 0$  for any  $x \in [a, b]$ . Let  $x_i \in [a, b]$  be a zero of  $f_2$ . Then

$$0 \neq W(f_1, f_2; x_i) = f_1(x_i)f_2'(x_i) - f_2(x_i)f_1'(x_i) = f_1(x_i)f_2'(x_i).$$

Thus,  $f_1(x_i) \neq 0$  and  $f_2'(x_i) \neq 0$ . Suppose that  $x_1$  and  $x_2$ —where  $x_2 > x_1$ —are two successive zeros of  $f_2$ . Since  $f_2$  is continuous in  $[a, b]$  and  $f_2'(x_1) \neq 0$ ,  $f_2$  has to be either increasing [ $f_2'(x_1) > 0$ ] or decreasing [ $f_2'(x_1) < 0$ ] at  $x_1$ . For  $f_2$  to be zero at  $x_2$ , the next point,  $f_2'(x_2)$  must have the *opposite* sign from  $f_2'(x_1)$  (see Figure 13.1). We proved earlier that the sign of the Wronskian does not change in  $[a, b]$  (see Proposition 13.4.2 and comments after it). The above equation then says that  $f_1(x_1)$  and  $f_1(x_2)$  also have opposite signs. The continuity of  $f_1$  then implies that  $f_1$  must cross the  $x$ -axis somewhere between  $x_1$  and  $x_2$ . A similar argument shows that there exists one zero of  $f_2$  between any two zeros of  $f_1$ .  $\square$



**Figure 13.1** If  $f_2'(x_1) > 0 > f_2'(x_2)$ , then (assuming that the Wronskian is positive)  $f_1(x_1) > 0 > f_1(x_2)$ .

**13.4.10. Example.** Two linearly independent solutions of  $y'' + y = 0$  are  $\sin x$  and  $\cos x$ . The separation theorem suggests that the zeros of  $\sin x$  and  $\cos x$  must alternate, a fact known from elementary trigonometry: The zeros of  $\cos x$  occur at odd multiples of  $\pi/2$ , and those of  $\sin x$  occur at even multiples of  $\pi/2$ . ■

A second useful result is known as the comparison theorem (for a proof, see [Birk 78, p. 38]).

the comparison theorem

**13.4.11. Theorem.** (the comparison theorem) *Let  $f$  and  $g$  be nontrivial solutions of  $u'' + p(x)u = 0$  and  $v'' + q(x)v = 0$ , respectively, where  $p(x) \geq q(x)$  for all  $x \in [a, b]$ . Then  $f$  vanishes at least once between any two zeros of  $g$ , unless  $p = q$  and  $f$  is a constant multiple of  $g$ .*

The form of the differential equations used in the comparison theorem is not restrictive because any HSOLDE can be cast in this form, as the following example shows.

**13.4.12. Example.** We show that  $y'' + p(x)y' + q(x)y = 0$  can be cast in the form  $u'' + S(x)u = 0$  by an appropriate functional transformation. Define  $w(x)$  by  $y = wu$ , and substitute in the HSOLDE to obtain

$$(u'w + w'u)' + p(u'w + w'u) + quw = 0,$$

or

$$wu'' + (2w' + pw)u' + (qw + pw' + w'')u = 0. \tag{13.17}$$

If we demand that the coefficient of  $u'$  be zero, we obtain the DE  $2w' + pw = 0$ , whose solution is

$$w(x) = C \exp \left[ -\frac{1}{2} \int_a^x p(t) dt \right].$$

Dividing (13.17) by this  $w$  and substituting for  $w$  yields

$$u'' + S(x)u = 0, \quad \text{where} \quad S(x) = q + p \frac{w'}{w} + \frac{w''}{w} = q - \frac{1}{4}p^2 - \frac{1}{2}p'. \quad \blacksquare$$

A useful special case of the comparison theorem is given as the following corollary whose straightforward but instructive proof is left as a problem.

**13.4.13. Corollary.** *If  $q(x) \leq 0$  for all  $x \in [a, b]$ , then no nontrivial solution of the differential equation  $v'' + q(x)v = 0$  can have more than one zero.*

**13.4.14. Example.** It should be clear from the preceding discussion that the oscillations of the solutions of  $v'' + q(x)v = 0$  are mostly determined by the sign and magnitude of  $q(x)$ . For  $q(x) \leq 0$  there is no oscillation; that is, there is no solution that changes sign more than once. Now suppose that  $q(x) \geq k^2 > 0$  for some real  $k$ . Then, by Theorem 13.4.11, any solution of  $v'' + q(x)v = 0$  must have at least one zero between any two successive zeros of the solution  $\sin kx$  of  $u'' + k^2u = 0$ . This means that any solution of  $v'' + q(x)v = 0$  has a zero in any interval of length  $\pi/k$  if  $q(x) \geq k^2 > 0$ .

Let us apply this to the Bessel DE,

$$y'' + \frac{1}{x}y' + \left(1 - \frac{n^2}{x^2}\right)y = 0.$$

We can eliminate the  $y'$  term by substituting  $v/\sqrt{x}$  for  $y$ .<sup>5</sup> This transforms the Bessel DE into.

$$v'' + \left(1 - \frac{4n^2 - 1}{4x^2}\right)v = 0.$$

oscillation of the  
Bessel function of  
order zero

We compare this, for  $n = 0$ , with  $u'' + u = 0$ , which has a solution  $u = \sin x$ , and conclude that each interval of length  $\pi$  of the positive  $x$ -axis contains at least one zero of any solution of order zero ( $n = 0$ ) of the Bessel equation. Thus, in particular, the zeroth Bessel function, denoted by  $J_0(x)$ , has a zero in each interval of length  $\pi$  of the  $x$ -axis.

On the other hand, for  $4n^2 - 1 > 0$ , or  $n > \frac{1}{2}$ , we have  $1 > [1 - (4n^2 - 1)/4x^2]$ . This implies that  $\sin x$  has *at least* one zero between any two successive zeros of the Bessel functions of order greater than  $\frac{1}{2}$ . It follows that such a Bessel function can have *at most* one zero between any two successive zeros of  $\sin x$  (or in each interval of length  $\pi$  on the positive  $x$ -axis).  $\blacksquare$

**13.4.15. Example.** Let us apply Corollary 13.4.13 to  $v'' - v = 0$  in which  $q(x) = -1 < 0$ . According to the corollary, the most general solution,  $c_1e^x + c_2e^{-x}$ , can have at most one zero. Indeed,

$$c_1e^x + c_2e^{-x} = 0 \Rightarrow x = \frac{1}{2} \ln \left| -\frac{c_2}{c_1} \right|,$$

and this (real)  $x$  (if it exists) is the only possible solution, as predicted by the corollary.  $\blacksquare$

<sup>5</sup>Because of the square root in the denominator, the range of  $x$  will have to be restricted to positive values.

### 13.5 Adjoint Differential Operators

We discussed adjoint operators in detail in the context of finite-dimensional vector spaces in Chapter 2. In particular, the importance of self-adjoint, or hermitian, operators was clearly spelled out by the spectral decomposition theorem of Chapter 4. A consequence of that theorem is the completeness of the eigenvectors of a hermitian operator, the fact that an arbitrary vector can be expressed as a linear combination of the (orthonormal) eigenvectors of a hermitian operator.

Self-adjoint differential operators are equally important because their “eigenfunctions” also form complete orthogonal sets, as we shall see later. This section will generalize the concept of the adjoint to the case of a differential operator (of second degree).

**13.5.1. Definition.** *The HSOLDE*

$$\mathbf{L}[y] \equiv p_2(x)y'' + p_1(x)y' + p_0(x)y = 0 \tag{13.18}$$

exact SOLDE is said to be *exact* if

$$\mathbf{L}[f] \equiv p_2(x)f'' + p_1(x)f' + p_0(x)f = \frac{d}{dx}[A(x)f' + B(x)f] \tag{13.19}$$

integrating factor for SOLDE for all  $f \in \mathcal{C}^2[a, b]$  and for some  $A, B \in \mathcal{C}^1[a, b]$ . An *integrating factor* for  $\mathbf{L}[y]$  is a function  $\mu(x)$  such that  $\mu(x)\mathbf{L}[y]$  is exact.

If an integrating factor exists, then Equation (13.18) reduces to

$$\frac{d}{dx}[A(x)y' + B(x)y] = 0 \Rightarrow A(x)y' + B(x)y = C,$$

a FOLDE with a constant inhomogeneous term. Even the ISOLDE corresponding to Equation (13.18) can be solved, because

$$\begin{aligned} \mu(x)\mathbf{L}[y] = \mu(x)r(x) &\Rightarrow \frac{d}{dx}[A(x)y' + B(x)y] = \mu(x)r(x) \\ &\Rightarrow A(x)y' + B(x)y = \int_a^x \mu(t)r(t) dt, \end{aligned}$$

which is a general FOLDE. Thus, the existence of an integrating factor completely solves a SOLDE. It is therefore important to know whether or not a SOLDE admits an integrating factor. First let us give a criterion for the *exactness* of a SOLDE.

**13.5.2. Proposition.** *The SOLDE of Equation (13.18) is exact if and only if  $p_2'' - p_1' + p_0 = 0$ .*

*Proof.* If the SOLDE is exact, then Equation (13.19) holds for all  $f$ , implying that  $p_2 = A$ ,  $p_1 = A' + B$ , and  $p_0 = B'$ . It follows that  $p_2'' = A''$ ,  $p_1' = A'' + B'$ , and  $p_0 = B'$ , which in turn give  $p_2'' - p_1' + p_0 = 0$ .

Conversely if  $p_2'' - p_1' + p_0 = 0$ , then, substituting  $p_0 = -p_2'' + p_1'$  in the LHS of Equation (13.18), we obtain

$$\begin{aligned} p_2y'' + p_1y' + p_0y &= p_2y'' + p_1y' + (-p_2'' + p_1')y \\ &= p_2y'' - p_2''y + (p_1y)' = (p_2y' - p_2'y)' + (p_1y)' \\ &= \frac{d}{dx}(p_2y' - p_2'y + p_1y), \end{aligned}$$

and the DE is exact. □

A general SOLDE is clearly not exact. Can we make it exact by multiplying it by an integrating factor as we did with a FOLDE? The following proposition contains the answer.

**13.5.3. Proposition.** *A function  $\mu$  is an integrating factor of the SOLDE of Equation (13.18) if and only if it is a solution of the HSOLDE*

$$\mathbf{M}[\mu] \equiv (p_2\mu)'' - (p_1\mu)' + p_0\mu = 0. \tag{13.20}$$

*Proof.* This is an immediate consequence of Proposition 13.5.2. □

We can expand Equation (13.20) to obtain the equivalent equation

$$p_2\mu'' + (2p_2' - p_1)\mu' + (p_2'' - p_1' + p_0)\mu = 0. \tag{13.21}$$

adjoint of a  
second-order linear  
differential operator

The operator  $\mathbf{M}$  given by

$$\mathbf{M} \equiv p_2 \frac{d^2}{dx^2} + (2p_2' - p_1) \frac{d}{dx} + (p_2'' - p_1' + p_0) \tag{13.22}$$

is called the **adjoint** of the operator  $\mathbf{L}$  and denoted by  $\mathbf{M} \equiv \mathbf{L}^\dagger$ . The reason for the use of the word “adjoint” will be made clear below.

Proposition 13.5.3 confirms the existence of an integrating factor. However, the latter can be obtained only by solving Equation (13.21), which is at least as difficult as solving the original differential equation! In contrast, the integrating factor for a FOLDE can be obtained by a mere integration [see Equation (13.8)].

Although integrating factors for SOLDEs are not as useful as their counterparts for FOLDEs, they can facilitate the study of SOLDEs. Let us first note that the adjoint of the adjoint of a differential operator is the original operator:  $(\mathbf{L}^\dagger)^\dagger = \mathbf{L}$  (see Problem 13.11). This suggests that if  $v$  is an integrating factor of  $\mathbf{L}[u]$ , then  $u$  will be an integrating factor of  $\mathbf{M}[v] \equiv \mathbf{L}^\dagger[v]$ . In particular, multiplying the first one by  $v$  and the second one by  $u$  and subtracting the results, we obtain [see Equations (13.18) and (13.20)]  $v\mathbf{L}[u] - u\mathbf{M}[v] = (vp_2)u'' - u(p_2v)'' + (vp_1)u' + u(p_1v)'$ , which can be simplified to

$$v\mathbf{L}[u] - u\mathbf{M}[v] = \frac{d}{dx}[p_2vu' - (p_2v)'u + p_1uv]. \tag{13.23}$$

Lagrange identities Integrating this from  $a$  to  $b$  yields

$$\int_a^b (v\mathbf{L}[u] - u\mathbf{M}[v]) dx = [p_2vu' - (p_2v)'u + p_1uv]_a^b. \quad (13.24)$$

Equations (13.23) and (13.24) are called the **Lagrange identities**. Equation (13.24) embodies the reason for calling  $\mathbf{M}$  the adjoint of  $\mathbf{L}$ : If we consider  $u$  and  $v$  as abstract vectors  $|u\rangle$  and  $|v\rangle$ ,  $\mathbf{L}$  and  $\mathbf{M}$  as operators in a Hilbert space with the inner product  $\langle u|v\rangle = \int_a^b u^*(x)v(x) dx$ , then Equation (13.24) can be written as

$$\langle v|\mathbf{L}|u\rangle - \langle u|\mathbf{M}|v\rangle = \langle u|\mathbf{L}^\dagger|v\rangle^* - \langle u|\mathbf{M}|v\rangle = [p_2vu' - (p_2v)'u + p_1uv]_a^b.$$

If the RHS is zero, then  $\langle u|\mathbf{L}^\dagger|v\rangle^* = \langle u|\mathbf{M}|v\rangle$  for all  $|u\rangle, |v\rangle$ , and since all these operators and functions are real,  $\mathbf{L}^\dagger = \mathbf{M}$ .

As in the case of finite-dimensional vector spaces, a self-adjoint differential operator merits special consideration. For  $\mathbf{M}[v] \equiv \mathbf{L}^\dagger[v]$  to be equal to  $\mathbf{L}$ , we must have [see Equations (13.18) and (13.21)]  $2p_2' - p_1 = p_1$  and  $p_2'' - p_1' + p_0 = p_0$ . The first equation gives  $p_2' = p_1$ , which also solves the second equation. If this condition holds, then we can write Equation (13.18) as  $\mathbf{L}[y] = p_2y'' + p_2'y' + p_0y$ , or

$$\mathbf{L}[y] = \frac{d}{dx} \left[ p_2(x) \frac{dy}{dx} \right] + p_0(x)y = 0.$$

Can we make all SOLDEs self-adjoint? Let us multiply both sides of Equation (13.18) by a function  $h(x)$ , to be determined later. We get the new DE

$$h(x)p_2(x)y'' + h(x)p_1(x)y' + h(x)p_0(x)y = 0,$$

which we desire to be self-adjoint. This will be accomplished if we choose  $h(x)$  such that  $hp_1 = (hp_2)'$ , or  $p_2h' + h(p_2' - p_1) = 0$ , which can be readily integrated to give

$$h(x) = \frac{1}{p_2} \exp \left[ \int^x \frac{p_1(t)}{p_2(t)} dt \right].$$

We have just proved the following:

**13.5.4. Theorem.** *The SOLDE of Equation (13.18) is self-adjoint if and only if  $p_2' = p_1$ , in which case the DE has the form*

$$\frac{d}{dx} \left[ p_2(x) \frac{dy}{dx} \right] + p_0(x)y = 0.$$

*If it is not self-adjoint, it can be made so by multiplying it through by*

$$h(x) = \frac{1}{p_2} \exp \left[ \int^x \frac{p_1(t)}{p_2(t)} dt \right].$$



**13.5.5. Example.** (a) The Legendre equation in normal form,

$$y'' - \frac{2x}{1-x^2}y' + \frac{\lambda}{1-x^2}y = 0,$$

is not self-adjoint. However, we get a self-adjoint version if we multiply through by  $h(x) = 1 - x^2$ :

$$(1-x^2)y'' - 2xy' + \lambda y = 0,$$

or  $[(1-x^2)y']' + \lambda y = 0$ .

(b) Similarly, the normal form of the Bessel equation

$$y'' + \frac{1}{x}y' + \left(1 - \frac{n^2}{x^2}\right)y = 0$$

is not self-adjoint, but multiplying through by  $h(x) = x$  yields

$$\frac{d}{dx} \left( x \frac{dy}{dx} \right) + \left( x - \frac{n^2}{x} \right) y = 0,$$

which is clearly self-adjoint. ■

## 13.6 Power-Series Solutions of SOLDES

Analysis is one of the richest branches of mathematics, focusing on the endless variety of objects we call functions. The simplest kind of function is a polynomial, which is obtained by performing the simple algebraic operations of addition and multiplication on the independent variable  $x$ . The next in complexity are the trigonometric functions, which are obtained by taking ratios of geometric objects. If we demand a simplistic, intuitive approach to functions, the list ends there. It was only with the advent of derivatives, integrals, and differential equations that a vastly rich variety of functions exploded into existence in the eighteenth and nineteenth centuries. For instance,  $e^x$ , nonexistent before the invention of calculus, can be thought of as the function that solves  $dy/dx = y$ .

Although the definition of a function in terms of DEs and integrals seems a bit artificial, for most applications it is the only way to define a function. For instance, the error function, used in statistics, is defined as

$$\operatorname{erf}(x) \equiv \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-t^2} dt.$$

Such a function cannot be expressed in terms of elementary functions. Similarly, functions (of  $x$ ) such as

$$\int_x^\infty \frac{\sin t}{t} dt, \quad \int_0^{\pi/2} \sqrt{1-x^2 \sin^2 t} dt, \quad \int_0^{\pi/2} \frac{dt}{\sqrt{1-x^2 \sin^2 t}},$$

and so on are encountered frequently in applications. None of these functions can be expressed in terms of other well-known functions.

An effective way of studying such functions is to study the differential equations they satisfy. In fact, the majority of functions encountered in mathematical physics obey the HSOLDE of Equation (13.18) in which the  $p_i(x)$  are elementary functions, mostly ratios of polynomials (of degree at most 2). Of course, to specify functions completely, appropriate boundary conditions are necessary. For instance, the error function mentioned above satisfies the HSOLDE  $y'' + 2xy' = 0$  with the boundary conditions  $y(0) = \frac{1}{2}$  and  $y'(0) = 1/\sqrt{\pi}$ .

The natural tendency to resist the idea of a function as a solution of a SOLDE is mostly due to the abstract nature of differential equations. After all, it is easier to imagine constructing functions by simple multiplications or with simple geometric figures that have been around for centuries. The following beautiful example (see [Birk 78, pp. 85–87]) should overcome this resistance and convince the skeptic that differential equations contain all the information about a function.

**13.6.1. Example.** We can show that the solutions to  $y'' + y = 0$  have all the properties we expect of  $\sin x$  and  $\cos x$ . Let us denote the two linearly independent solutions of this equation by  $C(x)$  and  $S(x)$ . To specify these functions completely, we set  $C(0) = S'(0) = 1$ , and  $C'(0) = S(0) = 0$ . We claim that this information is enough to identify  $C(x)$  and  $S(x)$  as  $\cos x$  and  $\sin x$ , respectively.

First, let us show that the solutions exist and are well-behaved functions. With  $C(0)$  and  $C'(0)$  given, the equation  $y'' + y = 0$  can generate all derivatives of  $C(x)$  at zero:  $C''(0) = -C(0) = -1$ ,  $C'''(0) = -C'(0) = 0$ ,  $C^{(4)}(0) = -C''(0) = +1$ , and, in general,

$$C^{(n)}(0) = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ (-1)^k & \text{if } n = 2k \text{ where } k = 0, 1, 2, \dots \end{cases}$$

Thus, the Taylor expansion of  $C(x)$  is

$$C(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}. \quad (13.25)$$

Similarly,

$$S(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}. \quad (13.26)$$

Example illustrates that all information about sine and cosine is hidden in their differential equation

A simple ratio test on the series representation of  $C(x)$  yields

$$\lim_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} = \lim_{k \rightarrow \infty} \frac{(-1)^{k+1} x^{2(k+1)} / (2k+2)!}{(-1)^k x^{2k} / (2k)!} = \lim_{k \rightarrow \infty} \frac{-x^2}{(2k+2)(2k+1)} = 0,$$

which shows that the series for  $C(x)$  converges for all values of  $x$ . Similarly, the series for  $S(x)$  is also convergent. Thus, we are dealing with well-defined finite-valued functions.

Let us now enumerate and prove some properties of  $C(x)$  and  $S(x)$ .

(a)  $C'(x) = -S(x)$ .

We prove this relation by differentiating  $C''(x) + C(x) = 0$  and writing the result as

$[C'(x)]'' + C'(x) = 0$  to make evident the fact that  $C'(x)$  is also a solution. Since  $C'(0) = 0$  and  $[C'(0)]' = C''(0) = -1$ , and since  $-S(x)$  satisfies the same initial conditions, the uniqueness theorem implies that  $C'(x) = -S(x)$ . Similarly,  $S'(x) = C(x)$ .

(b)  $C^2(x) + S^2(x) = 1$ .

Since the  $p(x)$  term is absent from the SOLDE, Proposition 13.4.2 implies that the Wronskian of  $C(x)$  and  $S(x)$  is constant. On the other hand,

$$\begin{aligned} W(C, S; x) &= C(x)S'(x) - C'(x)S(x) = C^2(x) + S^2(x) \\ &= W(C, S; 0) = C^2(0) + S^2(0) = 1. \end{aligned}$$

(c)  $S(a+x) = S(a)C(x) + C(a)S(x)$ .

The use of the chain rule easily shows that  $S(a+x)$  is a solution of the equation  $y'' + y = 0$ . Thus, it can be written as a linear combination of  $C(x)$  and  $S(x)$  [which are linearly independent because their Wronskian is nonzero by (b)]:

$$S(a+x) = AS(x) + BC(x). \quad (13.27)$$

This is a functional identity, which for  $x = 0$  gives  $S(a) = BC(0) = B$ . If we differentiate both sides of Equation (13.27), we get

$$C(a+x) = AS'(x) + BC'(x) = AC(x) - BS(x),$$

which for  $x = 0$  gives  $C(a) = A$ . Substituting the values of  $A$  and  $B$  in Equation (13.27) yields the desired identity. A similar argument leads to

$$C(a+x) = C(a)C(x) - S(a)S(x).$$

(d) Periodicity of  $C(x)$  and  $S(x)$ .

Let  $x_0$  be the smallest positive real number such that  $S(x_0) = C(x_0)$ . Then property (b) implies that  $C(x_0) = S(x_0) = 1/\sqrt{2}$ . On the other hand,

$$\begin{aligned} S(x_0+x) &= S(x_0)C(x) + C(x_0)S(x) = C(x_0)C(x) + S(x_0)S(x) \\ &= C(x_0)C(x) - S(x_0)S(-x) = C(x_0-x). \end{aligned}$$

The third equality follows because by Equation (13.26),  $S(x)$  is an odd function of  $x$ . This is true for all  $x$ ; in particular, for  $x = x_0$  it yields  $S(2x_0) = C(0) = 1$ , and by property (b),  $C(2x_0) = 0$ . Using property (c) once more, we get

$$\begin{aligned} S(2x_0+x) &= S(2x_0)C(x) + C(2x_0)S(x) = C(x), \\ C(2x_0+x) &= C(2x_0)C(x) - S(2x_0)S(x) = -S(x). \end{aligned}$$

Substituting  $x = 2x_0$  yields  $S(4x_0) = C(2x_0) = 0$  and  $C(4x_0) = -S(2x_0) = -1$ . Continuing in this manner, we can easily obtain

$$S(8x_0+x) = S(x), \quad C(8x_0+x) = C(x),$$

which prove the periodicity of  $S(x)$  and  $C(x)$  and show that their period is  $8x_0$ . It is even possible to determine  $x_0$ . This determination is left as a problem, but the result is

$$x_0 = \int_0^{1/\sqrt{2}} \frac{dt}{\sqrt{1-t^2}}.$$

A numerical calculation will show that this is  $\pi/4$ . ■

### 13.6.1 Frobenius Method of Undetermined Coefficients

A proper treatment of SOLDEs requires the medium of complex analysis and will be undertaken in the next chapter. At this point, however, we are seeking a *formal* infinite series solution to the SOLDE

$$y'' + p(x)y' + q(x)y = 0,$$

where  $p(x)$  and  $q(x)$  are real and analytic. This means that  $p(x)$  and  $q(x)$  can be represented by convergent power series in some interval  $(a, b)$ . [The interesting case where  $p(x)$  and  $q(x)$  may have singularities will be treated in the context of complex solutions.]

The general procedure is to write the expansions<sup>6</sup>

$$p(x) = \sum_{k=0}^{\infty} a_k x^k, \quad q(x) = \sum_{k=0}^{\infty} b_k x^k, \quad y = \sum_{k=0}^{\infty} c_k x^k \quad (13.28)$$

for the coefficient functions  $p$  and  $q$  and the solution  $y$ , substitute them in the SOLDE, and equate the coefficient of each power of  $x$  to zero. For this purpose, we need expansions for derivatives of  $y$ :

$$y' = \sum_{k=1}^{\infty} k c_k x^{k-1} = \sum_{k=0}^{\infty} (k+1) c_{k+1} x^k,$$

$$y'' = \sum_{k=1}^{\infty} (k+1) k c_{k+1} x^{k-1} = \sum_{k=0}^{\infty} (k+2)(k+1) c_{k+2} x^k.$$

Thus

$$p(x)y' = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} a_m x^m (k+1) c_{k+1} x^k = \sum_{k,m} (k+1) a_m c_{k+1} x^{k+m}.$$

Let  $k+m \equiv n$  and sum over  $n$ . Then the other sum, say  $m$ , cannot exceed  $n$ . Thus,

$$p(x)y' = \sum_{n=0}^{\infty} \sum_{m=0}^n (n-m+1) a_m c_{n-m+1} x^n.$$

Similarly,  $q(x)y = \sum_{n=0}^{\infty} \sum_{m=0}^n b_m c_{n-m} x^n$ . Substituting these sums and the series for  $y''$  in the SOLDE, we obtain

$$\sum_{n=0}^{\infty} \left\{ (n+1)(n+2)c_{n+2} + \sum_{m=0}^n [(n-m+1)a_m c_{n-m+1} + b_m c_{n-m}] \right\} x^n = 0.$$

<sup>6</sup>Here we are expanding about the origin. If such an expansion is impossible or inconvenient, one can expand about another point, say  $x_0$ . One would then replace all powers of  $x$  in all expressions below with powers of  $x - x_0$ . These expansions assume that  $p$ ,  $q$ , and  $y$  have no singularity at  $x = 0$ . In general, this assumption is not valid, and a different approach, in which the whole series is multiplied by a (not necessarily positive integer) power of  $x$ , ought to be taken. Details are provided in Chapter 14.

For this to be true for all  $x$ , the coefficient of each power of  $x$  must vanish:

$$(n+1)(n+2)c_{n+2} = -\sum_{m=0}^n [(n-m+1)a_m c_{n-m+1} + b_m c_{n-m}] \quad \text{for } n \geq 0,$$

or

$$n(n+1)c_{n+1} = -\sum_{m=0}^{n-1} [(n-m)a_m c_{n-m} + b_m c_{n-m-1}] \quad \text{for } n \geq 1. \quad (13.29)$$

If we know  $c_0$  and  $c_1$  (for instance from boundary conditions), we can uniquely determine  $c_n$  for  $n \geq 2$  from Equation (13.29). This, in turn, gives a unique power-series expansion for  $y$ , and we have the following theorem.

the SOLDE existence  
theorem

**13.6.2. Theorem.** (the existence theorem) *For any SOLDE of the form  $y'' + p(x)y' + q(x)y = 0$  with analytic coefficient functions given by the first two equations of (13.28), there exists a unique power series, given by the third equation of (13.28) that formally satisfies the SOLDE for each choice of  $c_0$  and  $c_1$ .*

This theorem merely states the existence of a formal power series and says nothing about its convergence. The following example will demonstrate that convergence is not necessarily guaranteed.

**13.6.3. Example.** The formal power-series solution for  $x^2 y' - y + x = 0$  can be obtained by letting  $y = \sum_{n=0}^{\infty} c_n x^n$ . Then  $y' = \sum_{n=0}^{\infty} (n+1)c_{n+1} x^n$ , and substitution in the DE gives  $\sum_{n=0}^{\infty} (n+1)c_{n+1} x^{n+2} - \sum_{n=0}^{\infty} c_n x^n + x = 0$ , or

$$\sum_{n=0}^{\infty} (n+1)c_{n+1} x^{n+2} - c_0 - c_1 x - \sum_{n=2}^{\infty} c_n x^n + x = 0.$$

We see that  $c_0 = 0$ ,  $c_1 = 1$ , and  $(n+1)c_{n+1} = c_{n+2}$  for  $n \geq 0$ . Thus, we have the recursion relation  $nc_n = c_{n+1}$  for  $n \geq 1$  whose unique solution is  $c_n = (n-1)!$ , which generates the following solution for the DE:

$$y = x + x^2 + (2!)x^3 + (3!)x^4 + \cdots + (n-1)!x^n + \cdots$$

This series is not convergent for any nonzero  $x$ . ■

As we shall see later, for *normal* SOLDEs, the power series of  $y$  in Equation (13.28) converges to an analytic function. The SOLDE solved in the preceding example is not normal.

**13.6.4. Example.** As an application of Theorem 13.6.2, let us consider the Legendre equation in its normal form

$$y'' - \frac{2x}{1-x^2}y' + \frac{\lambda}{1-x^2}y = 0.$$

For  $|x| < 1$  both  $p$  and  $q$  are analytic, and

$$p(x) = -2x \sum_{m=0}^{\infty} (x^2)^m = \sum_{m=0}^{\infty} (-2)x^{2m+1},$$

$$q(x) = \lambda \sum_{m=0}^{\infty} (x^2)^m = \sum_{m=0}^{\infty} \lambda x^{2m}.$$

Thus, the coefficients of Equation (13.28) are

$$a_m = \begin{cases} 0 & \text{if } m \text{ is even,} \\ -2 & \text{if } m \text{ is odd} \end{cases} \quad \text{and} \quad b_m = \begin{cases} \lambda & \text{if } m \text{ is even,} \\ 0 & \text{if } m \text{ is odd.} \end{cases}$$

We want to substitute for  $a_m$  and  $b_m$  in Equation (13.29) to find  $c_{n+1}$ . It is convenient to consider two cases: when  $n$  is odd and when  $n$  is even. For  $n = 2r + 1$ , Equation (13.29)—after some algebra—yields

$$(2r + 1)(2r + 2)c_{2r+2} = \sum_{m=0}^r (4r - 4m - \lambda)c_{2(r-m)}. \quad (13.30)$$

With  $r \rightarrow r + 1$ , this becomes

$$\begin{aligned} (2r + 3)(2r + 4)c_{2r+4} &= \sum_{m=0}^{r+1} (4r + 4 - 4m - \lambda)c_{2(r+1-m)} \\ &= (4r + 4 - \lambda)c_{2(r+1)} + \sum_{m=1}^{r+1} (4r + 4 - 4m - \lambda)c_{2(r+1-m)} \\ &= (4r + 4 - \lambda)c_{2r+2} + \sum_{m=0}^r (4r - 4m - \lambda)c_{2(r-m)} \\ &= (4r + 4 - \lambda)c_{2r+2} + (2r + 1)(2r + 2)c_{2r+2} \\ &= [-\lambda + (2r + 3)(2r + 2)]c_{2r+2}, \end{aligned}$$

where in going from the second equality to the third we changed the dummy index, and in going from the third equality to the fourth we used Equation (13.30). Now we let  $2r + 2 \equiv k$  to obtain  $(k + 1)(k + 2)c_{k+2} = [k(k + 1) - \lambda]c_k$ , or

$$c_{k+2} = \frac{k(k + 1) - \lambda}{(k + 1)(k + 2)} c_k \quad \text{for even } k.$$

It is not difficult to show that starting with  $n = 2r$ , the case of even  $n$ , we obtain this same equation for odd  $k$ . Thus, we can write

$$c_{n+2} = \frac{n(n + 1) - \lambda}{(n + 1)(n + 2)} c_n. \quad (13.31)$$

For arbitrary  $c_0$  and  $c_1$ , we obtain two independent solutions, one of which has only even powers of  $x$  and the other only odd powers. The generalized ratio test (see [Hass 99, Chapter 5]) shows that the series is divergent for  $x = \pm 1$  unless  $\lambda = l(l + 1)$  for some positive

integer  $l$ . In that case the infinite series becomes a polynomial, the Legendre polynomial encountered in Chapter 7.

Equation (13.31) could have been obtained by substituting Equation (13.28) directly into the Legendre equation. The roundabout way to (13.31) taken here shows the generality of Equation (13.29). With specific differential equations it is generally better to substitute (13.28) directly. ■

quantum harmonic oscillator: power series method

**13.6.5. Example.** We studied Hermite polynomials in Chapter 7 in the context of classical orthogonal polynomials. Let us see how they arise in physics.

The one-dimensional time-independent Schrödinger equation for a particle of mass  $m$  in a potential  $V(x)$  is

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V(x)\psi = E\psi,$$

where  $E$  is the total energy of the particle.

For a harmonic oscillator,  $V(x) = \frac{1}{2}kx^2 \equiv \frac{1}{2}m\omega^2x^2$  and

$$\psi'' - \frac{m^2\omega^2}{\hbar^2}x^2\psi + \frac{2m}{\hbar^2}E\psi = 0.$$

Substituting  $\psi(x) = H(x) \exp(-m\omega x^2/2\hbar)$  and then making the change of variables  $x = (1/\sqrt{m\omega/\hbar})y$  yields

$$H'' - 2yH' + \lambda H = 0 \quad \text{where} \quad \lambda = \frac{2E}{\hbar\omega} - 1. \tag{13.32}$$

This is the Hermite differential equation in normal form. We assume the expansion  $H(y) = \sum_{n=0}^{\infty} c_n y^n$  which yields

$$H'(y) = \sum_{n=1}^{\infty} n c_n y^{n-1} = \sum_{n=0}^{\infty} (n+1) c_{n+1} y^n,$$

$$H''(y) = \sum_{n=1}^{\infty} n(n+1) c_{n+1} y^{n-1} = \sum_{n=0}^{\infty} (n+1)(n+2) c_{n+2} y^n.$$

Substituting in Equation (13.32) gives

$$\sum_{n=0}^{\infty} [(n+1)(n+2)c_{n+2} + \lambda c_n] y^n - 2 \sum_{n=0}^{\infty} (n+1)c_{n+1} y^{n+1} = 0,$$

or

$$2c_2 + \lambda c_0 + \sum_{n=0}^{\infty} [(n+2)(n+3)c_{n+3} + \lambda c_{n+1} - 2(n+1)c_{n+1}] y^{n+1} = 0.$$

Setting the coefficients of powers of  $y$  equal to zero, we obtain

$$c_2 = -\frac{\lambda}{2}c_0,$$

$$c_{n+3} = \frac{2(n+1) - \lambda}{(n+2)(n+3)} c_{n+1} \quad \text{for } n \geq 0,$$

or, replacing  $n$  with  $n - 1$ ,

$$c_{n+2} = \frac{2n - \lambda}{(n + 1)(n + 2)} c_n, \quad n \geq 1. \tag{13.33}$$

The ratio test yields easily that the series is convergent for all values of  $y$ .

Thus, the infinite series whose coefficients obey the recursive relation in Equation (13.33) converges for all  $y$ . However, on physical grounds, i.e., the demand that  $\lim_{x \rightarrow \infty} \psi(x) = 0$ , the series must be truncated. This happens only if  $\lambda = 2l$  for some integer  $l$  (see Problem 13.20 and [Hass 99, Chapter 13]), and in that case we obtain a polynomial, the Hermite polynomial of order  $l$ . A consequence of such a truncation is the quantization of harmonic oscillator energy:

$$2l = \lambda = \frac{2E}{\hbar\omega} - 1 \Rightarrow E = (l + \frac{1}{2})\hbar\omega.$$

Two solutions are generated from Equation (13.33), one including only even powers and the other only odd powers. These are clearly linearly independent. Thus, knowledge of  $c_0$  and  $c_1$  determines the general solution of the HSOLDE of (13.32). ■

The preceding two examples show how certain special functions used in mathematical physics are obtained in an analytic way, by solving a differential equation. We saw in Chapter 12 how to obtain spherical harmonics and Legendre polynomials by algebraic methods. It is instructive to solve the harmonic oscillator problem using algebraic methods, as the following example demonstrates.

quantum harmonic oscillator: algebraic method

**13.6.6. Example.** The Hamiltonian of a one-dimensional harmonic oscillator is

$$H = \frac{\mathbf{p}^2}{2m} + \frac{1}{2}m\omega^2 x^2,$$

where  $\mathbf{p} = -i\hbar d/dx$  is the momentum operator. Let us find the eigenvectors and eigenvalues of  $H$ .

We define the operators

$$\mathbf{a} \equiv \sqrt{\frac{m\omega}{2\hbar}} x + i \frac{\mathbf{p}}{\sqrt{2m\hbar\omega}} \quad \text{and} \quad \mathbf{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}} x - i \frac{\mathbf{p}}{\sqrt{2m\hbar\omega}}.$$

Using the commutation relation  $[x, \mathbf{p}] = i\hbar 1$ , we can show that

$$[\mathbf{a}, \mathbf{a}^\dagger] = 1 \quad \text{and} \quad H = \hbar\omega \mathbf{a}^\dagger \mathbf{a} + \frac{1}{2}\hbar\omega 1. \tag{13.34}$$

creation and annihilation operators

Furthermore, one can readily show that

$$[H, \mathbf{a}] = -\hbar\omega \mathbf{a}, \quad [H, \mathbf{a}^\dagger] = \hbar\omega \mathbf{a}^\dagger. \tag{13.35}$$

Let  $|\psi_E\rangle$  be the eigenvector corresponding to the eigenvalue  $E$ :  $H|\psi_E\rangle = E|\psi_E\rangle$ , and note that Equation (13.35) gives  $H\mathbf{a}|\psi_E\rangle = (\mathbf{a}H - \hbar\omega \mathbf{a})|\psi_E\rangle = (E - \hbar\omega)\mathbf{a}|\psi_E\rangle$  and  $H\mathbf{a}^\dagger|\psi_E\rangle = (E + \hbar\omega)\mathbf{a}^\dagger|\psi_E\rangle$ . Thus,  $\mathbf{a}|\psi_E\rangle$  is an eigenvector of  $H$ , with eigenvalue  $E - \hbar\omega$ , and  $\mathbf{a}^\dagger|\psi_E\rangle$  is an eigenvector with eigenvalue  $E + \hbar\omega$ . That is why  $\mathbf{a}^\dagger$  and  $\mathbf{a}$  are called the **raising** and **lowering** (or creation and annihilation) operators, respectively. We can write

$$\mathbf{a}|\psi_E\rangle = c_E |\psi_{E-\hbar\omega}\rangle$$



By applying  $\mathbf{a}$  repeatedly, we obtain states of lower and lower energies. But there is a limit to this because  $\mathbf{H}$  is a positive operator: It cannot have a negative eigenvalue. Thus, there must exist a **ground state**,  $|\psi_0\rangle$ , such that  $\mathbf{a}|\psi_0\rangle = 0$ . The energy of this ground state (or the eigenvalue corresponding to  $|\psi_0\rangle$ ) can be obtained:<sup>7</sup>

$$\mathbf{H}|\psi_0\rangle = (\hbar\omega\mathbf{a}^\dagger\mathbf{a} + \frac{1}{2}\hbar\omega)|\psi_0\rangle = \frac{1}{2}\hbar\omega|\psi_0\rangle.$$

Repeated application of the raising operator yields both higher-level states and eigenvalues. We thus define  $|\psi_n\rangle$  by

$$(\mathbf{a}^\dagger)^n|\psi_0\rangle = c_n|\psi_n\rangle, \quad (13.36)$$

where  $c_n$  is a normalizing constant. The energy of  $|\psi_n\rangle$  is  $n$  units higher than the ground state's, or

$$E_n = (n + \frac{1}{2})\hbar\omega,$$

which is what we obtained in the preceding example.

To find  $c_n$ , we demand orthonormality for the  $|\psi_n\rangle$ . Taking the inner product of (13.36) with itself, we can show (see Problem 13.21) that  $|c_n|^2 = n|c_{n-1}|^2$ , or  $|c_n|^2 = n!|c_0|^2$ , which for  $|c_0| = 1$  and real  $c_n$  yields  $c_n = \sqrt{n!}$ . It follows, then, that

$$|\psi_n\rangle = \frac{1}{\sqrt{n!}}(\mathbf{a}^\dagger)^n|\psi_0\rangle. \quad (13.37)$$

quantum harmonic oscillator: connection between algebraic and analytic methods

In terms of functions and derivative operators,  $\mathbf{a}|\psi_0\rangle = 0$  gives

$$\left(\sqrt{\frac{m\omega}{2\hbar}}x + \sqrt{\frac{\hbar}{2m\omega}}\frac{d}{dx}\right)\psi_0(x) = 0$$

with the solution  $\psi_0(x) = c \exp(-m\omega x^2/2\hbar)$ . Normalizing  $\psi_0(x)$  gives

$$1 = \langle\psi_0|\psi_0\rangle = c^2 \int_{-\infty}^{\infty} \exp\left(-\frac{m\omega x^2}{\hbar}\right) dx = c^2 \left(\frac{\hbar\pi}{m\omega}\right)^{1/2}.$$

Thus,

$$\psi_0(x) = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} e^{-m\omega x^2/(2\hbar)}.$$

We can now write Equation (13.37) in terms of differential operators:

$$\psi_n(x) = \frac{1}{\sqrt{n!}} \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \left(\sqrt{\frac{m\omega}{2\hbar}}x - \sqrt{\frac{\hbar}{2m\omega}}\frac{d}{dx}\right)^n e^{-m\omega x^2/(2\hbar)}$$

Defining a new variable  $y = \sqrt{m\omega/\hbar}x$  transforms this equation into

$$\psi_n = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} \left(y - \frac{d}{dy}\right)^n e^{-y^2/2}.$$

<sup>7</sup>From here on, the unit operator  $\mathbf{1}$  will not be shown explicitly.

From this, the relation between Hermite polynomials, and the solutions of the one-dimensional harmonic oscillator as given in the previous example, we can obtain a general formula for  $H_n(x)$ . In particular, if we note that (see Problem 13.21)

$$e^{y^2/2} \left( y - \frac{d}{dy} \right) e^{-y^2/2} = -e^{y^2} \frac{d}{dy} e^{-y^2}$$

and, in general,

$$e^{y^2/2} \left( y - \frac{d}{dy} \right)^n e^{-y^2/2} = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2},$$

we recover the generalized Rodriguez formula of Chapter 7. ■

To end this section, we simply quote the following important theorem (for a proof, see [Birk 78, p. 95]):

**13.6.7. Theorem.** *For any choice of  $c_0$  and  $c_1$ , the radius of convergence of any power series solution  $y = \sum_{k=0}^{\infty} c_k x^k$  for the normal HSOLDE*

$$y'' + p(x)y' + q(x)y = 0$$

*whose coefficients satisfy the recursion relation of (13.29) is at least as large as the smaller of the two radii of convergence of the two series for  $p(x)$  and  $q(x)$ .*

In particular, if  $p(x)$  and  $q(x)$  are analytic in an interval around  $x = 0$ , then the solution of the normal HSOLDE is also analytic in a neighborhood of  $x = 0$ .

## 13.7 SOLDEs with Constant Coefficients

The solution to a SOLDE with constant coefficients can always be found in closed form. In fact, we can treat an  $n$ th-order linear differential equation (NOLDE) with constant coefficients with no extra effort. This brief section outlines the procedure for solving such an equation. For details, the reader is referred to any elementary book on differential equations (see also [Hass 99]). The most general  $n$ th-order linear differential equation (NOLDE) with constant coefficients can be written as

$$\mathbf{L}[y] \equiv y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = r(x). \quad (13.38)$$

The corresponding homogeneous NOLDE (HNOLDE) is obtained by setting  $r(x) = 0$ . Let us consider such a homogeneous case first. The solution to the homogeneous NOLDE

$$\mathbf{L}[y] \equiv y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0 \quad (13.39)$$

can be found by making the exponential substitution  $y = e^{\lambda x}$ , which results in the equation  $\mathbf{L}[e^{\lambda x}] = (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0)e^{\lambda x} = 0$ . This equation will hold only if  $\lambda$  is a root of the **characteristic polynomial**

characteristic  
polynomial of an  
HNOLDE

$$p(\lambda) \equiv \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0,$$

which, by the fundamental theorem of algebra, can be written as

$$p(\lambda) = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_m)^{k_m} \quad (13.40)$$

The  $\lambda_j$  are the distinct (complex) roots of  $p(\lambda)$  and have multiplicity  $k_j$ .

**13.7.1. Theorem.** Let  $\{\lambda_j\}_{j=1}^m$  be the roots of the characteristic polynomial of the real HNLDE of Equation (13.39), and let the respective roots have multiplicities  $\{k_j\}_{j=1}^m$ . Then the functions

$$\{e^{\lambda_j x}, xe^{\lambda_j x}, \dots, x^{k_j-1}e^{\lambda_j x}\}_{j=1}^m$$

are a basis of solutions of Equation (13.39).

When a  $\lambda$  is complex, one can write its corresponding solution in terms of trigonometric functions.

**13.7.2. Example.** An equation that is used in both mechanics and circuit theory is

$$\frac{d^2y}{dt^2} + a\frac{dy}{dt} + by = 0 \quad \text{for } a, b > 0. \quad (13.41)$$

Its characteristic polynomial is  $p(\lambda) = \lambda^2 + a\lambda + b$ , which has the roots

$$\lambda_1 = \frac{1}{2}(-a + \sqrt{a^2 - 4b}) \quad \text{and} \quad \lambda_2 = \frac{1}{2}(-a - \sqrt{a^2 - 4b}).$$

We can distinguish three different possible motions depending on the relative sizes of  $a$  and  $b$ .

(a)  $a^2 > 4b$  (overdamped): Here we have two distinct simple roots. The multiplicities are both one ( $k_1 = k_2 = 1$ ); therefore, the power of  $x$  for both solutions is zero ( $r_1 = r_2 = 0$ ). Let  $\gamma \equiv \frac{1}{2}\sqrt{a^2 - 4b}$ . Then the most general solution is

$$y(t) = e^{-at/2}(c_1e^{\gamma t} + c_2e^{-\gamma t}).$$

Since  $a > 2\gamma$ , this solution starts at  $y = c_1 + c_2$  at  $t = 0$  and continuously decreases; so, as  $t \rightarrow \infty$ ,  $y(t) \rightarrow 0$ .

(b)  $a^2 = 4b$  (critically damped): In this case we have one multiple root of order 2 ( $k_1 = 2$ ); therefore, the power of  $x$  can be zero or 1 ( $r_1 = 0, 1$ ). Thus, the general solution is

$$y(t) = c_1te^{-at/2} + c_0e^{-at/2}.$$

This solution starts at  $y(0) = c_0$  at  $t = 0$ , reaches a maximum (or minimum) at  $t = 2/a - c_0/c_1$ , and subsequently decays (grows) exponentially to zero.

(c)  $a^2 < 4b$  (underdamped): Once more, we have two distinct simple roots. The multiplicities are both one ( $k_1 = k_2 = 1$ ); therefore, the power of  $x$  for both solutions is zero

( $r_1 = r_2 = 0$ ). Let  $\omega \equiv \frac{1}{2}\sqrt{4b - a^2}$ . Then  $\lambda_1 = -a/2 + i\omega$  and  $\lambda_2 = \lambda_1^*$ . The roots are complex, and the most general solution is thus of the form

$$y(t) = e^{-at/2}(c_1 \cos \omega t + c_2 \sin \omega t) = Ae^{-at/2} \cos(\omega t + \alpha).$$

The solution is a harmonic variation with a decaying amplitude  $A \exp(-at/2)$ . Note that if  $a = 0$ , the amplitude does not decay. That is why  $a$  is called the **damping factor** (or the damping constant).

These equations describe either a mechanical system oscillating (with no external force) in a viscous (dissipative) fluid, or an electrical circuit consisting of a resistance  $R$ , an inductance  $L$ , and a capacitance  $C$ . For  $RLC$  circuits,  $a = R/L$  and  $b = 1/(LC)$ . Thus, the damping factor depends on the relative magnitudes of  $R$  and  $L$ . On the other hand, the frequency

$$\omega \equiv \sqrt{b - \left(\frac{a}{2}\right)^2} = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$$

depends on all three elements. In particular, for  $R \geq 2\sqrt{L/C}$  the circuit does not oscillate. ■

A physical system whose behavior in the absence of a driving force is described by a NOLDE will obey an inhomogeneous NOLDE in the presence of the driving force. This driving force is simply the inhomogeneous term of the NOLDE. The best way to solve such an inhomogeneous NOLDE in its most general form is by using Fourier transforms and Green's functions, as we will do in Chapter 20. For the particular, but important, case in which the inhomogeneous term is a product of polynomials and exponentials, the solution can be found in closed form.

**13.7.3. Theorem.** *The NOLDE  $\mathbf{L}[y] = e^{\lambda x} S(x)$ , where  $S(x)$  is a polynomial, has the particular solution  $e^{\lambda x} q(x)$ , where  $q(x)$  is also a polynomial. The degree of  $q(x)$  equals that of  $S(x)$  unless  $\lambda = \lambda_j$ , a root of the characteristic polynomial of  $\mathbf{L}$ , in which case the degree of  $q(x)$  exceeds that of  $S(x)$  by  $k_j$ , the multiplicity of  $\lambda_j$ .*

Once we know the form of the particular solution of the NOLDE, we can find the coefficients in the polynomial of the solution by substituting in the NOLDE and matching the powers on both sides.

**13.7.4. Example.** Let us find the most general solutions for the following two differential equations subject to the boundary conditions  $y(0) = 0$  and  $y'(0) = 1$ .

(a) The first DE we want to consider is

$$y'' + y = xe^x. \tag{13.42}$$

The characteristic polynomial is  $\lambda^2 + 1$ , whose roots are  $\lambda_1 = i$  and  $\lambda_2 = -i$ . Thus, a basis of solutions is  $\{\cos x, \sin x\}$ . To find the particular solution we note that  $\lambda$  (the coefficient of  $x$  in the exponential part of the inhomogeneous term) is 1, which is neither of the roots  $\lambda_1$  and  $\lambda_2$ . Thus, the particular solution is of the form  $q(x)e^x$ , where  $q(x) = Ax + B$  is

of degree 1 [same degree as that of  $S(x) = x$ ]. We now substitute  $u = (Ax + B)e^x$  in Equation (13.42) to obtain the relation

$$2Ax e^x + (2A + 2B)e^x = x e^x.$$

Matching the coefficients, we have

$$2A = 1 \quad \text{and} \quad 2A + 2B = 0 \Rightarrow A = \frac{1}{2} = -B.$$

Thus, the most general solution is

$$y = c_1 \cos x + c_2 \sin x + \frac{1}{2}(x - 1)e^x.$$

Imposing the given boundary conditions yields  $0 = y(0) = c_1 - \frac{1}{2}$  and  $1 = y'(0) = c_2$ . Thus,

$$y = \frac{1}{2} \cos x + \sin x + \frac{1}{2}(x - 1)e^x$$

is the unique solution.

(b) The next DE we want to consider is

$$y'' - y = x e^x. \quad (13.43)$$

Here  $p(\lambda) = \lambda^2 - 1$ , and the roots are  $\lambda_1 = 1$  and  $\lambda_2 = -1$ . A basis of solutions is  $\{e^x, e^{-x}\}$ . To find a particular solution, we note that  $S(x) = x$  and  $\lambda = 1 = \lambda_1$ . Theorem 13.7.3 then implies that  $q(x)$  must be of degree 2, because  $\lambda_1$  is a simple root, i.e.,  $k_1 = 1$ . We therefore try

$$q(x) = Ax^2 + Bx + C \Rightarrow u = (Ax^2 + Bx + C)e^x.$$

Taking the derivatives and substituting in Equation (13.43) yields two equations,

$$4A = 1 \quad \text{and} \quad A + B = 0,$$

whose solution is  $A = -B = \frac{1}{4}$ . Note that  $C$  is not determined, because  $Ce^x$  is a solution of the homogeneous DE corresponding to Equation (13.43), so when  $L$  is applied to  $u$ , it eliminates the term  $Ce^x$ . Another way of looking at the situation is to note that the most general solution to (13.43) is of the form

$$y = c_1 e^x + c_2 e^{-x} + \left(\frac{1}{4}x^2 - \frac{1}{4}x + C\right)e^x.$$

The term  $Ce^x$  could be absorbed in  $c_1 e^x$ . We therefore set  $C = 0$ , apply the boundary conditions, and find the unique solution

$$y = \frac{5}{4} \sinh x + \frac{1}{4}(x^2 - x)e^x. \quad \blacksquare$$

## 13.8 The WKB Method

In this section, we treat the somewhat specialized method—due to Wentzel, Kramers, and Brillouin—of obtaining an approximate solution to a particular type of second-order DE arising from the Schrödinger equation in one dimension. Suppose we are interested in finding approximate solutions of the DE

$$\frac{d^2y}{dx^2} + q(x)y = 0 \quad (13.44)$$

in which  $q$  varies “slowly” with respect to  $x$  in the sense discussed below. If  $q$  varies infinitely slowly, i.e., if it is a constant, the solution to Equation (13.44) is simply an imaginary exponential (or trigonometric). So, let us define  $\phi(x)$  by  $y = e^{i\phi(x)}$  and rewrite the DE as

$$(\phi')^2 + i\phi'' - q = 0. \quad (13.45)$$

Assuming that  $\phi''$  is small (compared to  $q$ ), so that  $y$  does not oscillate too rapidly, we can find an approximate solution to the DE:

$$\phi' = \pm\sqrt{q} \Rightarrow \phi = \pm \int \sqrt{q(x)} dx. \quad (13.46)$$

The condition of validity of our assumption is obtained by differentiating (13.46):

$$|\phi''| \approx \frac{1}{2} \left| \frac{q'}{\sqrt{q}} \right| \ll |q|.$$

It follows from Equation (13.46) and the definition of  $\phi$  that  $1/\sqrt{q}$  is approximately  $1/(2\pi)$  times one “wavelength” of the solution  $y$ . Therefore, the approximation is valid if the change in  $q$  in one wavelength is small compared to  $|q|$ .

The approximation can be improved by inserting the derivative of (13.46) in the DE and solving for a new  $\phi$ :

$$(\phi')^2 \approx q \pm \frac{i}{2} \frac{q'}{\sqrt{q}} \Rightarrow \phi' \approx \pm \left( q \pm \frac{i}{2} \frac{q'}{\sqrt{q}} \right)^{1/2},$$

or

$$\begin{aligned} \phi' &\approx \pm\sqrt{q} \left( 1 \pm \frac{i}{2} \frac{q'}{q^{3/2}} \right)^{1/2} = \pm\sqrt{q} \left( 1 \pm \frac{i}{4} \frac{q'}{q^{3/2}} \right) \\ &= \pm\sqrt{q} + \frac{i}{4} \frac{q'}{q} \Rightarrow \phi(x) \approx \pm \int \sqrt{q} dx + \frac{i}{4} \ln q. \end{aligned}$$

The two choices give rise to two different solutions, a linear combination of which gives the most general solution. Thus,

$$y \approx \frac{1}{\sqrt[4]{q(x)}} \left\{ c_1 \exp \left[ i \int \sqrt{q} dx \right] + c_2 \exp \left[ -i \int \sqrt{q} dx \right] \right\}. \quad (13.47)$$

Equation (13.47) gives an approximate solution to (13.44) in any region in which the condition of validity holds. The method fails if  $q$  changes too rapidly or if it is zero at some point of the region. The latter is a serious difficulty, since we often wish to join a solution in a region in which  $q(x) > 0$  to one in a region in which  $q(x) < 0$ . There is a general procedure for deriving the so-called *connection formulas* relating the constants  $c_1$  and  $c_2$  of the two solutions on either side of the point where  $q(x) = 0$ . We shall not go into the details of such a derivation, as it is not particularly illuminating.<sup>8</sup> We simply quote a particular result that is useful in applications.

Suppose that  $q$  passes through zero at  $x_0$ , is positive to the right of  $x_0$ , and satisfies the condition of validity in regions both to the right and to the left of  $x_0$ . Furthermore, assume that the solution of the DE decreases exponentially to the left of  $x_0$ . Under such conditions, the solution to the left will be of the form

$$\frac{1}{\sqrt[4]{-q(x)}} \exp \left[ - \int_x^{x_0} \sqrt{-q(x)} \, dx \right], \quad (13.48)$$

while to the right, we have

$$2 \frac{1}{\sqrt[4]{q(x)}} \cos \left[ \int_{x_0}^x \sqrt{q(x)} \, dx - \frac{\pi}{4} \right]. \quad (13.49)$$

A similar procedure gives connection formulas for the case where  $q$  is positive on the left and negative on the right of  $x_0$ .

**13.8.1. Example.** Consider the Schrödinger equation in one dimension

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} [E - V(x)]\psi = 0$$

where  $V(x)$  is a potential well meeting the horizontal line of constant  $E$  at  $x = a$  and  $x = b$ , so that

$$q(x) = \frac{2m}{\hbar^2} [E - V(x)] \begin{cases} > 0 & \text{if } a < x < b, \\ < 0 & \text{if } x < a \text{ or } x > b. \end{cases}$$

The solution that is bounded to the left of  $a$  must be exponentially decaying. Therefore, in the interval  $(a, b)$  the approximate solution, as given by Equation (13.49), is

$$\psi(x) \approx \frac{A}{(E - V)^{1/4}} \cos \left( \int_a^x \sqrt{\frac{2m}{\hbar^2} [E - V(x)]} \, dx - \frac{\pi}{4} \right),$$

where  $A$  is some arbitrary constant. The solution that is bounded to the right of  $b$  must also be exponentially decaying. Hence, the solution for  $a < x < b$  is

$$\psi(x) \approx \frac{B}{(E - V)^{1/4}} \cos \left( \int_x^b \sqrt{\frac{2m}{\hbar^2} [E - V(x)]} \, dx - \frac{\pi}{4} \right).$$

<sup>8</sup>The interested reader is referred to the book by Mathews and Walker, pp. 27–37.

Since these two expressions give the same function in the same region, they must be equal. Thus,  $A = B$ , and, more importantly,

$$\cos\left(\int_a^x \sqrt{\frac{2m}{\hbar^2}[E - V(x)]} dx - \frac{\pi}{4}\right) = \cos\left(\int_x^b \sqrt{\frac{2m}{\hbar^2}[E - V(x)]} dx - \frac{\pi}{4}\right),$$

or

$$\int_a^b \sqrt{2m[E - V(x)]} dx = (n + \frac{1}{2})\pi\hbar.$$

This is essentially the Bohr–Sommerfeld quantization condition of pre-1925 quantum mechanics. ■

### 13.8.1 Classical Limit of the Schrödinger Equation

As long as we are approximating solutions of second-order DEs that arise naturally from the Schrödinger equation, it is instructive to look at another approximation to the Schrödinger equation, its classical limit in which the Planck constant goes to zero.

The idea is to note that since  $\psi(\mathbf{r}, t)$  is a complex function, one can write it as

$$\psi(\mathbf{r}, t) = A(\mathbf{r}, t) \exp\left[\frac{i}{\hbar}S(\mathbf{r}, t)\right], \quad (13.50)$$

where  $A(\mathbf{r}, t)$  and  $S(\mathbf{r}, t)$  are real-valued functions. Substituting (13.50) in the Schrödinger equation and separating the real and the imaginary parts yields

$$\begin{aligned} \frac{\partial S}{\partial t} + \frac{\nabla S \cdot \nabla S}{2m} + V &= \frac{\hbar^2}{2m} \frac{\nabla^2 A}{A}, \\ m \frac{\partial A}{\partial t} + \nabla S \cdot \nabla A + \frac{A}{2} \nabla^2 S &= 0. \end{aligned} \quad (13.51)$$

These two equations are completely equivalent to the Schrödinger equation. The second equation has a direct physical interpretation. Define

$$\rho(\mathbf{r}, t) \equiv A^2(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2 \quad \text{and} \quad \mathbf{J}(\mathbf{r}, t) \equiv A^2(\mathbf{r}, t) \underbrace{\frac{\nabla S}{m}}_{\equiv \mathbf{v}} = \rho \mathbf{v}, \quad (13.52)$$

multiply the second equation in (13.51) by  $2A/m$ , and note that it then can be written as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0, \quad (13.53)$$

which is the continuity equation for probability. The fact that  $\mathbf{J}$  is indeed the probability current density is left for Problem 13.30.



The first equation of (13.51) gives an interesting result when  $\hbar \rightarrow 0$  because in this limit, the RHS of the equation will be zero, and we get

$$\frac{\partial S}{\partial t} + \frac{1}{2}mv^2 + V = 0.$$

Taking the gradient of this equation, we obtain

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right) m\mathbf{v} + \nabla V = 0,$$

which is the equation of motion of a classical fluid with velocity field  $\mathbf{v} = \nabla S/m$ . We thus have the following:

Schrödinger equation describes a classical statistical mixture when  $\hbar \rightarrow 0$ .

**13.8.2. Proposition.** *In the classical limit, the solution of the Schrödinger equation describes a fluid (statistical mixture) of noninteracting classical particles of mass  $m$  subject to the potential  $V(\mathbf{r})$ . The density and the current density of this fluid are, respectively, the probability density  $\rho = |\psi|^2$  and the probability current density  $\mathbf{J}$  of the quantum particle.*

## 13.9 Numerical Solutions of DEs

The majority of differential equations encountered in physics do not have known analytic solutions. One therefore resorts to numerical solutions. There are a variety of methods having various degrees of simplicity of use and accuracy. This section considers a few representatives applicable to the solution of ODEs. We make frequent use of techniques developed in Section 2.6. Therefore, the reader is urged to consult that section as needed.

Any normal differential equation of  $n$ th order,

$$\frac{d^n x}{dt^n} = F(x, \dot{x}, \dots, x^{(n-1)}, t),$$

can be reduced to a system of  $n$  first-order differential equations by defining  $x_1 = x$ ,  $x_2 = \dot{x}$ ,  $\dots$ ,  $x_n = x^{(n-1)}$ . This gives the system

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = x_3, \quad \dots, \quad \dot{x}_{n-1} = x_n, \quad \dot{x}_n = F(x_1, x_2, \dots, x_n; t).$$

We restrict ourselves to a FODE of the form  $\dot{x} = f(x, t)$  in which  $f$  is a well-behaved function of two real variables. At the end of the section, we briefly outline a technique for solving second-order DEs.

Two general types of problems are encountered in applications. An initial value problem (IVP) gives  $x(t)$  at an initial time  $t_0$  and asks for the value of  $x$  at other times. The second type, the boundary value problem (BVP), applies only to differential equations of higher order than first. A second-order BVP specifies the value of  $x(t)$  and/or  $\dot{x}(t)$  at one or more points and asks for  $x$  or  $\dot{x}$  at other values of  $t$ . We shall consider only IVPs.

### 13.9.1 Using the Backward Difference Operator

Let us consider the IVP

$$\dot{x} = f(x, t), \quad x(t_0) = x_0. \quad (13.54)$$

The problem is to find  $\{x_k = x(t_0 + kh)\}_{k=1}^N$ , given (13.54).

Let us begin by integrating (13.54) between  $t_n$  and  $t_n + h$ :

$$x(t_n + h) - x(t_n) = \int_{t_n}^{t_n+h} \dot{x}(t) dt.$$

Changing the variable of integration to  $s = (t - t_n)/h$  and using the shift operator  $\mathbf{E}$  introduced in Section 2.6 yields

$$x_{n+1} - x_n = h \int_0^1 \dot{x}(t_n + sh) ds = h \int_0^1 [\mathbf{E}^s \dot{x}(t_n)] ds. \quad (13.55)$$

Since a typical situation involves calculating  $x_{n+1}$  from the values of  $x(t)$  and  $\dot{x}(t)$  at preceding steps, we want an expression in which the RHS of Equation (13.55) contains such preceding terms. This suggests expressing  $\mathbf{E}$  in terms of the backward difference operator. It will also be useful to replace the lower limit of integration to  $-p$ , where  $p$  is a number to be chosen later for convenience. Thus, Equation (13.55) becomes

$$\begin{aligned} x_{n+1} &= x_{n-p} + h \left[ \int_{-p}^1 (1 - \nabla)^{-s} ds \right] \dot{x}_n \\ &= x_{n-p} + h \left[ \int_{-p}^1 \sum_{k=0}^{\infty} \frac{\Gamma(-s+1) ds}{k! \Gamma(-s-k+1)} (-\nabla)^k \right] \dot{x}_n \\ &= x_{n-p} + h \left( \sum_{k=0}^{\infty} a_k^{(p)} \nabla^k \right) \dot{x}_n, \end{aligned} \quad (13.56)$$

where

$$a_k^{(p)} \equiv \frac{(-1)^k}{k!} \int_{-p}^1 \frac{\Gamma(-s+1) ds}{\Gamma(-s-k+1)} = \frac{1}{k!} \int_{-p}^1 s(s+1) \cdots (s+k-1) ds. \quad (13.57)$$

Keeping the first few terms for  $p = 0$ , we obtain the useful formula

$$x_{n+1} \approx x_n + h \left( 1 + \frac{\nabla}{2} + \frac{5\nabla^2}{12} + \frac{3\nabla^3}{8} + \frac{251\nabla^4}{720} + \frac{95\nabla^5}{288} + \cdots \right) \dot{x}_n. \quad (13.58)$$

Due to the presence of  $\nabla$  in Equation (13.58), finding the value of  $x(t)$  at  $t_{n+1}$  requires a knowledge of  $x(t)$  and  $\dot{x}(t)$  at points  $t_0, t_1, \dots, t_n$ . Because of this,

formulas of open and closed type

Equation (13.58) is called a formula of **open type**. In contrast, in formulas of **closed type**, the RHS contains values at  $t_{n+1}$  as well. We can obtain a formula of closed type by changing  $E^s \dot{x}_n$  to its equivalent form,  $E^{s-1} \dot{x}_{n+1}$ . The result is

$$x_{n+1} = x_{n-p} + h \sum_{k=0}^{\infty} b_k^{(p)} \nabla^k \dot{x}_{n+1}, \quad (13.59)$$

where

$$b_k^{(p)} \equiv \frac{(-1)^k}{k!} \int_{-p}^1 \frac{\Gamma(-s+2) ds}{\Gamma(-s-k+2)}.$$

Keeping the first few terms for  $p = 0$ , we obtain

$$x_{n+1} \approx x_n + h \left( 1 - \frac{\nabla}{2} - \frac{\nabla^2}{12} - \frac{\nabla^3}{24} - \frac{19\nabla^4}{720} - \frac{3\nabla^5}{160} - \dots \right) \dot{x}_{n+1}, \quad (13.60)$$

which involves evaluation at  $t_{n+1}$  on the RHS.

For  $p = 1$  ( $p = 3$ ), Equation (13.56) [(13.59)] results in an expansion in powers of  $\nabla$  in which the coefficient of  $\nabla^p$  ( $\nabla^{p+2}$ ) is zero. Thus, retaining terms up to the  $(p-1)$ st [ $(p+1)$ st] power of  $\nabla$  automatically gives us an accuracy of  $h^p$  ( $h^{p+2}$ ). This is the advantage of using nonzero values of  $p$  and the reason we considered such cases. The reason for the use of formulas of the closed type is the smallness of the error involved. All the formulas derived in this section involve powers of  $\nabla$  operating on  $\dot{x}_n$  or  $\dot{x}_{n+1}$ . This means that to find  $x_{n+1}$ , we must know the values of  $\dot{x}_k$  for  $k \leq n+1$ . However,  $\dot{x} = f(x, t)$  or  $\dot{x}_k = f(x_k, t_k)$  implies that knowledge of  $\dot{x}_k$  requires knowledge of  $x_k$ . Therefore, to find  $x_{n+1}$ , we must know not only the values of  $\dot{x}$  but also the values of  $x(t)$  at  $t_k$  for  $k \leq n+1$ . In particular, we cannot start with  $n = 0$  because we would get negative indices for  $x$  due to the high powers of  $\nabla$ . This means that the first few values of  $x_k$  must be obtained using a different method. One common method of *starting the solution* is to use a Taylor series expansion:

$$x_k = x(t_0 + kh) = x_0 + h\dot{x}_0 k + \frac{h^2 \ddot{x}_0}{2} k^2 + \dots, \quad (13.61)$$

where

$$\dot{x}_0 = f(x_0, t_0), \quad \ddot{x}_0 = \left( \frac{\partial f}{\partial x} \Big|_{x_0, t_0} \right) \dot{x}_0 + \frac{\partial f}{\partial t} \Big|_{x_0, t_0}, \quad \dots$$

For the general case, it is clear that the derivatives required for the RHS of Equation (13.61) involve very complicated expressions. The following example illustrates the procedure for a specific case.

**13.9.1. Example.** Let us solve the IVP  $\dot{x} + x + e^t x^2 = 0$  with  $x(0) = 1$ . We can obtain a Taylor series expansion for  $x$  by noting that

$$\begin{aligned} \dot{x}_0 &= -x_0 - x_0^2, & \ddot{x}_0 &= \dot{x}(0) = -\dot{x}_0 - 2x_0\dot{x}_0 - x_0^2, \\ \ddot{x}_0 &= -\dot{x}_0 - 2x_0^2 - 2x_0\dot{x}_0 - 4x_0\dot{x}_0 - x_0^2. \end{aligned}$$

Continuing in this way, we can obtain derivatives of all orders. Substituting  $x_0 = 1$  and keeping terms up to the fifth order, we obtain

$$\dot{x}_0 = -2, \quad \ddot{x}_0 = 5, \quad \dddot{x}_0 = -16, \quad \left. \frac{d^4 x}{dt^4} \right|_{t=0} = 65, \quad \left. \frac{d^5 x}{dt^5} \right|_{t=0} = -326.$$

Substituting these values in a Taylor series expansion with  $h = 0.1$  yields

$$x_k = 1 - 0.2k + 0.025k^2 - 0.0027k^3 + (2.7 \times 10^{-4})k^4 - (2.7 \times 10^{-5})k^5 + \dots$$

Thus,  $x_1 = 0.82254$ ,  $x_2 = 0.68186$ , and  $x_3 = 0.56741$ . The corresponding values of  $\dot{x}$  can be calculated using the DE. We simply quote the result:  $\dot{x}_1 = -1.57026$ ,  $\dot{x}_2 = -1.24973$ ,  $\dot{x}_3 = -1.00200$ . ■

Once the starting values are obtained, either a formula of open type or one of closed type is used to find the next  $x$  value. Only formulas of open type will be discussed here. However, as mentioned earlier, the accuracy of closed-type formulas is better. The price one pays for having  $x_{n+1}$  on the RHS is that using closed-type formulas requires *estimating*  $x_{n+1}$ . This estimate is then substituted in the RHS, and an improved estimate is found. The process is continued until no further improvement in the estimate is achieved.

The use of open-type formulas involves simple substitution of the known quantities  $x_0, x_1, \dots, x_n$  on the RHS to obtain  $x_{n+1}$ . The master equation (for  $p = 0$ ) is (13.58). The number of powers of  $\nabla$  that are retained gives rise to different methods. For instance, when no power is retained, the method is called **Euler's method** for which we use  $x_{n+1} \approx x_n + h\dot{x}_n$ . A more commonly used method is **Adam's method**, for which all powers of  $\nabla$  up to and including the third are retained. We then have

$$x_{n+1} \approx x_n + h\left(1 + \frac{1}{2}\nabla + \frac{5}{12}\nabla^2 + \frac{3}{8}\nabla^3\right)\dot{x}_n,$$

or, in terms of values of  $\dot{x}$ ,

$$x_{n+1} \approx x_n + \frac{h}{24}(55\dot{x}_n - 59\dot{x}_{n-1} + 37\dot{x}_{n-2} - 9\dot{x}_{n-3}). \tag{13.62}$$

Recall that  $\dot{x}_k = f(x_k, t_k)$ . Thus, if we know the values  $x_n, x_{n-1}, x_{n-2}$ , and  $x_{n-3}$ , we can obtain  $x_{n+1}$ .

**13.9.2. Example.** Knowing  $\dot{x}_0, \dot{x}_1, \dot{x}_2$ , and  $\dot{x}_3$ , we can use Equation (13.62) to calculate  $x_4$  for Example 13.9.1:

$$x_4 \approx x_3 + \frac{0.1}{24}(55\dot{x}_3 - 59\dot{x}_2 + 37\dot{x}_1 - 9\dot{x}_0) = 0.47793.$$

With  $x_4$  at our disposal, we can evaluate  $\dot{x}_4 = -x_4 - x_4^2 e^{t_4}$ , and substitute it in

$$x_5 \approx x_4 + \frac{0.1}{24}(55\dot{x}_4 - 59\ddot{x}_3 + 37\ddot{x}_2 - 9\ddot{x}_1)$$

to find  $x_5$ , and so on. ■

A crucial fact about such methods is that every value obtained is in error by some amount, and using such values to obtain new values propagates the error. Thus, error can accumulate rapidly and make approximations worse at each step. Discussion of error propagation and error analysis—topics that we have not, and shall not, cover—is common in the literature (see, for example, [Hild 87, pp. 267–268]).

### 13.9.2 The Runge–Kutta Method

The FODE of Equation (13.54) leads to a unique Taylor series,

$$x(t_0 + h) = x_0 + h\dot{x}_0 + \frac{h^2}{2}\ddot{x}_0 + \dots,$$

where  $\dot{x}_0$ ,  $\ddot{x}_0$ , and all the rest of the derivatives can be evaluated by differentiating  $\dot{x} = f(x, t)$ . Thus, theoretically, the Taylor series gives the solution (for  $t_0 + h$ ; but  $t_0 + 2h$ ,  $t_0 + 3h$ , and so on can be obtained similarly). However, in practice, the Taylor series converges slowly, and the accuracy involved is not high. Thus one resorts to other methods of solution such as described earlier.

Runge–Kutta method

Another method, known as the **Runge–Kutta method**, replaces the Taylor series

$$x_{n+1} = x_n + h\dot{x}_n + \frac{h^2}{2}\ddot{x}_n + \frac{h^3}{3!}\ddot{x}_n + \dots \quad (13.63)$$

with

$$x_{n+1} = x_n + h \left[ \alpha_0 f(x_n, t_n) + \sum_{j=1}^p \alpha_j f(x_n + b_j h, t_n + \mu_j h) \right], \quad (13.64)$$

where  $\alpha_0$  and  $\{\alpha_j, b_j, \mu_j\}_{j=1}^p$  are constants chosen such that if the RHS of (13.64) were expanded in powers of the spacing  $h$ , the coefficients of a certain number of the leading terms would agree with the corresponding expansion coefficients of the RHS of (13.63). It is customary to express the  $b$ 's as linear combinations of preceding values of  $f$ :

$$hb_i = \sum_{r=0}^{i-1} \lambda_{ir} k_r, \quad i = 1, 2, \dots, p.$$

The  $k_r$  are recursively defined as

$$k_0 = hf(x_n, t_n), \quad k_r = hf(x_n + b_r h, t_n + \mu_r h).$$

Then Equation (13.64) gives  $x_{n+1} = x_n + \sum_{r=0}^p \alpha_r k_r$ . The (nontrivial) task now is to determine the parameters  $\alpha_r$ ,  $\mu_r$ , and  $\lambda_{ij}$ .

**Carle David Tolmé Runge** (1856–1927), after returning from a six-month vacation in Italy, enrolled at the University of Munich to study literature. However, after six weeks of the course he changed to mathematics and physics.

Runge attended courses with Max Planck, and they became close friends. In 1877 both went to Berlin, but Runge turned to pure mathematics after attending Weierstrass's lectures. His doctoral dissertation (1880) dealt with differential geometry.

After taking a secondary-school teachers certification test, he returned to Berlin, where he was influenced by Kronecker. Runge then worked on a procedure for the numerical solution of algebraic equations in which the roots were expressed as infinite series of rational functions of the coefficients. In the area of numerical analysis, he is credited with an efficient method of solving differential equations numerically, work he did with Martin Kutta.

Runge published little at that stage, but after visiting Mittag-Leffler in Stockholm in September 1884 he produced a large number of papers in Mittag-Leffler's journal *Acta mathematica*. In 1886, Runge obtained a chair at Hanover and remained there for 18 years. Within a year Runge had moved away from pure mathematics to study the wavelengths of the spectral lines of elements other than hydrogen. He did a great deal of experimental work and published a great quantity of results, including a separation of the spectral lines of helium in two spectral series.

In 1904 Klein persuaded Göttingen to offer Runge a chair of applied mathematics, a post that Runge held until he retired in 1925.

Runge was always a fit and active man, and on his 70th birthday he entertained his grandchildren by doing handstands. However, a few months later he had a heart attack and died.



In general, the determination of these constants is extremely tedious. Let us consider the very simple case where  $p = 1$ , and let  $\lambda \equiv \lambda_{01}$  and  $\mu \equiv \mu_1$ . Then we obtain

$$x_{n+1} = x_n + \alpha_0 k_0 + \alpha_1 k_1, \tag{13.65}$$

where  $k_0 = hf(x_n, t_n)$  and  $k_1 = hf(x_n + \lambda k_0, t_n + \mu h)$ .

Taylor-expanding  $k_1$ , a function of two variables, gives<sup>9</sup>

$$k_1 = hf + h^2(\mu f_t + \lambda f f_x) + \frac{h^3}{2}(\mu^2 f_{tt} + 2\lambda \mu f f_{xt} + \lambda^2 f^2 f_{xx}) + O(h^4),$$

<sup>9</sup>The symbol  $O(h^m)$  means that all terms of order  $h^m$  and higher have been neglected.

where  $f_t \equiv \partial f / \partial t$ , etc. Substituting this in the first equation of (13.65), we get

$$x_{n+1} = x_n + h(\alpha_0 + \alpha_1)f + h^2\alpha_1(\mu f_t + \lambda f f_x) + \frac{h^3}{2}\alpha_1(\mu^2 f_{tt} + 2\lambda\mu f f_{xt} + \lambda^2 f^2 f_{xx}) + O(h^4). \tag{13.66}$$

On the other hand, with

$$\begin{aligned} \dot{x} &= f, & \ddot{x} &= \frac{df}{dt} = \frac{\partial f}{\partial t} \frac{dx}{dt} + \frac{\partial f}{\partial t} = \dot{x} f_x + f_t = f f_x + f_t, \\ \ddot{x} &= f_{tt} + 2f f_{xt} + f^2 f_{xx} + f_x(f f_x + f_t), \end{aligned}$$

Equation (13.63) gives

$$x_{n+1} = x_n + hf + \frac{h^2}{2}(f f_x + f_t) + \frac{h^3}{6}[f_{tt} + 2f f_{xt} + f^2 f_{xx} + f_x(f f_x + f_t)] + O(h^4). \tag{13.67}$$

If we demand that (13.66) and (13.67) agree up to the  $h^2$  term (we cannot demand agreement for  $h^3$  or higher because of overspecification), then we must have  $\alpha_0 + \alpha_1 = 1$ ,  $\alpha_1\mu = \frac{1}{2}$ ,  $\alpha_1\lambda = \frac{1}{2}$ . There are only three equations for four unknowns. Therefore, there will be an arbitrary parameter  $\beta$  in terms of which the unknowns can be written:

$$\alpha_0 = 1 - \beta, \quad \alpha_1 = \beta, \quad \mu = \frac{1}{\beta}, \quad \lambda = \frac{1}{2\beta}.$$

Substituting these values in Equation (13.65) gives

$$x_{n+1} = x_n + h \left[ (1 - \beta)f(x_n, t_n) + \beta f \left( x_n + \frac{hf}{2\beta}, t_n + \frac{h}{2\beta} \right) \right] + O(h^3).$$

This formula becomes useful if we let  $\beta = \frac{1}{2}$ . Then  $t_n + h/(2\beta) = t_n + h = t_{n+1}$ , which makes evaluation of the second term in square brackets convenient. For  $\beta = \frac{1}{2}$ , we have

$$x_{n+1} = x_n + \frac{h}{2}[f(x_n, t_n) + f(x_n + hf, t_{n+1})] + O(h^3). \tag{13.68}$$

What is nice about this equation is that it needs no starting up! We can plug in the known quantities  $t_n$ ,  $t_{n+1}$ , and  $x_n$  on the RHS and find  $x_{n+1}$  starting with  $n = 0$ . However, the result is not very accurate, and we cannot make it any more accurate by demanding agreement for higher powers of  $h$ , because, as mentioned earlier, such a demand overspecifies the unknowns.

**Martin Wilhelm Kutta** (1867–1944) lost his parents when he was still a child, and together with his brother went to his uncle in Breslau to go to *gymnasium*. He attended the University of Breslau from 1885–1890, and the University of Munich from 1891–1894 concentrating mainly on mathematics, but he was also interested in languages, music, and art. Although he completed the certification for teaching mathematics and physics in 1894, he did not start teaching immediately. Instead, he assisted von Dyck at the Technische Hochschule München until 1897 (and then again from 1899 to 1903).

From 1898 to 1899 he studied at Cambridge, and a year later, he finished his Ph.D. at the University of Munich. In 1902, he completed his habilitation in pure and applied mathematics at the Technische Hochschule München, where he became professor of applied mathematics five years later. In 1909 he accepted an offer from University of Vienna, but a year later he went to the Technische Hochschule Aachen as a professor. From 1912 until his retirement in 1935 he worked at the Technische Hochschule Stuttgart.

Kutta's name is well known not only to physicists, applied mathematicians, and engineers, but also to specialists in aerospace science and fluid mechanics. The first group use the *Runge–Kutta* method, developed at the beginning of the twentieth century to obtain numerical solutions to ordinary differential equations. The second group use the Kutta–Zhukovskii formula for the theoretical description of the buoyancy of a body immersed in a nonturbulent moving fluid. Kutta's work on the application of conformal mapping to the study of airplane wings was later applied to the flight of birds, and further developed by L. Prandtl in the theory of wings.

Kutta obtained the motivation for his first scientific publication from Boltzmann and others (including a historian of mathematics) when working on the theoretical determination of the heat exchanged between two concentric cylinders kept at constant temperatures. By applying the conformal mapping technique, Kutta managed to obtain numerical values for the heat conductivity of air that agreed well with the experimental values of the time.

Three of Kutta's publications dealt with the history of mathematics, for which he profited greatly because of his knowledge of the Arabic language.

One of the most important tasks of applied mathematics is to approximate numerically the initial value problem of ODEs whose solutions cannot be found in closed form. After Euler (1770) had already expressed the basic idea, Runge (1895) and Heun (1900) wrote down the appropriate formulas. Kutta's contribution was to considerably increase the accuracy, and allow for a larger selection of the parameters involved. After accepting a professorship in Stuttgart in 1912, Kutta devoted all his time to teaching. He was very much in demand as a teacher, and it is said that his lectures were so good that even engineering students took an interest in mathematics.

(Taken from W. Schulz, "Martin Wilhelm Kutta," *Neue Deutsche Biographie* 13, Berlin, (1952–) 348–350.)

Formulas that give more accurate results can be obtained by retaining terms beyond  $p = 1$ . Thus, for  $p = 2$ , if we write  $x_{n+1} = x_n + \sum_{r=0}^2 \alpha_r k_r$ , there will be eight unknowns (three  $\alpha$ 's, three  $\lambda_{ij}$ 's, and two  $\mu$ 's), and the demand for agreement between the Taylor expansion and the expansion of  $f$  up to  $h^3$  will yield only six equations. Therefore, there will be two arbitrary parameters whose specification



results in various formulas. The details of this kind of algebraic derivation are very messy, so we will merely consider two specific formulas. One such formula, due to Kutta, is

$$x_{n+1} = x_n + \frac{1}{6}(k_0 + 4k_1 + k_2) + O(h^4), \quad (13.69)$$

where

$$\begin{aligned} k_0 &= hf(x_n, t_n), & k_1 &= hf(x_n + \frac{1}{2}k_0, t_n + \frac{1}{2}h), \\ k_2 &= hf(x_n + 2k_1 - k_0, t_n + h). \end{aligned}$$

A second formula, due to Heun, has the form

$$x_{n+1} = x_n + \frac{1}{4}(k_0 + 3k_2) + O(h^4),$$

where

$$\begin{aligned} k_0 &= hf(x_n, t_n), & k_1 &= hf(x_n + \frac{1}{3}k_0, t_n + \frac{1}{3}h), \\ k_2 &= hf(x_n + \frac{2}{3}k_1 - k_0, t_n + \frac{2}{3}h). \end{aligned}$$

These two formulas are of about the same order of accuracy.

**13.9.3. Example.** Let us solve the DE of Example 13.9.1 using the Runge–Kutta method. With  $t_0 = 0$ ,  $x_0 = 1$ ,  $h = 0.1$ , and  $n = 0$ , Equation (13.69) gives  $k_0 = -0.2$ ,  $k_1 = -0.17515$ ,  $k_2 = -0.16476$ , so that

$$x_1 = 1 + \frac{1}{6}(-0.2 + 4(-0.17515) - 0.16476) = 0.82244.$$

This  $x_1$ ,  $h = 0.1$ , and  $t_1 = t_0 + h = 0.1$  yield the following new values:  $k_0 = -0.15700$ ,  $k_1 = -0.13870$ ,  $k_2 = -0.13040$ , which in turn give

$$x_2 = 0.82244 + \frac{1}{6}[-0.15700 - 4(0.13870) - 0.13040] = 0.68207.$$

We similarly obtain  $x_3 = 0.56964$  and  $x_4 = 0.47858$ . On the other hand, solving the FODE analytically gives the exact result  $x(t) = e^{-t}/(1+t)$ .

Table 13.1 compares the values obtained here, those obtained using Adam's method, and the exact values to five decimal places. It is clear that the Runge–Kutta method is more accurate than the methods discussed earlier. ■

The accuracy of the Runge–Kutta method and the fact that it requires no startup procedure make it one of the most popular methods for solving differential equations. The Runge–Kutta method can be made more accurate by using higher values of  $p$ . For instance, a formula that is used for  $p = 3$  is

$$x_{n+1} = x_n + \frac{1}{6}(k_0 + 2k_1 + 2k_2 + k_3) + O(h^5), \quad (13.70)$$

where

$$\begin{aligned} k_0 &= hf(x_n, t_n), & k_1 &= hf(x_n + \frac{1}{2}k_0, t_n + \frac{1}{2}h), \\ k_2 &= hf(x_n + \frac{1}{2}k_1, t_n + \frac{1}{2}h) & k_3 &= hf(x_n + k_2, t_n + h). \end{aligned}$$

$t$	Analytical	Runge-Kutta	Adam's method
0.	1	1	1
0.1	0.82258	0.82244	0.82254
0.2	0.68228	0.68207	0.68186
0.3	0.56986	0.56964	0.56741
0.4	0.47880	0.47858	0.47793

**Table 13.1** Solutions to the differential equation of Example 13.9.1 obtained in three different ways.

### 13.9.3 Higher-Order Equations

Any  $n$ th-order differential equation is equivalent to  $n$  first-order differential equations in  $n + 1$  variables. Thus, for instance, the most general SODE,  $F(\ddot{x}, \dot{x}, x, t) = 0$ , can be reduced to two FODEs by solving for  $\ddot{x}$  to obtain  $\ddot{x} = G(\dot{x}, x, t)$ , and defining  $\dot{x} = u$  to get the system of equations

$$\dot{u} = G(u, x, t), \quad \dot{x} = u.$$

These two equations are completely equivalent to the original SODE. Thus, it is appropriate to discuss numerical solutions of systems of FODEs in several variables. The discussion here will be limited to systems consisting of two equations. The generalization to several equations is not difficult.

Consider the IVP of the following system of equations

$$\dot{x} = f(x, u, t), \quad x(t_0) = x_0; \quad \dot{u} = g(x, u, t), \quad u(t_0) = u_0. \quad (13.71)$$

Using an obvious generalization of Equation (13.70), we can write

$$\begin{aligned} x_{n+1} &= x_n + \frac{1}{6}(k_0 + 2k_1 + 2k_2 + k_3) + O(h^5), \\ u_{n+1} &= u_n + \frac{1}{6}(m_0 + 2m_1 + 2m_2 + m_3) + O(h^5), \end{aligned} \quad (13.72)$$

where

$$\begin{aligned} k_0 &= hf(x_n, u_n, t_n), & k_1 &= hf(x_n + \frac{1}{2}k_0, u_n + \frac{1}{2}m_0, t_n + \frac{1}{2}h), \\ k_2 &= hf(x_n + \frac{1}{2}k_1, u_n + \frac{1}{2}m_1, t_n + \frac{1}{2}h), \\ k_3 &= hf(x_n + k_2, u_n + m_2, t_n + h), \end{aligned}$$

and

$$\begin{aligned} m_0 &= hg(x_n, u_n, t_n), & m_1 &= hg(x_n + \frac{1}{2}k_0, u_n + \frac{1}{2}m_0, t_n + \frac{1}{2}h), \\ m_2 &= hg(x_n + \frac{1}{2}k_1, u_n + \frac{1}{2}m_1, t_n + \frac{1}{2}h), \\ m_3 &= hg(x_n + k_2, u_n + m_2, t_n + h). \end{aligned}$$

These formulas are more general than needed for a SODE, since, as mentioned above, such a SODE is equivalent to the simpler system in which  $f(x, u, t) \equiv u$ . Therefore, Equation (13.72) specializes to

$$\begin{aligned} k_0 &= hu_n = h\dot{x}_n, & k_1 &= h(u_n + \frac{1}{2}m_0) = h\dot{x}_n + \frac{1}{2}hm_0, \\ k_2 &= h\dot{x}_n + \frac{1}{2}hm_1, & k_3 &= h\dot{x}_n + hm_2, \end{aligned}$$

and

$$\begin{aligned} x_{n+1} &= x_n + h\dot{x}_n + \frac{1}{6}h(m_0 + m_1 + m_2) + O(h^5), \\ \dot{x}_{n+1} &= \dot{x}_n + \frac{1}{6}(m_0 + 2m_1 + 2m_2 + m_3) + O(h^5), \end{aligned} \tag{13.73}$$

where

$$\begin{aligned} m_0 &= hg(x_n, \dot{x}_n, t_n), & m_1 &= hg(x_n + \frac{1}{2}h\dot{x}_n, \dot{x}_n + \frac{1}{2}m_0, t_n + \frac{1}{2}h), \\ m_2 &= hg(x_n + \frac{1}{2}h\dot{x}_n + \frac{1}{4}hm_0, \dot{x}_n + \frac{1}{2}m_1, t_n + \frac{1}{2}h), \\ m_3 &= hg(x_n + h\dot{x}_n + \frac{1}{2}hm_1, \dot{x}_n + m_2, t_n + h). \end{aligned}$$

**13.9.4. Example.** The IVP  $\ddot{x} + x = 0$ ,  $x(0) = 0$ ,  $\dot{x}(0) = 1$  clearly has the analytic solution  $x(t) = \sin t$ . Nevertheless, let us use Equation (13.73) to illustrate the Runge-Kutta method and compare the result with the exact solution.

For this problem  $g(x, \dot{x}, t) = -x$ . Therefore, we can easily calculate the  $m$ 's:

$$\begin{aligned} m_0 &= -hx_n, & m_1 &= -h(x_n + \frac{1}{2}h\dot{x}_n), \\ m_2 &= -h(x_n + \frac{1}{2}h\dot{x}_n - \frac{1}{4}h^2\ddot{x}_n), \\ m_3 &= -h[x_n + h\dot{x}_n - \frac{1}{2}h^2(\ddot{x}_n + \frac{1}{2}h\ddot{\dot{x}}_n)]. \end{aligned}$$

These lead to the following expressions for  $x_{n+1}$  and  $\dot{x}_{n+1}$ :

$$\begin{aligned} x_{n+1} &= x_n + h\dot{x}_n - \frac{1}{6}h^2(3x_n + h\dot{x}_n - \frac{1}{4}h^2\ddot{x}_n), \\ \dot{x}_{n+1} &= \dot{x}_n - \frac{1}{6}h[6x_n + 3h\dot{x}_n - h^2(\ddot{x}_n + \frac{1}{4}h\ddot{\dot{x}}_n)]. \end{aligned}$$

Starting with  $x_0 = 0$  and  $\dot{x}_0 = 1$ , we can generate  $x_1, x_2$ , and so on by using the last two equations successively. The results for 10 values of  $x$  with  $h = 0.1$  are given to five significant figures in Table 13.2. Note that up to  $x_5$  there is complete agreement with the exact result. ■

The Runge-Kutta method lends itself readily to use in computer programs. Because Equation (13.73) does not require any startups, it can be used directly to generate solutions to any IVP involving a SODE.

Another, more direct, method of solving higher-order differential equations is to substitute  $\mathbf{D} = -(1/h)\ln(\mathbf{1} - \nabla)$  for the derivative operator in the differential equation, expand in terms of  $\nabla$ , and keep an appropriate number of terms. Problem 13.33 illustrates this point for a linear SODE.

$t$	Runge–Kutta	$\sin t$
0.1	0.09983	0.09983
0.2	0.19867	0.19867
0.3	0.29552	0.29552
0.4	0.38942	0.38942
0.5	0.47943	0.47943
0.6	0.56466	0.56466
0.7	0.64425	0.64422
0.8	0.71741	0.71736
0.9	0.78342	0.78333
1.0	0.84161	0.84147

**Table 13.2** Comparison of the Runge–Kutta and exact solutions to the second order DE of Example 13.9.4.

## 13.10 Problems

**13.1.** Let  $u(x)$  be a differentiable function satisfying the *differential inequality*  $u'(x) \leq Ku(x)$  for  $x \in [a, b]$ , where  $K$  is a constant. Show that  $u(x) \leq u(a)e^{K(x-a)}$ . Hint: Multiply both sides of the inequality by  $e^{-Kx}$ , and show that the result can be written as the derivative of a nonincreasing function. Then use the fact that  $a \leq x$  to get the final result.

**13.2.** Prove Proposition 13.4.2.

**13.3.** Let  $f$  and  $g$  be two differentiable functions that are linearly dependent. Show that their Wronskian vanishes.

**13.4.** Show that if  $(f_1, f_1')$  and  $(f_2, f_2')$  are linearly dependent at one point, then  $f_1$  and  $f_2$  are linearly dependent at all  $x \in [a, b]$ . Here  $f_1$  and  $f_2$  are solutions of the DE of (13.12). Hint: Derive the identity

$$W(f_1, f_2; x_2) = W(f_1, f_2; x_1) \exp \left\{ - \int_{x_1}^{x_2} p(t) dt \right\}.$$

**13.5.** Show that the solutions to the SOLDE  $y'' + q(x)y = 0$  have a constant Wronskian.

**13.6.** Find (in terms of an integral)  $G_n(x)$ , the linearly independent “partner” of the Hermite polynomial  $H_n(x)$ . Specialize this to  $n = 0, 1$ . Is it possible to find  $G_0(x)$  and  $G_1(x)$  in terms of elementary functions?

**13.7.** Let  $f_1, f_2,$  and  $f_3$  be any three solutions of  $y'' + py' + qy = 0$ . Show that the (generalized  $3 \times 3$ ) Wronskian of these solutions is zero. Thus, any three solutions of the HSOLDE are linearly dependent.

13.8. For the HSOLDE  $y'' + py' + qy = 0$ , show that

$$p = -\frac{f_1 f_2'' - f_2 f_1''}{W(f_1, f_2)} \quad \text{and} \quad q = \frac{f_1' f_2'' - f_2' f_1''}{W(f_1, f_2)}.$$

Thus, knowing two solutions of an HSOLDE allows us to reconstruct the DE.

13.9. Let  $f_1$ ,  $f_2$ , and  $f_3$  be three solutions of the third-order linear differential equation  $y''' + p_2(x)y'' + p_1(x)y' + p_0(x)y = 0$ . Derive a FODE satisfied by the (generalized  $3 \times 3$ ) Wronskian of these solutions.

13.10. Prove Corollary 13.4.13. Hint: Consider the solution  $u = 1$  of the DE  $u'' = 0$  and apply Theorem 13.4.11.

13.11. Show that the adjoint of  $\mathbf{M}$  given in Equation (13.21) is the original  $\mathbf{L}$ .

13.12. Show that if  $u(x)$  and  $v(x)$  are solutions of the self-adjoint DE  $(pu')' + qu = 0$ , then **Abel's identity**,  $p(uv' - vu') = \text{constant}$ , holds.

13.13. Reduce each DE to self-adjoint form.

$$(a) \quad x^2 y'' + x y' + y = 0.$$

$$(b) \quad y'' + y' \tan x = 0.$$

13.14. Reduce the self-adjoint DE  $(py')' + qy = 0$  to  $u'' + S(x)u = 0$  by an appropriate change of the dependent variable. What is  $S(x)$ ? Apply this reduction to the Legendre DE for  $P_n(x)$ , and show that

$$S(x) = \frac{1 + n(n+1) - n(n+1)x^2}{(1-x^2)^2}.$$

Now use this result to show that every solution of the Legendre equation has at least  $(2n+1)/\pi$  zeros on  $(-1, +1)$ .

13.15. Substitute  $v = y'/y$  in the *homogeneous* SOLDE

$$y'' + p(x)y' + q(x)y = 0$$

and:

Riccati equation (a) Show that it turns into  $v' + v^2 + p(x)v + q(x) = 0$ , which is a first-order *nonlinear* equation called the **Riccati equation**. Would the same substitution work if the DE were inhomogeneous?

(b) Show that by an appropriate transformation, the Riccati equation can be directly cast in the form  $u' + u^2 + S(x) = 0$ .

13.16. For the function  $S(x)$  defined in Example 13.6.1, let  $S^{-1}(x)$  be the inverse, i.e.,  $S^{-1}(S(x)) = x$ . Show that

$$\frac{d}{dx}[S^{-1}(x)] = \frac{1}{\sqrt{1-x^2}},$$

and given that  $S^{-1}(0) = 0$ , conclude that

$$S^{-1}(x) = \int_0^x \frac{dt}{\sqrt{1-t^2}}.$$

**13.17.** Define  $\sinh x$  and  $\cosh x$  as the solutions of  $y'' = y$  satisfying the boundary conditions  $y(0) = 0$ ,  $y'(0) = 1$  and  $y(0) = 1$ ,  $y'(0) = 0$ , respectively. Using Example 13.6.1 as a guide, show that

$$(a) \cosh^2 x - \sinh^2 x = 1. \quad (b) \cosh(-x) = \cosh x. \\ (c) \sinh(-x) = -\sinh x. \quad (d) \sinh(a+x) = \sinh a \cosh x + \cosh a \sinh x.$$

**13.18.** (a) Derive Equation (13.30) of Example 13.6.4.

(b) Derive Equation (13.31) of Example 13.6.4 by direct substitution.

(c) Let  $\lambda = l(l+1)$  in Example 13.6.4 and calculate the Legendre polynomials  $P_l(x)$  for  $l = 0, 1, 2, 3$ , subject to the condition  $P_l(1) = 1$ .

**13.19.** Use Equation (13.33) of Example 13.6.5 to generate the first three Hermite polynomials. Use the normalization

$$\int_{-\infty}^{\infty} [H_n(x)]^2 e^{-x^2} dx = \sqrt{\pi} 2^n n!$$

to determine the arbitrary constant.

**13.20.** The function defined by

$$f(x) = \sum_{n=0}^{\infty} c_n x^n, \quad \text{where } c_{n+2} = \frac{2n - \lambda}{(n+1)(n+2)} c_n,$$

can be written as  $f(x) = c_0 g(x) + c_1 h(x)$ , where  $g$  is even and  $h$  is odd in  $x$ . Show that  $f(x)$  goes to infinity at least as fast as  $e^{x^2}$  does, i.e.,  $\lim_{x \rightarrow \infty} f(x)e^{-x^2} \neq 0$ . Hint: Consider  $g(x)$  and  $h(x)$  separately and show that

$$g(x) = \sum_{n=0}^{\infty} b_n x^n, \quad \text{where } b_{n+1} = \frac{4n - \lambda}{(2n+1)(2n+2)} b_n.$$

Then concentrate on the ratio  $g(x)/e^{x^2}$ , where  $g$  and  $e^{x^2}$  are approximated by polynomials of very high degrees. Take the limit of this ratio as  $x \rightarrow \infty$ , and use recursion relations for  $g$  and  $e^{x^2}$ . The odd case follows similarly.

**13.21.** Refer to Example 13.6.6 for this problem.

(a) Derive the commutation relation  $[\mathbf{a}, \mathbf{a}^\dagger] = 1$ .

(b) Show that the Hamiltonian can be written as given in Equation (13.34).

(c) Derive the commutation relation  $[\mathbf{a}, (\mathbf{a}^\dagger)^n] = n(\mathbf{a}^\dagger)^{n-1}$ .

(d) Take the inner product of Equation (13.36) with itself and use (c) to show that  $|c_n|^2 = n|c_{n-1}|^2$ . From this, conclude that  $|c_n|^2 = n!|c_0|^2$ .

(e) For any function  $f(y)$ , show that

$$\left(y - \frac{d}{dy}\right) (e^{y^2/2} f) = -e^{y^2/2} \frac{df}{dy}.$$

Apply  $(y - d/dy)$  repeatedly to both sides of the above equation to obtain

$$\left(y - \frac{d}{dy}\right)^n (e^{y^2/2} f) = (-1)^n e^{y^2/2} \frac{d^n f}{dy^n}.$$

(f) Choose an appropriate  $f(y)$  in part (e) and show that

$$e^{y^2/2} \left(y - \frac{d}{dy}\right)^n e^{-y^2/2} = (-1)^n e^{y^2/2} \frac{d^n}{dy^n} (e^{-y^2}).$$

**13.22.** Solve Airy's DE,  $y'' + xy = 0$ , by the power-series method. Show that the radius of convergence for both independent solutions is infinite. Use the comparison theorem to show that for  $x > 0$  these solutions have infinitely many zeros, but for  $x < 0$  they can have at most one zero.

**13.23.** Show that the functions  $x^r e^{\lambda x}$ , where  $r = 0, 1, 2, \dots, k$ , are linearly independent. Hint: Starting with  $(\mathbf{D} - \lambda)^k$ , apply powers of  $\mathbf{D} - \lambda$  to a linear combination of  $x^r e^{\lambda x}$  for all possible  $r$ 's.

**13.24.** Find a basis of real solutions for each DE.

(a)  $y'' + 5y' + 6 = 0$ .

(b)  $y''' + 6y'' + 12y' + 8y = 0$ .

(c)  $\frac{d^4 y}{dx^4} = y$ .

(d)  $\frac{d^4 y}{dx^4} = -y$ .

**13.25.** Solve the following initial value problems.

(a)  $\frac{d^4 y}{dx^4} = y$ ,  $y(0) = y'(0) = y'''(0) = 0$ ,  $y''(0) = 1$ .

(b)  $\frac{d^4 y}{dx^4} + \frac{d^2 y}{dx^2} = 0$ ,  $y(0) = y''(0) = y'''(0) = 0$ ,  $y'(0) = 1$ .

(c)  $\frac{d^4 y}{dx^4} = 0$ ,  $y(0) = y'(0) = y''(0) = 0$ ,  $y'''(0) = 2$ .

**13.26.** Solve  $y'' - 2y' + y = xe^x$  subject to the initial conditions  $y(0) = 0$ ,  $y'(0) = 1$ .

**13.27.** Find the general solution of each equation.

$$(a) y'' = xe^x.$$

$$(c) y'' + y = \sin x \sin 2x.$$

$$(e) y'' - y = e^x \sin 2x.$$

$$(g) y'' - 4y' + 4 = e^x + xe^{2x}.$$

$$(b) y'' - 4y' + 4y = x^2.$$

$$(d) y'' - y = (1 + e^{-x})^2.$$

$$(f) y^{(6)} - y^{(4)} = x^2.$$

$$(h) y'' + y = e^{2x}.$$

**13.28.** Consider the Euler equation,

$$x^n y^{(n)} + a_{n-1} x^{n-1} y^{(n-1)} + \dots + a_1 x y' + a_0 y = r(x).$$

Substitute  $x = e^t$  and show that such a substitution reduces this to a DE with constant coefficients. In particular, solve  $x^2 y'' - 4xy' + 6y = x$ .

**13.29.** (a) Show that the substitution (13.50) reduces the Schrödinger equation to (13.51).

(b) From the second equation of (13.51), derive the continuity equation for probability.

**13.30.** Show that the usual definition of probability current density,

$$\mathbf{J} = \text{Re} \left[ \psi^* \frac{\hbar}{im} \nabla \psi \right],$$

reduces to that in Equation (13.52) if we use (13.50).

**13.31.** Write a computer program that solves the following differential equations by

(a) Adam's method [Equation (13.62)] and

(b) the Runge-Kutta method [Equation (13.70)].

$$\dot{x} = t - x^2, \quad x(0) = 1$$

$$\dot{x} = e^{-xt}, \quad x(0) = 1$$

$$\dot{x} = x^2 t^2 + 1, \quad x(0) = 1$$

$$\dot{x} = t + \sin x, \quad x(0) = \pi/2$$

$$\dot{x} = \sin xt, \quad x(0) = 1$$

**13.32.** Solve the following IVPs numerically, with  $h = 0.1$ . Find the first ten values of  $x$ .

$$(a) \ddot{x} + 0.2\dot{x}^2 + 10x = 20t, \quad x(0) = 0, \quad \dot{x}(0) = 0.$$

$$(b) \ddot{x} + 4x = t^2, \quad x(0) = 1, \quad \dot{x}(0) = 0.$$

$$(c) \ddot{x} + \dot{x} + x = 0, \quad x(0) = 2, \quad \dot{x}(0) = 0.$$

$$(d) t\ddot{x} + \dot{x} + xt = 0, \quad x(0) = 1, \quad \dot{x}(0) = 0.$$

$$(e) \ddot{x} + \dot{x} + x^2 = t, \quad x(0) = 1, \quad \dot{x}(0) = 0.$$

$$(f) \ddot{x} + xt = 0, \quad x(0) = 0, \quad \dot{x}(0) = 1.$$

$$(g) \ddot{x} + \sin x = t0, \quad x(0) = \frac{\pi}{2}, \quad \dot{x}(0) = 0.$$



13.33. Substitute  $d/dt = \mathbf{D} = -(1/h)\ln(1 - \nabla)$  in the SOLDE  $\ddot{x} + p(t)\dot{x} + q(t)x = r(t)$  and expand the log terms to obtain

$$(\nabla^2 + \nabla^3)x_n - hp_n(\nabla + \frac{1}{2}\nabla^2)x_n + h^2q_nx_n = h^2r_n$$

Since  $\nabla$  is of order  $h$ , one has to keep one power fewer in the second term. Find an expression for  $x_n$  in terms of  $x_{n-1}$ ,  $x_{n-2}$ , and  $x_{n-3}$ , valid to  $h^2$ .

### Additional Reading

1. Birkhoff, G. and Rota, G.-C. *Ordinary Differential Equations*, 3rd ed., Wiley, 1978. The small size of this book is very deceptive. It is loaded with information. Written by two excellent mathematicians and authors, the book covers all the topics of this chapter and much more in a very clear and lucid style.
2. DeVries, P. *A First Course in Computational Physics*, Wiley, 1994. The numerical solutions of differential equations are discussed in detail. The approach is slightly different from the one used in this chapter.
3. Hildebrand, F. *Introduction to Numerical Analysis*, 2nd ed., Dover, 1987. Our treatment of numerical solutions of differential equations closely follows that of this reference.
4. Mathews, J. and Walker, R. *Mathematical Methods of Physics*, 2nd ed., Benjamin, 1970. A good source for WKB approximation.

## Complex Analysis of SOLDEs

We have familiarized ourselves with some useful techniques for finding solutions to differential equations. One powerful method that leads to formal solutions is power series. We also stated Theorem 13.6.7 which guarantees the convergence of the solution of the power series within a circle whose size is at least as large as the smallest of the circles of convergence of the coefficient functions. Thus, the convergence of the solution is related to the convergence of the coefficient functions. What about the nature of the convergence, or the analyticity of the solution? Is it related to the analyticity of the coefficient functions? If so, how? Are the singular points of the coefficients also singular points of the solution? Is the nature of the singularities the same? This chapter answers some of these questions.

Analyticity is best handled in the complex plane. An important reason for this is the property of analytic continuation discussed in Chapter 11. The differential equation  $du/dx = u^2$  has a solution  $u = -1/x$  for all  $x$  except  $x = 0$ . Thus, we have to “puncture” the real line by removing  $x = 0$  from it. Then we have two solutions, because the domain of definition of  $u = -1/x$  is not connected on the real line (technically, the definition of a function includes its domain as well as the rule for going from the domain to the range). In addition, if we confine ourselves to the real line, there is no way that we can connect the  $x > 0$  region to the  $x < 0$  region. However, in the complex plane the same equation,  $dw/dz = w^2$ , has the complex solution  $w = -1/z$ , which is analytic everywhere except at  $z = 0$ . Puncturing the complex plane does not destroy the connectivity of the region of definition of  $w$ . Thus, the solution in the  $x > 0$  region can be analytically continued to the solution in the  $x < 0$  region by going around the origin.

The aim of this chapter is to investigate the analytic properties of the solutions of some well known SOLDEs in mathematical physics. We begin with a result from differential equation theory (for a proof, see [Birk 78, p. 223]).

continuation  
principle

**14.0.1. Proposition.** (continuation principle) *The function obtained by analytic continuation of any solution of an analytic differential equation along any path in the complex plane is a solution of the analytic continuation of the differential equation along the same path.*

An analytic differential equation is one with analytic coefficient functions. This proposition makes it possible to find a solution in one region of the complex plane and then continue it analytically. The following example shows how the singularities of the coefficient functions affect the behavior of the solution.

**14.0.2. Example.** Let us consider the FODE  $w' - (\gamma/z)w = 0$  for  $\gamma \in \mathbb{R}$ . The coefficient function  $p(z) = -\gamma/z$  has a simple pole at  $z = 0$ . The solution to the FODE is easily found to be  $w = z^\gamma$ . Thus, depending on whether  $\gamma$  is a nonnegative integer, a negative integer  $-m$ , or a noninteger, the solution has a regular point, a pole of order  $m$ , or a branch point at  $z = 0$ , respectively. ■

This example shows that the singularities of the solution depend on the parameters of the differential equation.

## 14.1 Analytic Properties of Complex DEs

To prepare for discussing the analytic properties of the solutions of SOLDEs, let us consider some general properties of differential equations from a complex analytical point of view.

### 14.1.1 Complex FOLDEs

In the homogeneous FOLDE

$$\frac{dw}{dz} + p(z)w = 0, \quad (14.1)$$

$p(z)$  is assumed to have only isolated singular points. It follows that  $p(z)$  can be expanded about a point  $z_0$ —which may be a singularity of  $p(z)$ —as a Laurent series in some annular region  $r_1 < |z - z_0| < r_2$ :

$$p(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n \quad \text{where } r_1 < |z - z_0| < r_2.$$

The solution to Equation (14.1), as given in Theorem 13.2.1 with  $q = 0$ , is

$$\begin{aligned} w(z) &= \exp \left[ - \int p(z) dz \right] \\ &= C \exp \left[ -a_{-1} \int \frac{dz}{z - z_0} - \sum_{n=0}^{\infty} a_n \int (z - z_0)^n dz - \sum_{n=2}^{\infty} a_{-n} \int (z - z_0)^{-n} dz \right] \\ &= C \exp \left[ -a_{-1} \ln(z - z_0) - \sum_{n=0}^{\infty} \frac{a_n}{n+1} (z - z_0)^{n+1} + \sum_{n=1}^{\infty} \frac{a_{-n-1}}{n} (z - z_0)^{-n} \right]. \end{aligned}$$

We can write this solution as

$$w(z) = C(z - z_0)^\alpha g(z), \tag{14.2}$$

where  $\alpha \equiv -a_{-1}$  and  $g(z)$  is an analytic *single-valued* function in the annular region  $r_1 < |z - z_0| < r_2$  because  $g(z)$  is the exponential of an analytic function.

For the special case in which  $p$  has a simple pole, i.e., when  $a_{-n} = 0$  for all  $n \geq 2$ , the second sum in the exponent will be absent, and  $g$  will be analytic even at  $z_0$ . In fact,  $g(z_0) = 1$ , and choosing  $C = 1$ , we can write

$$w(z) = (z - z_0)^\alpha \left[ 1 + \sum_{k=1}^{\infty} b_k (z - z_0)^k \right]. \tag{14.3}$$

The singularity of the coefficient functions of an FOLDE determines the singularity of the solution.

Depending on the nature of the singularity of  $p(z)$  at  $z_0$ , the solutions given by Equation (14.2) have different classifications. For instance, if  $p(z)$  has a removable singularity (if  $a_{-n} = 0 \forall n \geq 1$ ), the solution is  $Cg(z)$ , which is analytic. In this case, we say that the FOLDE [Equation (14.1)] *has a removable singularity at  $z_0$* . If  $p(z)$  has a simple pole at  $z_0$  (if  $a_{-1} \neq 0$  and  $a_{-n} = 0 \forall n \geq 2$ ), then in general, the solution has a branch point at  $z_0$ . In this case we say that the FOLDE *has a regular singular point*. Finally, if  $p(z)$  has a pole of order  $m > 1$ , then the solution will have an essential singularity (see Problem 14.1). In this case the FOLDE is said to have an *irregular singular point*.

To arrive at the solution given by Equation (14.2), we had to solve the FOLDE. Since higher-order differential equations are not as easily solved, it is desirable to obtain such a solution through other considerations. The following example sets the stage for this endeavor.

**14.1.1. Example.** A FOLDE has a unique solution, to within a multiplicative constant, given by Theorem 13.2.1. Thus, given a solution  $w(z)$ , any other solution must be of the form  $Cw(z)$ . Let  $z_0$  be a singularity of  $p(z)$ , and let  $z - z_0 = re^{i\theta}$ . Start at a point  $z$  and circle  $z_0$  so that  $\theta \rightarrow \theta + 2\pi$ . Even though  $p(z)$  may have a simple pole at  $z_0$ , the solution may have a branch point there. This is clear from the general solution, where  $\alpha$  may be a noninteger. Thus,  $\tilde{w}(z) \equiv w(z_0 + re^{i(\theta+2\pi)})$  may be different from  $w(z)$ . To discover this branch point—without solving the DE—invoke Proposition 14.0.1 and conclude that  $\tilde{w}(z)$  is also a solution to the FOLDE. Thus,  $\tilde{w}(z)$  can be different from  $w(z)$  by at most a multiplicative constant:  $\tilde{w}(z) = Cw(z)$ . Define the *complex* number  $\alpha$  by  $C = e^{2\pi i\alpha}$ . Then the function  $g(z) \equiv (z - z_0)^{-\alpha} w(z)$  is single-valued around  $z_0$ . In fact,

$$\begin{aligned} g(z_0 + re^{i(\theta+2\pi)}) &= [re^{i(\theta+2\pi)}]^{-\alpha} w(z_0 + re^{i(\theta+2\pi)}) \\ &= (z - z_0)^{-\alpha} e^{-2\pi i\alpha} e^{2\pi i\alpha} w(z) = (z - z_0)^{-\alpha} w(z) = g(z). \end{aligned}$$

This argument shows that a solution  $w(z)$  of the FOLDE of Equation (14.1) can be written as  $w(z) = (z - z_0)^\alpha g(z)$ , where  $g(z)$  is single-valued. ■

### 14.1.2 The Circuit Matrix

The method used in Example 14.1.1 can be generalized to obtain a similar result for the NOLDE

$$\mathbb{L}[w] = \frac{d^n w}{dz^n} + p_{n-1}(z) \frac{d^{n-1} w}{dz^{n-1}} + \cdots + p_1(z) \frac{dw}{dz} + p_0(z)w = 0 \quad (14.4)$$

where all the  $p_i(z)$  are analytic in  $r_1 < |z - z_0| < r_2$ .

Let  $\{w_j(z)\}_{j=1}^n$  be a basis of solutions of Equation (14.4), and let  $z - z_0 = re^{i\theta}$ . Start at  $z$  and analytically continue the functions  $w_j(z)$  one complete turn to  $\theta + 2\pi$ . Let  $\tilde{w}_j(z) \equiv \tilde{w}_j(z_0 + re^{i\theta}) \equiv w_j(z_0 + re^{i(\theta+2\pi)})$ . Then, by a generalization of Proposition 14.0.1,  $\{\tilde{w}_j(z)\}_{j=1}^n$  are not only solutions, but they are linearly independent (because they are  $w_j$ 's evaluated at a different point). Therefore, they also form a basis of solutions. On the other hand,  $\tilde{w}_j(z)$  can be expressed as a linear combination of the  $w_j(z)$ . Thus,  $\tilde{w}_j(z) = w_j(z_0 + re^{i(\theta+2\pi)}) = \sum_{k=1}^n a_{jk} w_k(z)$ . The matrix  $A = (a_{jk})$ , called the **circuit matrix** of the NOLDE, is invertible, because it transforms one basis into another. Therefore, it has only nonzero eigenvalues. We let  $\lambda$  be one such eigenvalue, and choose the column vector  $C$ , with entries  $\{c_i\}_{i=1}^n$ , to be the corresponding eigenvector of the transpose of  $A$  (note that  $A$  and  $A^t$  have the same set of eigenvalues). At least one such eigenvector always exists, because the characteristic polynomial of  $A^t$  has at least one root. Now we let  $w(z) = \sum_{j=1}^n c_j w_j(z)$ . Clearly, this  $w(z)$  is a solution of (14.4), and

$$\begin{aligned} \tilde{w}(z) &\equiv w(z_0 + re^{i(\theta+2\pi)}) = \sum_{j=1}^n c_j w_j(z_0 + re^{i(\theta+2\pi)}) \\ &= \sum_{j=1}^n c_j \sum_{k=1}^n a_{jk} w_k(z) = \sum_{j,k} (A^t)_{kj} c_j w_k(z) = \sum_{k=1}^n \lambda c_k w_k(z) = \lambda w(z). \end{aligned}$$

If we define  $\alpha$  by  $\lambda = e^{2\pi i \alpha}$ , then  $w(z_0 + re^{i(\theta+2\pi)}) = e^{2\pi i \alpha} w(z)$ . Now we write  $f(z) \equiv (z - z_0)^{-\alpha} w(z)$ . Following the argument used in Example 14.1.1, we get  $f(z_0 + re^{i(\theta+2\pi)}) = f(z)$ ; that is,  $f(z)$  is single-valued around  $z_0$ . We thus have the following theorem.

**14.1.2. Theorem.** *Any homogeneous NOLDE with analytic coefficient functions in  $r_1 < |z - z_0| < r_2$  admits a solution of the form*

$$w(z) = (z - z_0)^\alpha f(z)$$

where  $f(z)$  is single-valued around  $z_0$  in  $r_1 < |z - z_0| < r_2$ .

An isolated singular point  $z_0$  near which an analytic function  $w(z)$  can be written as  $w(z) = (z - z_0)^\alpha f(z)$ , where  $f(z)$  is single-valued and analytic in the punctured neighborhood of  $z_0$ , is called a **simple branch point** of  $w(z)$ . The arguments leading to Theorem 14.1.2 imply that a solution with a simple branch

point exists if and only if the vector  $C$  whose components appear in  $w(z)$  is an eigenvector of  $A^t$ , the transpose of the circuit matrix. Thus, there are as many solutions with simple branch points as there are linearly independent eigenvectors of  $A^t$ .

### 14.2 Complex SOLDES

Let us now consider the SOLDE  $w'' + p(z)w' + q(z)w = 0$ . Given two linearly independent solutions  $w_1(z)$  and  $w_2(z)$ , we form the  $2 \times 2$  matrix  $A$  and try to diagonalize it. There are three possible outcomes:

1. The matrix  $A$  is diagonalizable, and we can find two eigenvectors,  $F(z)$  and  $G(z)$ , corresponding, respectively, to two distinct eigenvalues,  $\lambda_1$  and  $\lambda_2$ . This means that  $F(z_0 + re^{i(\theta+2\pi)}) = \lambda_1 F(z)$  and  $G(z_0 + re^{i(\theta+2\pi)}) = \lambda_2 G(z)$ . Defining  $\lambda_1 = e^{2\pi i\alpha}$  and  $\lambda_2 = e^{2\pi i\beta}$ , we get  $F(z) = (z - z_0)^\alpha f(z)$  and  $G(z) = (z - z_0)^\beta g(z)$ , as Theorem 14.1.2 suggests. The set  $\{F(z), G(z)\}$  is called a **canonical basis** of the SOLDE.

canonical basis of the SOLDE

2. The matrix  $A$  is diagonalizable, and the two eigenvalues are the same. In this case both  $F(z)$  and  $G(z)$  have the same constant  $\alpha$ :

$$F(z) = (z - z_0)^\alpha f(z) \quad \text{and} \quad G(z) = (z - z_0)^\alpha g(z).$$

3. We cannot find two eigenvectors. This corresponds to the case where  $A$  is not diagonalizable. However, we can always find one eigenvector, so  $A$  has only one eigenvalue,  $\lambda$ . We let  $w_1(z)$  be the solution of the form  $(z - z_0)^\alpha f(z)$ , where  $f(z)$  is single-valued and  $\lambda = e^{2\pi i\alpha}$ . The existence of such a solution is guaranteed by Theorem 14.1.2. Let  $w_2(z)$  be any other linearly independent solution (Theorem 13.3.5 ensures the existence of such a second solution). Then

$$w_2(z_0 + re^{i(\theta+2\pi)}) = aw_1(z) + bw_2(z),$$

and the circuit matrix will be  $A = \begin{pmatrix} \lambda & 0 \\ a & b \end{pmatrix}$ , which has eigenvalues  $\lambda$  and  $b$ . Since  $A$  is assumed to have only one eigenvalue (otherwise we would have the first outcome again), we must have  $b = \lambda$ . This reduces  $A$  to  $A = \begin{pmatrix} \lambda & 0 \\ a & \lambda \end{pmatrix}$ , where  $a \neq 0$ . The condition  $a \neq 0$  is necessary to distinguish this case from the second outcome. Now we analytically continue  $h(z) \equiv w_2(z)/w_1(z)$  one whole turn around  $z_0$ , obtaining

$$\begin{aligned} h(z_0 + re^{i(\theta+2\pi)}) &= \frac{w_2(z_0 + re^{i(\theta+2\pi)})}{w_1(z_0 + re^{i(\theta+2\pi)})} = \frac{aw_1(z) + \lambda w_2(z)}{\lambda w_1(z)} \\ &= \frac{a}{\lambda} + \frac{w_2(z)}{w_1(z)} = \frac{a}{\lambda} + h(z). \end{aligned}$$

It then follows that the function<sup>1</sup>

$$g_1(z) \equiv h(z) - \frac{a}{2\pi i \lambda} \ln(z - z_0)$$

is single-valued in  $r_1 < |z - z_0| < r_2$ . If we redefine  $g_1(z)$  and  $w_2(z)$  as  $(2\pi i \lambda/a)g_1(z)$  and  $(2\pi i \lambda/a)w_2(z)$ , respectively, we have the following:

**14.2.1. Theorem.** *If  $p(z)$  and  $q(z)$  are analytic in the annular region  $r_1 < |z - z_0| < r_2$ , then the SOLDE  $w'' + p(z)w' + q(z)w = 0$  admits a basis of solutions  $\{w_1, w_2\}$  in the neighborhood of the singular point  $z_0$ , where either*

$$w_1(z) = (z - z_0)^\alpha f(z), \quad w_2(z) = (z - z_0)^\beta g(z)$$

or, in exceptional cases (when the circuit matrix is not diagonalizable),

$$w_1(z) = (z - z_0)^\alpha f(z), \quad w_2(z) = w_1(z)[g_1(z) + \ln(z - z_0)].$$

The functions  $f(z)$ ,  $g(z)$ , and  $g_1(z)$  are analytic and single-valued in the annular region.

This theorem allows us to factor out the *branch point*  $z_0$  from the rest of the solutions. However, even though  $f(z)$ ,  $g(z)$ , and  $g_1(z)$  are analytic in the annular region  $r_1 < |z - z_0| < r_2$ , they may very well have poles of arbitrary orders at  $z_0$ . Can we also factor out the *poles*? In general, we cannot; however, under special circumstances, described in the following definition, we can.

**14.2.2. Definition.** *A SOLDE of the form  $w'' + p(z)w' + q(z)w = 0$  that is analytic in  $0 < |z - z_0| < r$  has a **regular singular point** at  $z_0$  if  $p(z)$  has at worst a simple pole and  $q(z)$  has at worst a pole of order 2 there.*

regular singular point  
of a SOLDE defined

In a neighborhood of a regular singular point  $z_0$ , the coefficient functions  $p(z)$  and  $q(z)$  have the power-series expansions

$$p(z) = \frac{a_{-1}}{z - z_0} + \sum_{k=0}^{\infty} a_k(z - z_0)^k,$$

$$q(z) = \frac{b_{-2}}{(z - z_0)^2} + \frac{b_{-1}}{z - z_0} + \sum_{k=0}^{\infty} b_k(z - z_0)^k.$$

Multiplying both sides of the first equation by  $z - z_0$  and the second by  $(z - z_0)^2$  and introducing  $P(z) \equiv (z - z_0)p(z)$ ,  $Q(z) \equiv (z - z_0)^2q(z)$ , we obtain

$$P(z) = \sum_{k=0}^{\infty} a_{k-1}(z - z_0)^k, \quad Q(z) = \sum_{k=0}^{\infty} b_{k-2}(z - z_0)^k.$$

<sup>1</sup>Recall that  $\ln(z - z_0)$  increases by  $2\pi i$  for each turn around  $z_0$ .

It is also convenient to multiply the SOLDE by  $(z - z_0)^2$  and write it as

$$(z - z_0)^2 w'' + (z - z_0)P(z)w' + Q(z)w = 0. \tag{14.5}$$

Inspired by the discussion leading to Theorem 14.2.1, we write

$$w(z) = (z - z_0)^\nu \sum_{k=0}^{\infty} C_k (z - z_0)^k, \quad C_0 = 1, \tag{14.6}$$

where we have chosen the arbitrary multiplicative constant in such a way that  $C_0 = 1$ . Substitute this in Equation (14.5), and change the dummy variable—so that all sums start at 0—to obtain

$$\sum_{n=0}^{\infty} \left\{ (n + \nu)(n + \nu - 1)C_n + \sum_{k=0}^n [(k + \nu)a_{n-k-1} + b_{n-k-2}]C_k \right\} (z - z_0)^{n+\nu} = 0,$$

which results in the recursion relation

$$(n + \nu)(n + \nu - 1)C_n = - \sum_{k=0}^n [(k + \nu)a_{n-k-1} + b_{n-k-2}]C_k. \tag{14.7}$$

indicial equation,  
indicial polynomial,  
characteristic  
exponents

For  $n = 0$ , this leads to what is known as the **indicial equation** for the exponent  $\nu$ :

$$I(\nu) \equiv \nu(\nu - 1) + a_{-1}\nu + b_{-2} = 0. \tag{14.8}$$

The roots of this equation are called the **characteristic exponents** of  $z_0$ , and  $I(\nu)$  is called its **indicial polynomial**. In terms of this polynomial, (14.7) can be expressed as

$$I(n + \nu)C_n = - \sum_{k=0}^{n-1} [(k + \nu)a_{n-k-1} + b_{n-k-2}]C_k \quad \text{for } n = 1, 2, \dots \tag{14.9}$$

Equation (14.8) determines what values of  $\nu$  are possible, and Equation (14.9) gives  $C_1, C_2, C_3, \dots$ , which in turn determine  $w(z)$ . Special care must be taken if the indicial polynomial vanishes at  $n + \nu$  for some positive integer  $n$ , that is, if  $n + \nu$ , in addition to  $\nu$ , is a root of the indicial polynomial:  $I(n + \nu) = 0 = I(\nu)$ .

If  $\nu_1$  and  $\nu_2$  are characteristic exponents of the indicial equation and  $\text{Re}(\nu_1) > \text{Re}(\nu_2)$ , then a solution for  $\nu_1$  always exists. A solution for  $\nu_2$  also exists if  $\nu_1 - \nu_2 \neq n$  for any (positive) integer  $n$ . In particular, if  $z_0$  is an ordinary point [a point at which both  $p(z)$  and  $q(z)$  are analytic], then only one solution is determined by (14.9). (Why?) The foregoing discussion is summarized in the following:



**14.2.3. Theorem.** *If the differential equation  $w'' + p(z)w' + q(z)w = 0$  has a regular singular point at  $z = z_0$ , then at least one power series of the form of (14.6) formally solves the equation. If  $\nu_1$  and  $\nu_2$  are the characteristic exponents of  $z_0$ , then there are two linearly independent formal solutions unless  $\nu_1 - \nu_2$  is an integer.*

**14.2.4. Example.** Let us consider some familiar differential equations.

(a) The Bessel equation is

$$w'' + \frac{1}{z}w' + \left(1 - \frac{\alpha^2}{z^2}\right)w = 0.$$

In this case, the origin is a regular singular point,  $a_{-1} = 1$ , and  $b_{-2} = -\alpha^2$ . Thus, the indicial equation is  $\nu(\nu - 1) + \nu - \alpha^2 = 0$ , and its solutions are  $\nu_1 = \alpha$  and  $\nu_2 = -\alpha$ . Therefore, there are two linearly independent solutions to the Bessel equation unless  $\nu_1 - \nu_2 = 2\alpha$  is an integer, i.e., unless  $\alpha$  is either an integer or a half-integer.

(b) For the Coulomb potential  $f(r) = \beta/r$ , the most general radial equation [Equation (12.14)] reduces to

$$w'' + \frac{2}{z}w' + \left(\frac{\beta}{z} - \frac{\alpha}{z^2}\right)w = 0.$$

The point  $z = 0$  is a regular singular point at which  $a_{-1} = 2$  and  $b_{-2} = -\alpha$ . The indicial polynomial is  $I(\nu) = \nu^2 + \nu - \alpha$  with characteristic exponents  $\nu_1 = -\frac{1}{2} + \frac{1}{2}\sqrt{1 + 4\alpha}$  and  $\nu_2 = -\frac{1}{2} - \frac{1}{2}\sqrt{1 + 4\alpha}$ . There are two independent solutions unless  $\nu_1 - \nu_2 = \sqrt{1 + 4\alpha}$  is an integer. In practice,  $\alpha = l(l + 1)$ , where  $l$  is some integer; so  $\nu_1 - \nu_2 = 2l + 1$ , and only one solution is obtained.

(c) The hypergeometric differential equation is

$$w'' + \frac{\gamma - (\alpha + \beta + 1)z}{z(1 - z)}w' - \frac{\alpha\beta}{z(1 - z)}w = 0.$$

A substantial number of functions in mathematical physics are solutions of this remarkable equation, with appropriate values for  $\alpha$ ,  $\beta$ , and  $\gamma$ . The regular singular points<sup>2</sup> are  $z = 0$  and  $z = 1$ . At  $z = 0$ ,  $a_{-1} = \gamma$  and  $b_{-2} = 0$ . The indicial polynomial is  $I(\nu) = \nu(\nu + \gamma - 1)$ , whose roots are  $\nu_1 = 0$  and  $\nu_2 = 1 - \gamma$ . Unless  $\gamma$  is an integer, we have two formal solutions. ■

It is shown in differential equation theory [Birk 78, pp. 40–242] that as long as  $\nu_1 - \nu_2$  is not an integer, the series solution of Theorem 14.2.3 is convergent for a neighborhood of  $z_0$ . What happens when  $\nu_1 - \nu_2$  is an integer? First, as a convenience, we translate the coordinate axes so that the point  $z_0$  coincides with the origin. This will save us some writing, because instead of powers of  $z - z_0$ , we will have powers of  $z$ . Next we let  $\nu_1 = \nu_2 + n$  with  $n$  a positive integer. Then, since it is impossible to encounter any new zero of the indicial polynomial beyond

<sup>2</sup>The coefficient of  $w$  need not have a pole of order 2. Its pole can be of order one as well.

$v_1$ , the recursion relation, Equation (14.9), will be valid for all values of  $n$ , and we obtain a solution:

$$w_1(z) = z^{v_1} f(z) = z^{v_1} \left( 1 + \sum_{k=1}^{\infty} C_k z^k \right),$$

which is convergent in the region  $0 < |z| < r$  for some  $r > 0$ . To investigate the nature and the possibility of the second solution, write the recursion relations of Equation (14.9) for the smaller characteristic root  $v_2$ :

$$\begin{aligned} I(v_2 + 1)C_1 &= \overbrace{-(v_2 a_0 + b_{-1})}^{\equiv \rho_1 I(v_2+1)} C_0 \Rightarrow C_1 = \rho_1, \\ I(v_2 + 2)C_2 &= -(v_2 a_1 + b_0)C_0 - [(v_2 + 1)a_0 + b_{-1}]C_1 \Rightarrow C_2 \equiv \rho_2, \\ &\vdots \\ I(v_2 + n - 1)C_{n-1} &\equiv \rho_{n-1} I(v_2 + n - 1)C_0 \Rightarrow C_{n-1} = \rho_{n-1}, \\ I(v_2 + n)C_n &= I(v_1)C_n = \rho_n C_0 \Rightarrow 0 = \rho_n, \end{aligned} \tag{14.10}$$

where in each step, we have used the result of the previous step in which  $C_k$  is given as a multiple of  $C_0 = 1$ . Here, the  $\rho$ 's are constants depending (possibly in a very complicated way) on the  $a_k$ 's and  $b_k$ 's.

Theorem 14.2.3 guarantees two power series solutions only when  $v_1 - v_2$  is not an integer. When  $v_1 - v_2$  is an integer, Equation (14.10) shows that a necessary condition for a second *power series* solution to exist is that  $\rho_n = 0$ . Therefore, when  $\rho_n \neq 0$ , we have to resort to other means of obtaining the second solution.

Let us define the second solution as

$$w_2(z) \equiv \overbrace{w_1(z)}^{\equiv w_1(z)} h(z) = z^{v_1} f(z) h(z) \tag{14.11}$$

and substitute in the SOLDE to obtain a FOLDE in  $h'$ , namely,  $h'' + (p + 2w'_1/w_1)h' = 0$ , or, by substituting  $w'_1/w_1 = v_1/z + f'/f$ , the equivalent FOLDE

$$h'' + \left( \frac{2v_1}{z} + \frac{2f'}{f} + p \right) h' = 0. \tag{14.12}$$

**14.2.5. Lemma.** *The coefficient of  $h'$  in Equation (14.12) has a residue of  $n + 1$ .*

*Proof.* Recall that the residue of a function is the coefficient of  $z^{-1}$  in the Laurent expansion of the function (about  $z = 0$ ). Let us denote this residue for the coefficient of  $h'$  by  $A_{-1}$ . Since  $f(0) = 1$ , the ratio  $f'/f$  is analytic at  $z = 0$ . Thus, the simple pole at  $z = 0$  comes from the other two terms. Substituting the Laurent expansion of  $p(z)$  gives

$$\frac{2v_1}{z} + p = \frac{2v_1}{z} + \frac{a_{-1}}{z} + a_0 + a_1 z + \dots$$

This shows that  $A_{-1} = 2v_1 + a_{-1}$ . On the other hand, comparing the two versions of the indicial polynomial  $v^2 + (a_{-1} - 1)v + b_{-2}$  and  $(v - v_1)(v - v_2) = v^2 - (v_1 + v_2)v + v_1v_2$  gives  $v_1 + v_2 = -(a_{-1} - 1)$ , or  $2v_1 - n = -(a_{-1} - 1)$ . Therefore,  $A_{-1} = 2v_1 + a_{-1} = n + 1$ .  $\square$

**14.2.6. Theorem.** *Suppose that the characteristic exponents of a SOLDE with a regular singular point at  $z = 0$  are  $v_1$  and  $v_2$ . Consider three cases:*

1.  $v_1 - v_2$  is not an integer.
2.  $v_2 = v_1 - n$  where  $n$  is a nonnegative integer, and  $\rho_n$ , as defined in Equation (14.10), vanishes.
3.  $v_2 = v_1 - n$  where  $n$  is a nonnegative integer, and  $\rho_n$ , as defined in Equation (14.10), does not vanish.

Then, in the first two cases, there exists a basis of solutions  $\{w_1, w_2\}$  of the form

$$w_i(z) = z^{v_i} \left( 1 + \sum_{k=1}^{\infty} C_k^{(i)} z^k \right), \quad i = 1, 2,$$

and in the third case, the basis of solutions takes the form

$$w_1(z) = z^{v_1} \left( 1 + \sum_{k=1}^{\infty} a_k z^k \right), \quad w_2(z) = z^{v_2} \left( 1 + \sum_{k=1}^{\infty} b_k z^k \right) + C w_1(z) \ln z,$$

where the power series are convergent in a neighborhood of  $z = 0$ .

*Proof.* The first two cases have been shown before. For the third case, we use Lemma 14.2.5 and write

$$\frac{2v_1}{z} + \frac{2f'}{f} + p = \frac{n+1}{z} + \sum_{k=0}^{\infty} c_k z^k,$$

and the solution for the FOLDE in  $h'$  will be [see Equation (14.3) and the discussion preceding it]

$$h'(z) = z^{-n-1} \left( 1 + \sum_{k=1}^{\infty} b_k z^k \right).$$

For  $n = 0$ , i.e., when the indicial polynomial has a double root, this yields  $h'(z) = 1/z + \sum_{k=1}^{\infty} b_k z^{k-1}$ , or  $h(z) = \ln z + g_1(z)$ , where  $g_1$  is analytic in a neighborhood of  $z = 0$ . For  $n \neq 0$ , we have  $h'(z) = b_n/z + \sum_{k \neq n}^{\infty} b_k z^{k-n-1}$  and, by integration,

$$\begin{aligned} h(z) &= b_n \ln z + \sum_{k \neq n}^{\infty} \frac{b_k}{k-n} z^{k-n} \\ &= b_n \ln z + z^{-n} \sum_{k \neq n}^{\infty} \frac{b_k}{k-n} z^k = b_n \ln z + z^{-n} g_2(z), \end{aligned}$$

where  $g_2$  is analytic in a neighborhood of  $z = 0$ . Substituting  $h$  in Equation (14.11) and recalling that  $\nu_2 = \nu_1 - n$ , we obtain the desired results of the theorem.  $\square$

### 14.3 Fuchsian Differential Equations

In many cases of physical interest, the behavior of the solution of a SOLDE at infinity is important. For instance, bound state solutions of the Schrödinger equation describing the probability amplitudes of particles in quantum mechanics must tend to zero as the distance from the center of the binding force increases.

We have seen that the behavior of a solution is determined by the behavior of the coefficient functions. To determine the behavior at infinity, we substitute  $z = 1/t$  in the SOLDE

$$\frac{d^2 w}{dz^2} + p(z) \frac{dw}{dz} + q(z)w = 0 \quad (14.13)$$

and obtain

$$\frac{d^2 v}{dt^2} + \left[ \frac{2}{t} - \frac{1}{t^2} r(t) \right] \frac{dv}{dt} + \frac{1}{t^4} s(t)v = 0, \quad (14.14)$$

where  $v(t) = w(1/t)$ ,  $r(t) = p(1/t)$ , and  $s(t) = q(1/t)$ .

Clearly, as  $z \rightarrow \infty$ ,  $t \rightarrow 0$ . Thus, we are interested in the behavior of (14.14) at  $t = 0$ . We assume that both  $r(t)$  and  $s(t)$  are analytic at  $t = 0$ . Equation (14.14) shows, however, that the solution  $v(t)$  may still have singularities at  $t = 0$  because of the extra terms appearing in the coefficient functions.

We assume that infinity is a regular singular point of (14.13), by which we mean that  $t = 0$  is a regular singular point of (14.14). Therefore, in the Taylor expansions of  $r(t)$  and  $s(t)$ , the first (constant) term of  $r(t)$  and the first two terms of  $s(t)$  must be zero. Thus, we write

$$r(t) = a_1 t + a_2 t^2 + \cdots = \sum_{k=1}^{\infty} a_k t^k,$$

$$s(t) = b_2 t^2 + b_3 t^3 + \cdots = \sum_{k=2}^{\infty} b_k t^k.$$

By their definitions, these two equations imply that for  $p(z)$  and  $q(z)$ , and for large values of  $|z|$ , we must have expressions of the form

$$p(z) = \frac{a_1}{z} + \frac{a_2}{z^2} + \cdots = \sum_{k=1}^{\infty} \frac{a_k}{z^k},$$

$$q(z) = \frac{b_2}{z^2} + \frac{b_3}{z^3} + \cdots = \sum_{k=2}^{\infty} \frac{b_k}{z^k}. \quad (14.15)$$

When infinity is a regular singular point of Equation (14.13), or, equivalently, when the origin is a regular singular point of (14.14), it follows from Theorem 14.2.6 that there exists at least one solution of the form  $v_1(t) = t^\alpha (1 + \sum_{k=1}^{\infty} C_k t^k)$  or, in terms of  $z$ ,

$$w_1(z) = z^{-\alpha} \left( 1 + \sum_{k=1}^{\infty} \frac{C_k}{z^k} \right). \tag{14.16}$$

Here  $\alpha$  is a characteristic exponents at  $t = 0$  of (14.14), whose indicial polynomial is easily found to be  $\alpha(\alpha - 1) + (2 - a_1)\alpha + b_2 = 0$ .

**Fuchsian DE 14.3.1. Definition.** *A homogeneous differential equation with single-valued analytic coefficient functions is called a Fuchsian differential equation (FDE) if it has only regular singular points in the extended complex plane, i.e., the complex plane including the point at infinity.*

It turns out that a particular kind of FDE describes a large class of nonelementary functions encountered in mathematical physics. Therefore, it is instructive to classify various kinds of FDEs. A fact that is used in such a classification is that complex functions whose only singularities in the extended complex plane are poles are rational functions, i.e., ratios of polynomials (see Example 10.2.2). We thus expect FDEs to have only rational functions as coefficients.

A second-order Fuchsian DE with two regular singular points leads to uninteresting solutions!

Consider the case where the equation has at most two regular singular points at  $z_1$  and  $z_2$ . We introduce a new variable  $\xi(z) = \frac{z - z_1}{z - z_2}$ . The regular singular points at  $z_1$  and  $z_2$  are mapped onto the points  $\xi_1 = \xi(z_1) = 0$  and  $\xi_2 = \xi(z_2) = \infty$ , respectively, in the extended  $\xi$ -plane. Equation (14.13) becomes

$$\frac{d^2u}{d\xi^2} + \Phi(\xi) \frac{du}{d\xi} + \Theta(\xi)u = 0, \tag{14.17}$$

where  $u$ ,  $\Phi$ , and  $\Theta$  are functions of  $\xi$  obtained when  $z$  is expressed in terms of  $\xi$  in  $w(z)$ ,  $p(z)$ , and  $q(z)$ , respectively. From Equation (14.15) and the fact that  $\xi = 0$  is at most a simple pole of  $\Phi(\xi)$ , we obtain  $\Phi(\xi) = a_1/\xi$ . Similarly,  $\Theta(\xi) = b_2/\xi^2$ . Thus, a SOFDE with two regular singular points is equivalent to the DE  $w'' + (a_1/z)w' + (b_2/z^2)w = 0$ . Multiplying both sides by  $z^2$ , we obtain  $z^2w'' + a_1zw' + b_2w = 0$ , which is the second-order Euler differential equation. A general  $n$ th-order Euler differential equation is equivalent to a NOLDE with constant coefficients (see Problem 13.28). Thus, a second order Fuchsian DE (SOFDE) with two regular singular points is equivalent to a SOLDE with constant coefficients and produces nothing new.

A second-order Fuchsian DE with three regular singular points leads to interesting solutions!

The simplest SOFDE whose solutions may include nonelementary functions is therefore one having three regular singular points, at say  $z_1$ ,  $z_2$ , and  $z_3$ . By the transformation

$$\xi(z) = \frac{(z - z_1)(z_3 - z_2)}{(z - z_2)(z_3 - z_1)}$$

we can map  $z_1, z_2,$  and  $z_3$  onto  $\xi_1 = 0, \xi_2 = \infty,$  and  $\xi_3 = 1.$  Thus, we assume that the three regular singular points are at  $z = 0, z = 1,$  and  $z = \infty.$  It can be shown [see Problem (14.8)] that the most general  $p(z)$  and  $q(z)$  are

$$p(z) = \frac{A_1}{z} + \frac{B_1}{z-1} \quad \text{and} \quad q(z) = \frac{A_2}{z^2} + \frac{B_2}{(z-1)^2} - \frac{A_3}{z(z-1)}.$$

We thus have the following theorem.

Riemann differential equation

**14.3.2. Theorem.** *The most general second order Fuchsian DE with three regular singular points can be transformed into the form*

$$w'' + \left( \frac{A_1}{z} + \frac{B_1}{z-1} \right) w' + \left[ \frac{A_2}{z^2} + \frac{B_2}{(z-1)^2} - \frac{A_3}{z(z-1)} \right] w = 0, \tag{14.18}$$

where  $A_1, A_2, A_3, B_1,$  and  $B_2$  are constants. This equation is called the **Riemann differential equation.**

We can write the Riemann DE in terms of pairs of characteristic exponents,  $(\lambda_1, \lambda_2), (\mu_1, \mu_2),$  and  $(\nu_1, \nu_2),$  belonging to the singular points  $0, 1,$  and  $\infty,$  respectively. The indicial equations are easily found to be

$$\begin{aligned} \lambda^2 + (A_1 - 1)\lambda + A_2 &= 0, \\ \mu^2 + (B_1 - 1)\mu + B_2 &= 0, \\ \nu^2 + (1 - A_1 - B_1)\nu + A_2 + B_2 - A_3 &= 0. \end{aligned}$$

By writing the indicial equations as  $(\lambda - \lambda_1)(\lambda - \lambda_2) = 0,$  and so forth and comparing coefficients, we can find the following relations:

$$\begin{aligned} A_1 &= 1 - \lambda_1 - \lambda_2, & A_2 &= \lambda_1 \lambda_2, \\ B_1 &= 1 - \mu_1 - \mu_2, & B_2 &= \mu_1 \mu_2, \\ A_1 + B_1 &= \nu_1 + \nu_2 + 1, & A_2 + B_2 - A_3 &= \nu_1 \nu_2. \end{aligned}$$

These equations lead easily to the *Riemann identity*

$$\lambda_1 + \lambda_2 + \mu_1 + \mu_2 + \nu_1 + \nu_2 = 1. \tag{14.19}$$

Substituting these results in (14.18) gives the following result.

**14.3.3. Theorem.** *A second order Fuchsian DE with three regular singular points in the extended complex plane is equivalent to the Riemann DE,*

$$\begin{aligned} w'' + \left( \frac{1 - \lambda_1 - \lambda_2}{z} + \frac{1 - \mu_1 - \mu_2}{z-1} \right) w' \\ + \left[ \frac{\lambda_1 \lambda_2}{z^2} + \frac{\mu_1 \mu_2}{(z-1)^2} + \frac{\nu_1 \nu_2 - \lambda_1 \lambda_2 - \mu_1 \mu_2}{z(z-1)} \right] w = 0, \end{aligned} \tag{14.20}$$

which is uniquely determined by the pairs of characteristic exponents at each singular point. The characteristic exponents satisfy the Riemann identity, Equation (14.19).

The uniqueness of the Riemann DE allows us to derive identities for solutions and reduce the independent parameters of Equation (14.20) from five to three. We first note that if  $w(z)$  is a solution of the Riemann DE corresponding to  $(\lambda_1, \lambda_2)$ ,  $(\mu_1, \mu_2)$ , and  $(\nu_1, \nu_2)$ , then the function

$$v(z) = z^\lambda (z-1)^\mu w(z)$$

has branch points at  $z = 0, 1, \infty$  [because  $w(z)$  does]; therefore, it is a solution of the Riemann DE. Its pairs of characteristic exponents are (see Problem 14.10)

$$(\lambda_1 + \lambda, \lambda_2 + \lambda), \quad (\mu_1 + \mu, \mu_2 + \mu), \quad (\nu_1 - \lambda - \mu, \nu_2 - \lambda - \mu).$$

In particular, if we let  $\lambda = -\lambda_1$  and  $\mu = -\mu_1$ , then the pairs reduce to

$$(0, \lambda_2 - \lambda_1), \quad (0, \mu_2 - \mu_1), \quad (\nu_1 + \lambda_1 + \mu_1, \nu_2 + \lambda_1 + \mu_1).$$

Defining  $\alpha \equiv \nu_1 + \lambda_1 + \mu_1$ ,  $\beta \equiv \nu_2 + \lambda_1 + \mu_1$ , and  $\gamma \equiv 1 - \lambda_2 + \lambda_1$ , and using (14.19), we can write the pairs as

$$(0, 1 - \gamma), \quad (0, \gamma - \alpha - \beta), \quad (\alpha, \beta),$$

which yield the third version of the Riemann DE

$$w'' + \left( \frac{\gamma}{z} + \frac{1 - \gamma + \alpha + \beta}{z-1} \right) w' + \frac{\alpha\beta}{z(z-1)} w = 0.$$

hypergeometric DE This important equation is commonly written in the equivalent form

$$z(1-z)w'' + [\gamma - (1 + \alpha + \beta)z]w' - \alpha\beta w = 0 \quad (14.21)$$

and is called the **hypergeometric** differential equation (HGDE). We will study this equation next.

## 14.4 The Hypergeometric Function

The two characteristic exponents of Equation (14.21) at  $z = 0$  are 0 and  $1 - \gamma$ . It follows from Theorem 14.2.6 that there exists an *analytic* solution (corresponding to the characteristic exponent 0) at  $z = 0$ . Let us denote this solution, the **hypergeometric function**, by  $F(\alpha, \beta; \gamma; z)$  and write

hypergeometric  
function

$$F(\alpha, \beta; \gamma; z) = \sum_{k=0}^{\infty} a_k z^k \quad \text{where } a_0 = 1.$$

Substituting in the DE, we obtain the recurrence relation

$$a_{k+1} = \frac{(\alpha + k)(\beta + k)}{(k+1)(\gamma + k)} a_k \quad \text{for } k \geq 0.$$

hypergeometric series

These coefficients can be determined successively if  $\gamma$  is neither zero nor a negative integer:

$$\begin{aligned}
 F(\alpha, \beta; \gamma; z) &= 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)\beta(\beta+1)\cdots(\beta+k-1)}{k!\gamma(\gamma+1)\cdots(\gamma+k-1)} z^k \\
 &= \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(\alpha+k)\Gamma(\beta+k)}{\Gamma(k+1)\Gamma(\gamma+k)} z^k. \tag{14.22}
 \end{aligned}$$

The series in (14.22) is called the **hypergeometric series**, because it is the generalization of  $F(1, \beta; \beta; z)$ , which is simply the geometric series.

We note immediately from (14.22) that

**14.4.1. Box.** *The hypergeometric series becomes a polynomial if either  $\alpha$  or  $\beta$  is a negative integer.*

This is because for  $k < |\alpha|$  (or  $k < |\beta|$ ) both  $\Gamma(\alpha+k)$  [or  $\Gamma(\beta+k)$ ] and  $\Gamma(\alpha)$  [or  $\Gamma(\beta)$ ] have poles that cancel each other. However,  $\Gamma(\alpha+k)$  [or  $\Gamma(\beta+k)$ ] becomes finite for  $k > |\alpha|$  (or  $k > |\beta|$ ), and the pole in  $\Gamma(\alpha)$  [or  $\Gamma(\beta)$ ] makes the denominator infinite. Therefore, all terms of the series (14.22) beyond  $k = |\alpha|$  (or  $k = |\beta|$ ) will be zero.

Many of the properties of the hypergeometric function can be obtained directly from the HGDE, Equation (14.21). For instance, differentiating the HGDE and letting  $v = w'$ , we obtain

$$z(1-z)v'' + [\gamma + 1 - (\alpha + \beta + 3)z]v' - (\alpha + 1)(\beta + 1)v = 0,$$

which shows that  $F'(\alpha, \beta; \gamma; z) = CF(\alpha + 1, \beta + 1; \gamma + 1; z)$ . The constant  $C$  can be determined by differentiating Equation (14.22), setting  $z = 0$  in the result,<sup>3</sup> and noting that  $F(\alpha + 1, \beta + 1; \gamma + 1; 0) = 1$ . Then we obtain

$$F'(\alpha, \beta; \gamma; z) = \frac{\alpha\beta}{\gamma} F(\alpha + 1, \beta + 1; \gamma + 1; z). \tag{14.23}$$

Now assume that  $\gamma \neq 1$ , and make the substitution  $w = z^{1-\gamma}u$  in the HGDE to obtain<sup>4</sup>  $z(1-z)u'' + [\gamma_1 - (\alpha_1 + \beta_1 + 1)z]u' - \alpha_1\beta_1u = 0$ , where  $\alpha_1 = \alpha - \gamma + 1$ ,  $\beta_1 = \beta - \gamma + 1$ , and  $\gamma_1 = 2 - \gamma$ . Thus,

$$u = F(\alpha - \gamma + 1, \beta - \gamma + 1; 2 - \gamma; z),$$

<sup>3</sup>Note that the hypergeometric function evaluates to 1 at  $z = 0$  regardless of its parameters.

<sup>4</sup>In the following discussion,  $\alpha_1, \beta_1$ , and  $\gamma_1$  will represent the parameters of the new DE satisfied by the new function defined in terms of the old.



and  $u$  is therefore analytic at  $z = 0$ . This leads to an interesting result. Provided that  $\gamma$  is not an integer, the two functions

$$w_1(z) \equiv F(\alpha, \beta; \gamma; z), \quad w_2(z) \equiv z^{1-\gamma} F(\alpha - \gamma + 1, \beta - \gamma + 1; 2 - \gamma; z) \quad (14.24)$$

form a canonical basis of solutions to the HGDE at  $z = 0$ . This follows from Theorem 14.2.6 and the fact that  $(0, 1 - \gamma)$  are a pair of (different) characteristic exponents at  $z = 0$ .

**Johann Carl Friedrich Gauss** (1777–1855) was the greatest of all mathematicians and perhaps the most richly gifted genius of whom there is any record. He was born in the city of Brunswick in northern Germany. His exceptional skill with numbers was clear at a very early age, and in later life he joked that he knew how to count before he could talk. It is said that Goethe wrote and directed little plays for a puppet theater when he was 6 and that Mozart composed his first childish minuets when he was 5, but Gauss corrected an error in his father's payroll accounts at the age of 3. At the age of seven, when he started elementary school, his teacher was amazed when Gauss summed the integers from 1 to 100 instantly by spotting that the sum was 50 pairs of numbers each pair summing to 101.



His long professional life is so filled with accomplishments that it is impossible to give a full account of them in the short space available here. All we can do is simply give a chronology of his almost uncountable discoveries.

**1792–1794:** Gauss reads the works of Newton, Euler, and Lagrange; discovers the prime number theorem (at the age of 14 or 15); invents the method of least squares; conceives the Gaussian law of distribution in the theory of probability.

**1795:** (only 18 years old!) Proves that a regular polygon with  $n$  sides is constructible (by ruler and compass) if and only if  $n$  is the product of a power of 2 and distinct prime numbers of the form  $p_k = 2^{2^k} + 1$ , and completely solves the 2000-year old problem of ruler-and-compass construction of regular polygons. He also discovers the law of quadratic reciprocity.

**1799:** Proves the **fundamental theorem of algebra** in his doctoral dissertation using the then-mysterious complex numbers with complete confidence.

**1801:** Gauss publishes his *Disquisitiones Arithmeticae* in which he creates the modern rigorous approach to mathematics; predicts the exact location of the asteroid Ceres.

**1807:** Becomes professor of astronomy and the director of the new observatory at Göttingen.

**1809:** Publishes his second book, *Theoria motus corporum coelestium*, a major two-volume treatise on the motion of celestial bodies and the bible of planetary astronomers for the next 100 years.

**1812:** Publishes *Disquisitiones generales circa seriem infinitam*, a rigorous treatment of infinite series, and introduces the **hypergeometric function** for the first time, for which he uses the notation  $F(\alpha, \beta; \gamma; z)$ ; an essay on approximate integration.

**1820–1830:** Publishes over 70 papers, including *Disquisitiones generales circa superficies curvas*, in which he creates the intrinsic **differential geometry** of general curved surfaces,

the forerunner of Riemannian geometry and the general theory of relativity. From the 1830s on, Gauss was increasingly occupied with physics, and he enriched every branch of the subject he touched. In the theory of **surface tension**, he developed the fundamental idea of conservation of energy and solved the earliest problem in the **calculus of variations**. In **optics**, he introduced the concept of the focal length of a system of lenses. He virtually created the science of **geomagnetism**, and in collaboration with his friend and colleague Wilhelm Weber he invented the electromagnetic telegraph. In 1839 Gauss published his fundamental paper on the general theory of inverse square forces, which established **potential theory** as a coherent branch of mathematics and in which he established the **divergence theorem**.

Gauss had many opportunities to leave Göttingen, but he refused all offers and remained there for the rest of his life, living quietly and simply, traveling rarely, and working with immense energy on a wide variety of problems in mathematics and its applications. Apart from science and his family—he married twice and had six children, two of whom emigrated to America—his main interests were history and world literature, international politics, and public finance. He owned a large library of about 6000 volumes in many languages, including Greek, Latin, English, French, Russian, Danish, and of course German. His acuteness in handling his own financial affairs is shown by the fact that although he started with virtually nothing, he left an estate over a hundred times as great as his average annual income during the last half of his life. The foregoing list is the published portion of Gauss's total achievement; the unpublished and private part is almost equally impressive. His scientific diary, a little booklet of 19 pages, discovered in 1898, extends from 1796 to 1814 and consists of 146 very concise statements of the results of his investigations, which often occupied him for weeks or months. These ideas were so abundant and so frequent that he physically did not have time to publish them. Some of the ideas recorded in this diary:

**Cauchy Integral Formula:** Gauss discovers it in 1811, 16 years before Cauchy.

**Non-Euclidean Geometry:** After failing to prove Euclid's fifth postulate at the age of 15, Gauss came to the conclusion that the Euclidean form of geometry cannot be the only one possible.

**Elliptic Functions:** Gauss had found many of the results of Abel and Jacobi (the two main contributors to the subject) before these men were born. The facts became known partly through Jacobi himself. His attention was caught by a cryptic passage in the *Disquisitiones*, whose meaning can only be understood if one knows something about elliptic functions. He visited Gauss on several occasions to verify his suspicions and tell him about his own most recent discoveries, and each time Gauss pulled 30-year-old manuscripts out of his desk and showed Jacobi what Jacobi had just shown him. After a week's visit with Gauss in 1840, Jacobi wrote to his brother, "Mathematics would be in a very different position if practical astronomy had not diverted this colossal genius from his glorious career."

A possible explanation for not publishing such important ideas is suggested by his comments in a letter to Bolyai: "It is not knowledge but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment. When I have clarified and exhausted a subject, then I turn away from it in order to go into darkness again." His was the temperament of an explorer who is reluctant to take the time to write an account of his last expedition when he could be starting another. As it was, Gauss wrote a great deal, but to have published every fundamental discovery he made in a form satisfactory to himself would have required several long lifetimes.

---

A third relation can be obtained by making the substitution  $w = (1-z)^{\gamma-\alpha-\beta}u$ . This leads to a hypergeometric equation for  $u$  with  $\alpha_1 = \gamma - \alpha$ ,  $\beta_1 = \gamma - \beta$ , and  $\gamma_1 = \gamma$ . Furthermore,  $w$  is analytic at  $z = 0$ , and  $w(0) = 1$ . We conclude that  $w = F(\alpha, \beta; \gamma; z)$ . We therefore have the identity

$$F(\alpha, \beta; \gamma; z) = (1-z)^{\gamma-\alpha-\beta} F(\gamma-\alpha, \gamma-\beta; \gamma; z). \quad (14.25)$$

To obtain the canonical basis at  $z = 1$ , we make the substitution  $t = 1-z$ , and note that the result is again the HGDE, with  $\alpha_1 = \alpha$ ,  $\beta_1 = \beta$ , and  $\gamma_1 = \alpha + \beta - \gamma + 1$ . It follows from Equation (14.24) that

$$\begin{aligned} w_3(z) &\equiv F(\alpha, \beta; \alpha + \beta - \gamma + 1; 1-z), \\ w_4(z) &\equiv (1-z)^{\gamma-\alpha-\beta} F(\gamma-\beta, \gamma-\alpha; \gamma-\alpha-\beta+1; 1-z) \end{aligned} \quad (14.26)$$

form a canonical basis of solutions to the HGDE at  $z = 1$ .

A symmetry of the hypergeometric function that is easily obtained from the HGDE is

$$F(\alpha, \beta; \gamma; z) = F(\beta, \alpha; \gamma; z). \quad (14.27)$$

The six functions

$$F(\alpha \pm 1, \beta; \gamma; z), \quad F(\alpha, \beta \pm 1; \gamma; z), \quad F(\alpha, \beta; \gamma \pm 1; z)$$

are called hypergeometric functions **contiguous to**  $F(\alpha, \beta; \gamma; z)$ . The discussion above showed how to obtain the basis of solutions at  $z = 1$  from the regular solution to the HDE  $z = 0$ ,  $F(\alpha, \beta; \gamma; z)$ . We can show that the basis of solutions at  $z = \infty$  can also be obtained from the hypergeometric function.

Equation (14.16) suggests a function of the form

$$v(z) = z^r F\left(\alpha_1, \beta_1; \gamma_1; \frac{1}{z}\right) \equiv z^r w\left(\frac{1}{z}\right) \Rightarrow w(z) = z^r v\left(\frac{1}{z}\right), \quad (14.28)$$

where  $r, \alpha_1, \beta_1$ , and  $\gamma_1$  are to be determined. Since  $w(z)$  is a solution of the HGDE,  $v$  will satisfy the following DE (see Problem 14.15):

$$\begin{aligned} z(1-z)v'' + [1-\alpha-\beta-2r-(2-\gamma-2r)z]v' \\ - \left[ r^2 - r + r\gamma - \frac{1}{z}(r+\alpha)(r+\beta) \right] v = 0. \end{aligned} \quad (14.29)$$

This reduces to the HGDE if  $r = -\alpha$  or  $r = -\beta$ . For  $r = -\alpha$ , the parameters become  $\alpha_1 = \alpha$ ,  $\beta_1 = 1 + \alpha - \gamma$ , and  $\gamma_1 = \alpha - \beta + 1$ . For  $r = -\beta$ , the parameters are  $\alpha_1 = \beta$ ,  $\beta_1 = 1 + \beta - \gamma$ , and  $\gamma_1 = \beta - \alpha + 1$ . Thus,

$$\begin{aligned} v_1(z) &= z^{-\alpha} F\left(\alpha, 1 + \alpha - \gamma; \alpha - \beta + 1; \frac{1}{z}\right), \\ v_2(z) &= z^{-\beta} F\left(\beta, 1 + \beta - \gamma; \beta - \alpha + 1; \frac{1}{z}\right) \end{aligned} \quad (14.30)$$

form a canonical basis of solutions for the HGDE that are valid about  $z = \infty$ .

As the preceding discussion suggests, it is possible to obtain many relations among the hypergeometric functions with different parameters and independent variables. In fact, the nineteenth-century mathematician Kummer showed that there are 24 different (but linearly dependent, of course) solutions to the HGDE. These are collectively known as **Kummer's solutions**, and six of them were derived above. Another important relation (shown in Problem 14.16) is that

$$z^{\alpha-\gamma}(1-z)^{\gamma-\alpha-\beta} F\left(\gamma-\alpha, 1-\alpha; 1-\alpha+\beta; \frac{1}{z}\right) \quad (14.31)$$

also solves the HGDE.

Many of the functions that occur in mathematical physics are related to the hypergeometric function. Even some of the common elementary functions can be expressed in terms of the hypergeometric function with appropriate parameters. For example, when  $\beta = \gamma$ , we obtain

$$F(\alpha, \beta; \beta; z) = \sum_{k=0}^{\infty} \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(k+1)} z^k = (1-z)^{-\alpha}.$$

Similarly,  $F(\frac{1}{2}, \frac{1}{2}; \frac{3}{2}; z^2) = \sin^{-1} z/z$ , and  $F(1, 1; 2; -z) = \ln(1+z)/z$ . However, the real power of the hypergeometric function is that it encompasses almost all of the nonelementary functions encountered in physics. Let us look briefly at a few of these. **Jacobi functions** are solutions of the DE

$$(1-x^2)\frac{d^2u}{dx^2} + [\beta-\alpha-(\alpha+\beta+2)x]\frac{du}{dx} + \lambda(\lambda+\alpha+\beta+1)u = 0 \quad (14.32)$$

Defining  $x = 1 - 2z$  changes this DE into the HGDE with parameters  $\alpha_1 = \lambda$ ,  $\beta_1 = \lambda + \alpha + \beta + 1$ , and  $\gamma_1 = 1 + \alpha$ . The solutions of Equation (14.32), called the **Jacobi functions of the first kind**, are, with appropriate normalization,

$$P_{\lambda}^{(\alpha, \beta)}(z) = \frac{\Gamma(\lambda + \alpha + 1)}{\Gamma(\lambda + 1)\Gamma(\alpha + 1)} F\left(-\lambda, \lambda + \alpha + \beta + 1; 1 + \alpha; \frac{1-z}{2}\right).$$

When  $\lambda = n$ , a nonnegative integer, the Jacobi function turns into a polynomial of degree  $n$  with the following expansion:

$$P_n^{(\alpha, \beta)}(z) = \frac{\Gamma(n + \alpha + 1)}{\Gamma(n + 1)\Gamma(n + \alpha + \beta + 1)} \sum_{k=0}^n \frac{\Gamma(n + \alpha + \beta + k + 1)}{\Gamma(\alpha + k + 1)} \left(\frac{z-1}{2}\right)^k$$

These are the Jacobi polynomials discussed in Chapter 7. In fact, the DE satisfied by  $P_n^{(\alpha, \beta)}(x)$  of Chapter 7 is identical to Equation (14.32). Note that the transformation  $x = 1 - 2z$  translates the points  $z = 0$  and  $z = 1$  to the points  $x = 1$  and  $x = -1$ , respectively. Thus the regular singular points of the Jacobi functions of the first kind are at  $\pm 1$  and  $\infty$ .

A second, linearly independent, solution of Equation (14.32) is obtained by using (14.31). These are called the **Jacobi functions of the second kind**:

$$Q_{\lambda}^{(\alpha, \beta)}(z) = \frac{2^{\lambda+\alpha+\beta} \Gamma(\lambda + \alpha + 1) \Gamma(\lambda + \beta + 1)}{\Gamma(2\lambda + \alpha + \beta + 2)(z - 1)^{\lambda+\alpha+1}(z + 1)^{\beta}} \cdot F\left(\lambda + \alpha + 1, \lambda + 1; 2\lambda + \alpha + \beta + 2; \frac{2}{1 - z}\right). \tag{14.33}$$

Gegenbauer functions

**Gegenbauer functions**, or ultraspherical functions, are special cases of Jacobi functions for which  $\alpha = \beta = \mu - \frac{1}{2}$ . They are defined by

$$C_{\lambda}^{\mu}(z) = \frac{\Gamma(\lambda + 2\mu)}{\Gamma(\lambda + 1)\Gamma(2\mu)} F\left(-\lambda, \lambda + 2\mu; \mu + \frac{1}{2}; \frac{1 - z}{2}\right). \tag{14.34}$$

Legendre functions

Note the change in the normalization constant. Linearly independent Gegenbauer functions “of the second kind” can be obtained from the Jacobi functions of the second kind by the substitution  $\alpha = \beta = \mu - \frac{1}{2}$ . Another special case of the Jacobi functions is obtained when  $\alpha = \beta = 0$ . Those obtained from the Jacobi functions of the first kind are called **Legendre functions of the first kind**:

$$P_{\lambda}(z) \equiv P_{\lambda}^{(0,0)}(z) = C_{\lambda}^{1/2} = F\left(-\lambda, \lambda + 1; 1; \frac{1 - z}{2}\right). \tag{14.35}$$

Legendre functions of the second kind are obtained from the Jacobi functions of the second kind in a similar way:

$$Q_{\lambda}(z) = \frac{2^{\lambda} \Gamma^2(\lambda + 1)}{\Gamma(2\lambda + 2)(z - 1)^{\lambda+1}} F\left(\lambda + 1, \lambda + 1; 2\lambda + 2; \frac{2}{1 - z}\right).$$

Other functions derived from the Jacobi functions are obtained similarly (see Chapter 7).

## 14.5 Confluent Hypergeometric Functions

The transformation  $x = 1 - 2z$  translates the regular singular points of the HGDE by a finite amount. Consequently, the new functions still have two regular singular points,  $z = \pm 1$ , in the complex plane. In some physical cases of importance, only the origin, corresponding to  $r = 0$  in spherical coordinates (typically the location of the source of a central force), is the singular point. If we want to obtain a differential equation consistent with such a case, we have to “push” the singular point  $z = 1$  to infinity. This can be achieved by making the substitution  $t = rz$  in the HGDE and taking the limit  $r \rightarrow \infty$ . The substitution yields

$$\frac{d^2 w}{dt^2} + \left(\frac{\gamma}{t} + \frac{1 - \gamma + \alpha + \beta}{t - r}\right) \frac{dw}{dt} + \frac{\alpha\beta}{t(t - r)} w = 0. \tag{14.36}$$

If we blindly take the limit  $r \rightarrow \infty$  with  $\alpha$ ,  $\beta$ , and  $\gamma$  remaining finite, Equation (14.36) reduces to  $\ddot{w} + (\gamma/t)\dot{w} = 0$ , an elementary FODE in  $w$ . To obtain a nonelementary DE, we need to manipulate the parameters, to let some of them tend to infinity. We want  $\gamma$  to remain finite, because otherwise the coefficient of  $dw/dt$  will blow up. We therefore let  $\beta$  or  $\alpha$  tend to infinity. The result will be the same either way because  $\alpha$  and  $\beta$  appear symmetrically in the equation. It is customary to let  $\beta = r \rightarrow \infty$ . In that case, Equation (14.36) becomes

$$\frac{d^2w}{dt^2} + \left(\frac{\gamma}{t} - 1\right) \frac{dw}{dt} - \frac{\alpha}{t}w = 0.$$

confluent  
hypergeometric DE

Multiplying by  $t$  and changing the independent variable back to  $z$  yields

$$zw''(z) + (\gamma - z)w'(z) - \alpha w(z) = 0. \tag{14.37}$$

This is called the **confluent hypergeometric DE (CHGDE)**.

Since  $z = 0$  is still a regular singular point of the CHGDE, we can obtain expansions about that point. The characteristic exponents are 0 and  $1 - \gamma$ , as before. Thus, there is an analytic solution (corresponding to the characteristic exponent 0) to the CHGDE at the origin, which is called the **confluent hypergeometric function** and denoted by  $\Phi(\alpha; \gamma; z)$ . Since  $z = 0$  is the only possible (finite) singularity of the CHGDE,  $\Phi(\alpha; \gamma; z)$  is an entire function.

confluent  
hypergeometric  
function and series

We can obtain the series expansion of  $\Phi(\alpha; \gamma; z)$  directly from Equation (14.22) and the fact that  $\Phi(\alpha; \gamma; z) = \lim_{\beta \rightarrow \infty} F(\alpha, \beta; \gamma; z/\beta)$ . The result is

$$\Phi(\alpha; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)} \sum_{k=0}^{\infty} \frac{\Gamma(\alpha + k)}{\Gamma(k + 1)\Gamma(\gamma + k)} z^k. \tag{14.38}$$

This is called the **confluent hypergeometric series**. An argument similar to the one given in the case of the hypergeometric function shows that

**14.5.1. Box.** *The confluent hypergeometric function  $\Phi(\alpha; \gamma; z)$  reduces to a polynomial when  $\alpha$  is a negative integer.*

A second solution of the CHGDE can be obtained, as for the HGDE. If  $1 - \gamma$  is not an integer, then by taking the limit  $\beta \rightarrow \infty$  of Equation (14.24), we obtain the second solution  $z^{1-\gamma} \Phi(\alpha - \gamma + 1, 2 - \gamma; z)$ . Thus, any solution of the CHGDE can be written as a linear combination of  $\Phi(\alpha; \gamma; z)$  and  $z^{1-\gamma} \Phi(\alpha - \gamma + 1, 2 - \gamma; z)$ .

hydrogen-like atoms

**14.5.2. Example.** The time-independent Schrödinger equation for a central potential, in units in which  $\hbar = m = 1$ , is  $-\frac{1}{2}\nabla^2\Psi + V(r)\Psi = E\Psi$ . For the case of hydrogen-like atoms,  $V(r) = -Ze^2/r$ , where  $Z$  is the atomic number, and the equation reduces to

$$\nabla^2\Psi + \left(2E + \frac{2Ze^2}{r}\right)\Psi = 0.$$

The radial part of this equation is given by Equation (12.14) with  $f(r) = 2E + 2Ze^2/r$ . Defining  $u = rR(r)$ , we may write

$$\frac{d^2u}{dr^2} + \left( \lambda + \frac{a}{r} - \frac{b}{r^2} \right) u = 0, \tag{14.39}$$

where  $\lambda = 2E$ ,  $a = 2Ze^2$ , and  $b = l(l + 1)$ . This equation can be further simplified by defining  $r \equiv kz$  ( $k$  is an arbitrary constant to be determined later):

$$\frac{d^2u}{dz^2} + \left( \lambda k^2 + \frac{ak}{z} - \frac{b}{z^2} \right) u = 0.$$

Choosing  $\lambda k^2 = -\frac{1}{4}$  and introducing  $\alpha \equiv a/(2\sqrt{-\lambda})$  yields

$$\frac{d^2u}{dz^2} + \left( -\frac{1}{4} + \frac{\alpha}{z} - \frac{b}{z^2} \right) u = 0.$$

Equations of this form can be transformed into the CHGDE by making the substitution  $u(z) = z^\mu e^{-\nu z} f(z)$ . It then follows that

$$\frac{d^2f}{dz^2} + \left( \frac{2\mu}{z} - 2\nu \right) \frac{df}{dz} + \left[ -\frac{1}{4} + \frac{\mu(\mu - 1)}{z^2} - \frac{2\mu\nu}{z} + \frac{\alpha}{z} - \frac{b}{z^2} + \nu^2 \right] f = 0.$$

Choosing  $\nu^2 = \frac{1}{4}$  and  $\mu(\mu - 1) = b$  reduces this equation to

$$f'' + \left( \frac{2\mu}{z} - 2\nu \right) f' - \frac{2\mu\nu - \alpha}{z} f = 0,$$

which is in the form of (14.37).

quantization of the energy of the hydrogen atom

On physical grounds, we expect  $u(z) \rightarrow 0$  as  $z \rightarrow \infty$ .<sup>5</sup> Therefore,  $\nu = \frac{1}{2}$ . Similarly, with  $\mu(\mu - 1) = b = l(l + 1)$ , we obtain the two possibilities  $\mu = -l$  and  $\mu = l + 1$ . Again on physical grounds, we demand that  $u(0)$  be finite (the wave function must not blow up at  $r = 0$ ). This implies<sup>6</sup> that  $\mu = l + 1$ . We thus obtain

$$f'' + \left[ \frac{2(l + 1)}{z} - 1 \right] f' - \frac{l + 1 - \alpha}{z} f = 0.$$

Multiplying by  $z$  gives  $zf'' + [2(l + 1) - z]f' - (l + 1 - \alpha)f = 0$ . Comparing this with Equation (14.37) shows that  $f$  is proportional to  $\Phi(l + 1 - \alpha, 2l + 2; z)$ . Thus, the solution of (14.39) can be written as

$$u(z) = Cz^{l+1} e^{-z/2} \Phi(l + 1 - \alpha, 2l + 2; z).$$

An argument similar to that used in Problem 13.20 will reveal that the product  $e^{-z/2} \Phi(l + 1 - \alpha, 2l + 2; z)$  will be infinite unless the power series representing  $\Phi$  terminates (becomes a polynomial). It follows from Box 14.5.1 that this will take place if

$$l + 1 - \alpha = -N \tag{14.40}$$

<sup>5</sup>This is because the volume integral of  $|\Psi|^2$  over all space must be finite. The radial part of this integral is simply the integral of  $r^2 R^2(r) = u^2(r)$ . This latter integral will not be finite unless  $u(\infty) = 0$ .

<sup>6</sup>Recall that  $\mu$  is the exponent of  $z = r/k$ .

for some integer  $N \geq 0$ . In that case we obtain the Laguerre polynomials

$$L_N^j \equiv \frac{\Gamma(N+j+1)}{\Gamma(N+1)\Gamma(j+1)} \Phi(-N, j+1; z), \quad \text{where } j = 2l + 1.$$

Condition (14.40) is the quantization rule for the energy levels of a hydrogen-like atom. Writing everything in terms of the original parameters and defining  $n = N + l + 1$  yields—after restoring all the  $m$ 's and the  $\hbar$ 's—the energy levels of a hydrogen-like atom:

$$E = -\frac{Z^2 m e^4}{2\hbar^2 n^2} = -Z^2 \left(\frac{m c^2}{2}\right) \alpha^2 \frac{1}{n^2},$$

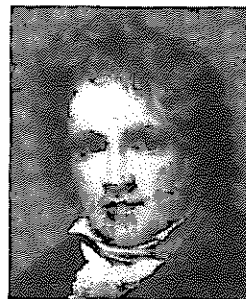
where  $\alpha = e^2/(\hbar c) = 1/137$  is the **fine structure constant**.

The radial wave functions can now be written as

$$R_{n,l}(r) = \frac{u_{n,l}(r)}{r} = C r^l e^{-Zr/(na_0)} \Phi\left(-n+l+1, 2l+2; \frac{2Zr}{na_0}\right),$$

where  $a_0 = \hbar^2/(me^2) = 0.529 \times 10^{-8}$  cm is the Bohr radius. ■

**Friedrich Wilhelm Bessel** (1784–1846) showed no signs of unusual academic ability in school, although he did show a liking for mathematics and physics. He left school intending to become a merchant's apprentice, a desire that soon materialized with a seven-year unpaid apprenticeship with a large mercantile firm in Bremen. The young Bessel proved so adept at accounting and calculation that he was granted a small salary, with raises, after only the first year. An interest in foreign trade led Bessel to study geography and languages at night, astonishingly learning to read and write English in only three months. He also studied navigation in order to qualify as a cargo officer aboard ship, but his innate curiosity soon compelled him to investigate astronomy at a more fundamental level. Still serving his apprenticeship, Bessel learned to observe the positions of stars with sufficient accuracy to determine the longitude of Bremen, checking his results against professional astronomical journals. He then tackled the more formidable problem of determining the orbit of Halley's comet from published observations. After seeing the close agreement between Bessel's calculations and those of Halley, the German astronomer Olbers encouraged Bessel to improve his already impressive work with more observations. The improved calculations, an achievement tantamount to a modern doctoral dissertation, were published with Olbers's recommendation. Bessel later received appointments with increasing authority at observatories near Bremen and in Königsberg, the latter position being accompanied by a professorship. (The title of doctor, required for the professorship, was granted by the University of Göttingen on the recommendation of Gauss.)



Bessel proved himself an excellent observational astronomer. His careful measurements coupled with his mathematical aptitude allowed him to produce accurate positions for a number of previously mapped stars, taking account of instrumental effects, atmospheric refraction, and the position and motion of the observation site. In 1820 he determined the position of the vernal equinox accurate to 0.01 second, in agreement with modern values.



His observation of the variation of the proper motion of the stars Sirius and Procyon led him to posit the existence of nearby, large, low-luminosity stars called dark companions. Between 1821 and 1833 he catalogued the positions of about 75,000 stars, publishing his measurements in detail. One of his most important contributions to astronomy was the determination of the distance to a star using parallax. This method uses triangulation, or the determination of the apparent positions of a distant object viewed from two points a known distance apart, in this case two diametrically opposed points of the Earth's orbit. The angle subtended by the baseline of Earth's orbit, viewed from the star's perspective, is known as the star's parallax. Before Bessel's measurement, stars were assumed to be so distant that their parallaxes were too small to measure, and it was further assumed that bright stars (thought to be nearer) would have the largest parallax. Bessel correctly reasoned that stars with large proper motions were more likely to be nearby ones and selected such a star, 61 Cygni, for his historic measurement. His measured parallax for that star differs by less than 8% from the currently accepted value.

Given such an impressive record in astronomy, it seems only fitting that the famous functions that bear Bessel's name grew out of his investigations of perturbations in planetary systems. He showed that such perturbations could be divided into two effects and treated separately: the obvious direct attraction due to the perturbing planet and an indirect effect caused by the sun's response to the perturber's force. The so-called Bessel functions then appear as coefficients in the series treatment of the indirect perturbation. Although special cases of Bessel functions were discovered by Bernoulli, Euler, and Lagrange, the systematic treatment by Bessel clearly established his preeminence, a fitting tribute to the creator of the most famous functions in mathematical physics.

### 14.5.1 Bessel Functions

Bessel differential equation

The Bessel differential equation is usually written as

$$w'' + \frac{1}{z}w' + \left(1 - \frac{\nu^2}{z^2}\right)w = 0 \quad (14.41)$$

As in the example above, the substitution  $w = z^\mu e^{-\eta z} f(z)$  transforms (14.41) into

$$\frac{d^2 f}{dz^2} + \left(\frac{2\mu + 1}{z} - 2\eta\right) \frac{df}{dz} + \left[\frac{\mu^2 - \nu^2}{z^2} - \frac{\eta(2\mu + 1)}{z} + \eta^2 + 1\right] f = 0,$$

which, if we set  $\mu = \nu$  and  $\eta = i$ , reduces to

$$f'' + \left(\frac{2\nu + 1}{z} - 2i\right) f' - \frac{(2\nu + 1)i}{z} f = 0.$$

Making the further substitution  $2iz = t$ , and multiplying out by  $t$ , we obtain

$$t \frac{d^2 f}{dt^2} + (2\nu + 1 - t) \frac{df}{dt} - \left(\nu + \frac{1}{2}\right) f = 0,$$

which is in the form of (14.37) with  $\alpha = \nu + \frac{1}{2}$  and  $\gamma = 2\nu + 1$ .

Bessel function of the first kind

Thus, solutions of the Bessel equation, Equation (14.41), can be written as constant multiples of  $z^\nu e^{-iz} \Phi(\nu + \frac{1}{2}, 2\nu + 1; 2iz)$ . With proper normalization, we define the **Bessel function of the first kind of order  $\nu$**  as

$$J_\nu(z) = \frac{1}{\Gamma(\nu + 1)} \left(\frac{z}{2}\right)^\nu e^{-iz} \Phi\left(\nu + \frac{1}{2}, 2\nu + 1; 2iz\right). \tag{14.42}$$

Using Equation (14.38) and the expansion for  $e^{-iz}$ , we can show that

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(\nu + k + 1)} \left(\frac{z}{2}\right)^{2k}. \tag{14.43}$$

The second linearly independent solution can be obtained as usual and is proportional to

$$\begin{aligned} z^{1-(2\nu+1)} \left(\frac{z}{2}\right)^\nu e^{-iz} \Phi\left(\nu + \frac{1}{2} - (2\nu + 1) + 1, 2 - (2\nu + 1); 2iz\right) \\ = C \left(\frac{z}{2}\right)^{-\nu} e^{-iz} \Phi\left(-\nu + \frac{1}{2}, -2\nu + 1; 2iz\right) = C J_{-\nu}(z), \end{aligned}$$

provided that  $1 - \gamma = 1 - (2\nu + 1) = -2\nu$  is not an integer. When  $\nu$  is an integer,  $J_{-n}(z) = (-1)^n J_n(z)$  (see Problem 14.25). Thus, when  $\nu$  is a noninteger, the most general solution is of the form  $AJ_\nu(z) + BJ_{-\nu}(z)$ .

Bessel function of the second kind, or Neumann function

How do we find a second linearly independent solution when  $\nu$  is an integer  $n$ ? We first define

$$Y_\nu(z) = \frac{J_\nu(z) \cos \nu\pi - J_{-\nu}(z)}{\sin \nu\pi}, \tag{14.44}$$

called the Bessel function of the second kind, or the **Neumann function**. For noninteger  $\nu$  this is simply a linear combination of the two linearly independent solutions. For integer  $\nu$  the function is indeterminate. Therefore, we use l'Hôpital's rule and define

$$Y_n(z) \equiv \lim_{\nu \rightarrow n} Y_\nu(z) = \frac{1}{\pi} \lim_{\nu \rightarrow n} \left[ \frac{\partial J_\nu}{\partial \nu} - (-1)^n \frac{\partial J_{-\nu}}{\partial \nu} \right].$$

Equation (14.43) yields

$$\frac{\partial J_\nu}{\partial \nu} = J_\nu(z) \ln \left(\frac{z}{2}\right) - \left(\frac{z}{2}\right)^\nu \sum_{k=0}^{\infty} (-1)^k \frac{\Psi(\nu + k + 1)}{k! \Gamma(\nu + k + 1)} \left(\frac{z}{2}\right)^{2k},$$

where  $\Psi(z) = (d/dz) \ln \Gamma(z)$ . Similarly,

$$\frac{\partial J_{-\nu}}{\partial \nu} = -J_{-\nu}(z) \ln \left(\frac{z}{2}\right) + \left(\frac{z}{2}\right)^{-\nu} \sum_{k=0}^{\infty} \frac{\Psi(-\nu + k + 1)}{k! \Gamma(-\nu + k + 1)} \left(\frac{z}{2}\right)^{2k}.$$

Substituting these expressions in the definition of  $Y_n(z)$  and using  $J_{-n}(z) = (-1)^n J_n(z)$ , we obtain

$$Y_n(z) = \frac{2}{\pi} J_n(z) \ln \left( \frac{z}{2} \right) - \frac{1}{\pi} \left( \frac{z}{2} \right)^n \sum_{k=0}^{\infty} (-1)^k \frac{\Psi(n+k+1)}{k! \Gamma(n+k+1)} \left( \frac{z}{2} \right)^{2k} - \frac{1}{\pi} \left( \frac{z}{2} \right)^{-n} (-1)^n \sum_{k=0}^{\infty} (-1)^k \frac{\Psi(k-n+1)}{k! \Gamma(k-n+1)} \left( \frac{z}{2} \right)^{2k}. \quad (14.45)$$

The natural log term is indicative of the solution suggested by Theorem 14.2.6. Since  $Y_\nu(z)$  is linearly independent of  $J_\nu(z)$  for any  $\nu$ , integer or noninteger, it is convenient to consider  $\{J_\nu(z), Y_\nu(z)\}$  as a basis of solutions for the Bessel equation.

Another basis of solutions is defined as

$$H_\nu^{(1)}(z) = J_\nu(z) + iY_\nu(z), \quad H_\nu^{(2)}(z) = J_\nu(z) - iY_\nu(z), \quad (14.46)$$

Bessel function of the third kind or Hankel function

which are called Bessel functions of the third kind, or **Hankel functions**.

Replacing  $z$  by  $iz$  in the Bessel equation yields

$$\frac{d^2 w}{dz^2} + \frac{1}{z} \frac{dw}{dz} - \left( 1 + \frac{\nu^2}{z^2} \right) w = 0,$$

whose basis of solutions consists of multiples of  $J_\nu(iz)$  and  $J_{-\nu}(iz)$ . Thus, the **modified Bessel functions of the first kind** are defined as

$$I_\nu(z) \equiv e^{-i\pi\nu/2} J_\nu(iz) = \left( \frac{z}{2} \right)^\nu \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(\nu+k+1)} \left( \frac{z}{2} \right)^{2k}.$$

Similarly, the **modified Bessel functions of the second kind** are defined as

$$K_\nu(z) = \frac{\pi}{2 \sin \nu\pi} [I_{-\nu}(z) - I_\nu(z)].$$

When  $\nu$  is an integer  $n$ ,  $I_n = I_{-n}$ , and  $K_n$  is indeterminate. Thus, we define  $K_n(z)$  as  $\lim_{\nu \rightarrow n} K_\nu(z)$ . This gives

$$K_n(z) = \frac{(-1)^n}{2} \lim_{\nu \rightarrow n} \left[ \frac{\partial I_{-\nu}}{\partial \nu} - \frac{\partial I_\nu}{\partial \nu} \right],$$

which has the power-series representation

$$K_n(z) = (-1)^{n+1} I_n(z) \ln \left( \frac{z}{2} \right) + \frac{1}{2} (-1)^n \left( \frac{z}{2} \right)^n \sum_{k=0}^{\infty} \frac{\Psi(n+k+1)}{k! \Gamma(n+k+1)} \left( \frac{z}{2} \right)^{2k} + \frac{1}{2} (-1)^n \left( \frac{z}{2} \right)^{-n} \sum_{k=0}^{\infty} \frac{\Psi(k-n+1)}{k! \Gamma(k-n+1)} \left( \frac{z}{2} \right)^{2k}.$$

recurrence relation  
for solutions of the  
Bessel DE

We can obtain a recurrence relation for solutions of the Bessel equation as follows. If  $Z_\nu(z)$  is a solution of order  $\nu$ , then (see Problem 14.28)

$$Z_{\nu+1} = C_1 z^\nu \frac{d}{dz} [z^{-\nu} Z_\nu(z)] \quad \text{and} \quad Z_{\nu-1} = C_2 z^{-\nu} \frac{d}{dz} [z^\nu Z_\nu(z)].$$

If the constants are chosen in such a way that  $Z_\nu$ ,  $Z_{-\nu}$ ,  $Z_{\nu+1}$ , and  $Z_{\nu-1}$  satisfy their appropriate series expansions, then  $C_1 = -1$  and  $C_2 = 1$ . Carrying out the differentiation in the equations for  $Z_{\nu+1}$  and  $Z_{\nu-1}$ , we obtain

$$Z_{\nu+1} = \frac{\nu}{z} Z_\nu - \frac{dZ_\nu}{dz}, \quad Z_{\nu-1} = \frac{\nu}{z} Z_\nu + \frac{dZ_\nu}{dz}. \quad (14.47)$$

Adding these two equations yields the recursion relation

$$Z_{\nu-1}(z) + Z_{\nu+1}(z) = \frac{2\nu}{z} Z_\nu(z), \quad (14.48)$$

where  $Z_\nu(z)$  can be any of the three kinds of Bessel functions.

## 14.6 Problems

**14.1.** Show that the solution of  $w' + w/z^2 = 0$  has an essential singularity at  $z = 0$ .

**14.2.** Derive the recursion relation of Equation (14.7) and express it in terms of the indicial polynomial, as in Equation (14.9).

**14.3.** Find the characteristic exponent associated with the solution of  $w'' + p(z)w' + q(z)w = 0$  at an ordinary point [a point at which  $p(z)$  and  $q(z)$  have no poles]. How many solutions can you find?

**14.4.** The Laplace equation in electrostatics when separated in spherical coordinates yields a DE in the radial coordinate given by

$$\frac{d}{dx} \left( x^2 \frac{dy}{dx} \right) - n(n+1)y = 0 \quad \text{for } n \geq 0.$$

Starting with an infinite series of the form (14.6), show that the two independent solutions of this ODE are of the form  $x^n$  and  $x^{-n-1}$ .

**14.5.** Find the indicial polynomial, characteristic exponents, and recursion relation at both of the regular singular points of the Legendre equation,

$$w'' - \frac{2z}{1-z^2}w' + \frac{\alpha}{1-z^2}w = 0.$$

What is  $a_k$ , the coefficient of the Laurent expansion, for the point  $z = +1$ ?

**14.6.** Show that the substitution  $z = 1/t$  transforms Equation (14.13) into Equation (14.14).

**14.7.** Obtain the indicial polynomial of Equation (14.14) for expansion about  $t = 0$ .

**14.8.** Show that Riemann DE represents the most general second order Fuchsian DE.

**14.9.** Derive the indicial equation for the Riemann DE.

**14.10.** Show that the transformation  $v(z) = z^\lambda(z-1)^\mu w(z)$  changes the pairs of characteristic exponents  $(\lambda_1, \lambda_2)$ ,  $(\mu_1, \mu_2)$ , and  $(\nu_1, \nu_2)$  for the Riemann DE to  $(\lambda_1 + \lambda, \lambda_2 + \lambda)$ ,  $(\mu_1 + \mu, \mu_2 + \mu)$ , and  $(\nu_1 - \lambda - \mu, \nu_2 - \lambda - \mu)$ .

**14.11.** Go through the steps leading to Equations (14.24), (14.25), and (14.26).

**14.12.** Show that the elliptic function of the first kind, defined as

$$K(z) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - z^2 \sin^2 \theta}},$$

can be expressed as  $(\pi/2)F(\frac{1}{2}, \frac{1}{2}; 1; z^2)$ .

**14.13.** By differentiating the hypergeometric series, show that

$$\frac{d^n}{dz^n} F(\alpha, \beta; \gamma; z) = \frac{\Gamma(\alpha + n)\Gamma(\beta + n)\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma + n)} F(\alpha + n, \beta + n; \gamma + n; z).$$

**14.14.** Use direct substitution in the hypergeometric series to show that

$$F(-\alpha, \beta; \beta; -z) = (1+z)^\alpha, \quad F(\frac{1}{2}, \frac{1}{2}; \frac{3}{2}; z^2) = \frac{1}{z} \sin^{-1} z,$$

$$F(1, 1; 2; -z) = \frac{1}{z} \ln(1+z).$$

**14.15.** Show that the substitution  $v(z) = z^r w(1/z)$  [see Equation (14.28)] transforms the HGDE into Equation (14.29).

**14.16.** Consider the function  $v(z) \equiv z^r(1-z)^s F(\alpha_1, \beta_1; \gamma_1; 1/z)$  and assume that it is a solution of HGDE. Find a relation among  $r, s, \alpha_1, \beta_1$ , and  $\gamma_1$  such that  $v(z)$  is written in terms of three parameters rather than five. In particular, show that one possibility is

$$v(z) = z^{\alpha-\gamma}(1-z)^{\gamma-\alpha-\beta} F(\gamma-\alpha, 1-\alpha; 1+\beta-\alpha; 1/z).$$

Find all such possibilities.

**14.17.** Show that the Jacobi functions are related to the hypergeometric functions.

**14.18.** Derive the expression for the Jacobi function of the second kind as given in Equation (14.33).

**14.19.** Show that  $z = \infty$  is not a regular singular point of the CHGDE.

**14.20.** Derive the confluent hypergeometric series from hypergeometric series.

**14.21.** Show that the **Weber–Hermite equation**,  $u'' + (\nu + \frac{1}{2} - \frac{1}{4}z^2)u = 0$  can be transformed into the CHGDE. Hint: Make the substitution  $u(z) = \exp(-\frac{1}{4}z^2)v(z)$ .

**14.22.** The linear combination

$$\Psi(\alpha, \gamma; z) \equiv \frac{\Gamma(1-\gamma)}{\Gamma(\alpha-\gamma+1)}\Phi(\alpha, \gamma; z) + \frac{\Gamma(\gamma-1)}{\Gamma(\alpha)}z^{1-\gamma}\Phi(\alpha-\gamma+1, 2-\gamma; z)$$

is also a solution of the CHGDE. Show that the Hermite polynomials can be written as

$$H_n\left(\frac{z}{\sqrt{2}}\right) = 2^n \Psi\left(-\frac{n}{2}, \frac{1}{2}; \frac{z^2}{2}\right).$$

**14.23.** Verify that the error function  $\operatorname{erf}(z) = \int_0^z e^{-t^2} dt$  satisfies the relation  $\operatorname{erf}(z) = z\Phi(\frac{1}{2}, \frac{3}{2}; -z^2)$ .

**14.24.** Derive the series expansion of the Bessel function of the first kind from that of the confluent hypergeometric series and the expansion of the exponential. Check your answer by obtaining the same result by substituting the power series directly in the Bessel DE.

**14.25.** Show that  $J_{-n}(z) = (-1)^n J_n(z)$ . Hint: Let  $\nu = -n$  in the expansion of  $J_\nu(z)$  and use  $\Gamma(m) = \infty$  for a nonpositive integer  $m$ .

**14.26.** In a potential-free region, the radial part of the Schrödinger equation reduces to

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left[ \lambda - \frac{\alpha}{r^2} \right] R = 0.$$

Write the solutions of this DE in terms of Bessel functions. Hint: Substitute  $R = u/\sqrt{r}$ . These solutions are called **spherical Bessel functions**.

spherical Bessel  
functions

**14.27.** Theorem 14.2.6 states that under certain conditions, linearly independent solutions of SOLDE at regular singular points exist even though the difference between the characteristic exponents is an integer. An example is the case of Bessel functions of half-odd-integer orders. Evaluate the Wronskian of the two linearly independent solutions,  $J_\nu$  and  $J_{-\nu}$ , of the Bessel equation and show that it vanishes *only if  $\nu$  is an integer*. This shows, in particular, that  $J_{n+1/2}$  and  $J_{-n-1/2}$  are linearly independent. Hint: Consider the value of the Wronskian at  $z = 0$ , and use the formula  $\Gamma(\nu)\Gamma(1-\nu) = \pi/\sin \nu\pi$ .

**14.28.** Show that  $z^{\pm\nu}(d/dz)[z^{\mp\nu}Z_\nu(z)]$  is a solution of the Bessel equation of order  $\nu \pm 1$  if  $Z_\nu$  is a solution of order  $\nu$ .

**14.29.** Use the recursion relation of Equation (14.47) to prove that

$$\left(\frac{1}{z} \frac{d}{dz}\right)^m [z^\nu Z_\nu(z)] = z^{\nu-m} Z_{\nu-m}(z),$$

$$\left(\frac{1}{z} \frac{d}{dz}\right)^m [z^{-\nu} Z_\nu(z)] = (-1)^m z^{-\nu-m} Z_{\nu+m}(z).$$

**14.30.** Using the series expansion of the Bessel function, write  $J_{1/2}(z)$  and  $J_{-1/2}(z)$  in terms of elementary functions. Hint: First show that  $\Gamma(k + \frac{3}{2}) = \sqrt{\pi} (2k + 1)! / (k! 2^{2k+1})$ .

**14.31.** From the results of the previous two problems, derive the relations

$$J_{-n-1/2}(z) = \sqrt{\frac{2}{\pi}} z^{n+1/2} \left(\frac{1}{z} \frac{d}{dz}\right)^n \left(\frac{\cos z}{z}\right),$$

$$J_{n+1/2}(z) = \sqrt{\frac{2}{\pi}} z^{n+1/2} \left(-\frac{1}{z} \frac{d}{dz}\right)^n \left(\frac{\sin z}{z}\right).$$

**14.32.** Obtain the following integral identities:

$$(a) \int z^{\nu+1} J_\nu(z) dz = z^{\nu+1} J_{\nu+1}(z).$$

$$(b) \int z^{-\nu+1} J_\nu(z) dz = -z^{-\nu+1} J_{\nu-1}(z).$$

$$(c) \int z^{\mu+1} J_\nu(z) dz = z^{\mu+1} J_{\nu+1}(z) + (\mu - \nu) z^\mu J_\nu(z) \\ - (\mu^2 - \nu^2) \int z^{\mu-1} J_\nu(z) dz,$$

and evaluate

$$(d) \int z^3 J_0(z) dz.$$

Hint: For (c) write  $z^{\mu+1} = z^{\mu-\nu} z^{\nu+1}$  and use integration by parts.

**14.33.** Use Theorem 14.2.6 and the fact that  $J_n(z)$  is entire to show that for integer  $n$ , a second solution to the Bessel equation exists and can be written as  $Y_n(z) = J_n(z)[f_n(z) + C_n \ln z]$ , where  $f_n(z)$  is analytic about  $z = 0$ .

**14.34.** (a) Show that the Wronskian  $W(J_\nu, Z; z)$  of  $J_\nu$  and any other solution  $Z$  of the Bessel equation, satisfies the equation

$$\frac{d}{dz}[zW(J_\nu, Z; z)] = 0.$$

(b) For some constant  $A$ , show that

$$\frac{d}{dz} \left[ \frac{Z}{J_\nu} \right] = \frac{W(z)}{J_\nu^2(z)} = \frac{A}{zJ_\nu^2(z)}.$$

(c) Show that the general second solution of the Bessel equation can be written as

$$Z_\nu(z) = J_\nu(z) \left[ B + A \int \frac{dz}{zJ_\nu^2(z)} \right].$$

**14.35.** Spherical Bessel functions are defined by

$$f_l(z) \equiv \sqrt{\frac{\pi}{2}} \left( \frac{Z_{l+1/2}(z)}{\sqrt{z}} \right).$$

Let  $f_l(z)$  denote a spherical Bessel function "of some kind." By direct differentiation and substitution in the Bessel equation, show that

$$(a) \frac{d}{dz} [z^{l+1} f_l(z)] = z^{l+1} f_{l-1}(z). \quad (b) \frac{d}{dz} [z^{-l} f_l(z)] = -z^{-l} f_{l+1}(z).$$

(c) Combine the results of parts (a) and (b) to derive the recursion relations

$$f_{l-1}(z) + f_{l+1}(z) = \frac{2l+1}{z} f_l(z), \quad l f_{l-1}(z) - (l+1) f_{l+1}(z) = (2l+1) \frac{df_l}{dz}.$$

**14.36.** Show that  $W(J_\nu, Y_\nu; z) = 2/(\pi z)$ ,  $W(H_\nu^{(1)}, H_\nu^{(2)}; z) = 4/(i\pi z)$ . Hint: Use Problem 14.34.

**14.37.** Verify the following relations:

$$(a) Y_{n+1/2}(z) = (-1)^{n+1} J_{-n-1/2}(z), \quad Y_{-n-1/2}(z) = (-1)^n J_{n+1/2}(z).$$

$$(b) Y_{-\nu}(z) = \sin \nu\pi J_\nu(z) + \cos \nu\pi Y_\nu(z) = \frac{J_\nu(z) - \cos \nu\pi J_{-\nu}(z)}{\sin \nu\pi}.$$

$$(c) Y_{-n}(z) = (-1)^n Y_n(z) \quad \text{in the limit } \nu \rightarrow n \text{ in part (b).}$$

**14.38.** Use the recurrence relation for the Bessel function to show that  $J_1(z) = -J_0'(z)$ .

**14.39.** Let  $u = J_\nu(\lambda z)$  and  $v = J_\nu(\mu z)$ . Multiply the Bessel DE for  $u$  by  $v/z$  and that of  $v$  by  $u/z$ . Subtract the two equations to obtain

$$(\lambda^2 - \mu^2)zuv = \frac{d}{dz} \left[ z \left( u \frac{dv}{dz} - v \frac{du}{dz} \right) \right].$$

(a) Write the above equation in terms of  $J_\nu(\lambda z)$  and  $J_\nu(\mu z)$  and integrate both sides with respect to  $z$ .



(b) Now divide both sides by  $\lambda^2 - \mu^2$  and take the limit as  $\mu \rightarrow \lambda$ . You will need to use L'Hôpital's rule.

(c) Substitute for  $J_v''(\lambda z)$  from the Bessel DE and simplify to get

$$\int z [J_v(\lambda z)]^2 dz = \frac{z^2}{2} \left\{ [J_v'(\lambda z)]^2 + \left( 1 - \frac{v^2}{\lambda^2 z^2} \right) [J_v(\lambda z)]^2 \right\}.$$

(d) Finally, let  $\lambda = x_{vn}/a$ , where  $x_{vn}$  is the  $n$ th root of  $J_v$ , and use Equation (14.47) to arrive at

$$\int_0^a z J_v^2 \left( \frac{x_{vn}}{a} z \right) dz = \frac{a^2}{2} J_{v+1}^2(x_{vn}).$$

**14.40.** The generating function for Bessel functions of integer order is  $\exp[\frac{1}{2}z(t - 1/t)]$ . To see this, rewrite the generating function as  $e^{zt/2}e^{-z/2t}$ , expand both factors, and write the product as powers of  $t^n$ . Now show that the coefficient of  $t^n$  is simply  $J_n(z)$ . Finally, use  $J_{-n}(z) = (-1)^n J_n(z)$  to derive the formula  $\exp[\frac{1}{2}z(t - 1/t)] = \sum_{n=-\infty}^{\infty} J_n(z)t^n$ .

**14.41.** Make the substitutions  $z = \beta t^\gamma$  and  $w = t^\alpha u$  to transform the Bessel DE into  $t^2 \frac{d^2 u}{dt^2} + (2\alpha + 1)t \frac{du}{dt} + (\beta^2 \gamma^2 t^{2\gamma} + \alpha^2 - v^2 \gamma^2)u = 0$ . Now show that **Airy's differential equation**  $\ddot{u} - tu = 0$  has solutions of the form  $J_{1/3}(\frac{2}{3}it^{3/2})$  and  $J_{-1/3}(\frac{2}{3}it^{3/2})$ .

**14.42.** Show that the general solution of  $\frac{d^2 w}{dt^2} + \frac{e^{2t} - v^2}{t^4} w = 0$  is  $w = t[AJ_v(e^{1/t}) + BY_v(e^{1/t})]$ .

**14.43.** Transform  $dw/dz + w^2 + z^m = 0$  by making the substitution  $w = (d/dz) \ln v$ . Now make the further substitutions

$$v = u\sqrt{z} \quad \text{and} \quad t = \frac{2}{m+2} z^{1+(1/2)m}$$

to show that the new DE can be transformed into a Bessel equation of order  $1/(m+2)$ .

**14.44.** Starting with the relation

$$\exp[\frac{1}{2}x(t - 1/t)] \exp[\frac{1}{2}y(t - 1/t)] = \exp[\frac{1}{2}(x + y)(t - 1/t)]$$

and the fact that the exponential function is the generating function for  $J_n(z)$ , prove the "addition theorem" for Bessel functions:

$$J_n(x + y) = \sum_{k=-\infty}^{\infty} J_k(x) J_{n-k}(y).$$

## Additional Reading

1. Birkhoff, G. and Rota, G.-C. *Ordinary Differential Equations*, 3rd ed., Wiley, 1978. The first two sections of this chapter closely follow their presentation.
2. Dennery, P. and Krzywicki, A. *Mathematics for Physicists*, Harper and Row, 1967.
3. Watson, G. *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, 1952. As the name suggests, the definitive text and reference on Bessel functions.

# Integral Transforms and Differential Equations

The discussion in Chapter 14 introduced a general method of solving differential equations by power series—also called the Frobenius method—which gives a solution that converges within a circle of convergence. In general, this circle of convergence may be small; however, the function represented by the power series can be analytically continued using methods presented in Chapter 11.

This chapter, which is a bridge between differential equations and operators on Hilbert spaces (to be developed in the next part), introduces another method of solving DEs, which uses **integral transforms** and incorporates the analytic continuation automatically. The integral transform of a function  $v$  is another function  $u$  given by

$$u(z) = \int_C K(z, t)v(t) dt, \quad (15.1)$$

kernel of integral transforms

where  $C$  is a convenient contour, and  $K(z, t)$ , called the **kernel** of the integral transform, is an appropriate function of two complex variables.

examples of integral transforms

**15.0.1. Example.** Let us consider some examples of integral transforms.

(a) The *Fourier transform* is familiar from the discussion of Chapter 8. The kernel is

$$K(x, y) = e^{ixy}.$$

(b) The **Laplace transform** is used frequently in electrical engineering. Its kernel is

$$K(x, y) = e^{-xy}.$$

(c) The **Euler transform** has the kernel

$$K(x, y) = (x - y)^y.$$

(d) The **Mellin transform** has the kernel

$$K(x, y) = G(x^y),$$

where  $G$  is an arbitrary function. Most of the time  $K(x, y)$  is taken to be simply  $x^y$ .

(e) The **Hankel transform** has the kernel

$$K(x, y) = yJ_n(xy),$$

where  $J_n$  is the  $n$ th-order Bessel function.

(f) A transform that is useful in connection with the Bessel equation has the kernel

$$K(x, y) = \left(\frac{x}{2}\right)^y e^{y-x^2/4y}. \quad \blacksquare$$

Strategy for solving  
DEs using integral  
transforms

The idea behind using integral transform is to write the solution  $u(z)$  of a DE in  $z$  in terms of an integral such as Equation (15.1) and choose  $v$  and the kernel in such a way as to render the DE more manageable. Let  $L_z$  be a differential operator (DO) in the variable  $z$ . We want to determine  $u(z)$  such that  $L_z[u] = 0$ , or equivalently, such that  $\int_C L_z[K(z, t)]v(t) dt = 0$ . Suppose that we can find  $M_t$ , a DO in the variable  $t$ , such that  $L_z[K(z, t)] = M_t[K(z, t)]$ . Then the DE becomes  $\int_C (M_t[K(z, t)])v(t) dt = 0$ . If  $C$  has  $a$  and  $b$  as initial and final points ( $a$  and  $b$  may be equal), then the Lagrange identity [see Equation (13.24)] yields

$$0 = L_z[u] = \int_a^b K(z, t)M_t^\dagger[v(t)] dt + Q[K, v]_a^b,$$

where  $Q[K, v]$  is the “surface term.” If  $v(t)$  and the contour  $C$  (or  $a$  and  $b$ ) are chosen in such a way that

$$Q[K, v]_a^b = 0 \quad \text{and} \quad M_t^\dagger[v(t)] = 0, \tag{15.2}$$

the problem is solved. The trick is to find an  $M_t$  such that Equation (15.2) is easier to solve than the original equation,  $L_z[u] = 0$ . This in turn demands a clever choice of the kernel,  $K(z, t)$ . This chapter discusses how to solve some common differential equations of mathematical physics using the general idea presented above.

## 15.1 Integral Representation of the Hypergeometric Function

Recall that for the hypergeometric function, the differential operator is

$$L_z = z(1-z)\frac{d^2}{dz^2} + [\gamma - (\alpha + \beta + 1)z]\frac{d}{dz} - \alpha\beta$$

For such operators—whose coefficient functions are polynomials—the proper choice for  $K(z, t)$  is the **Euler kernel**,  $(z - t)^s$ . Applying  $L_z$  to this kernel and

rearranging terms, we obtain

$$\begin{aligned} \mathbf{L}_z[K(z, t)] = & \{z^2[-s(s-1) - s(\alpha + \beta + 1) - \alpha\beta] + z[s(s-1) + s\gamma \\ & + st(\alpha + \beta + 1) + 2\alpha\beta t] - \gamma st - \alpha\beta t^2\}(z-t)^{s-2}. \end{aligned} \quad (15.3)$$

Note that except for a multiplicative constant,  $K(z, t)$  is symmetric in  $z$  and  $t$ . This suggests that the general form of  $\mathbf{M}_t$  may be chosen to be the same as that of  $\mathbf{L}_z$  except for the interchange of  $z$  and  $t$ . If we can manipulate the parameters in such a way that  $\mathbf{M}_t$  becomes simple, then we have a chance of solving the problem. For instance, if  $\mathbf{M}_t$  has the form of  $\mathbf{L}_z$  with the constant term absent, then the hypergeometric DE effectively reduces to a FODE (in  $dv/dt$ ). Let us exploit this possibility.

The general form of the  $\mathbf{M}_t$  that we are interested in is

$$\mathbf{M}_t = p_2(t) \frac{d^2}{dt^2} + p_1(t) \frac{d}{dt},$$

i.e., with no  $p_0$  term. By applying  $\mathbf{M}_t$  to  $K(z, t) = (z-t)^s$  and setting the result equal to the RHS of Equation (15.3), we obtain

$$\begin{aligned} s(s-1)p_2 - p_1sz + p_1st = & z^2[-s(s-1) - s(\alpha + \beta + 1) - \alpha\beta] \\ & + z[s(s-1) + s\gamma + st(\alpha + \beta + 1) + 2\alpha\beta t] \\ & - \gamma st - \alpha\beta t^2, \end{aligned}$$

for which the coefficients of equal powers of  $z$  on both sides must be equal:

$$\begin{aligned} -s(s-1) - s(\alpha + \beta + 1) - \alpha\beta = 0 & \Rightarrow s = -\alpha \quad \text{or} \quad s = -\beta, \\ -p_1s = s(s-1) + s\gamma + st(\alpha + \beta + 1) + 2\alpha\beta t, \\ s(s-1)p_2 + p_1st = -\gamma st - \alpha\beta t^2. \end{aligned}$$

If we choose  $s = -\alpha$  ( $s = -\beta$  leads to an equivalent representation), the coefficient functions of  $\mathbf{M}_t$  will be completely determined. In fact, the second equation gives  $p_1(t)$ , and the third determines  $p_2(t)$ . We finally obtain

$$p_1(t) = \alpha + 1 - \gamma + t(\beta - \alpha - 1), \quad p_2(t) = t - t^2,$$

and

$$\mathbf{M}_t = (t - t^2) \frac{d^2}{dt^2} + [\alpha + 1 - \gamma + t(\beta - \alpha - 1)] \frac{d}{dt}, \quad (15.4)$$

which, according to Equation (13.20), yields the following DE for the adjoint:

$$\mathbf{M}_t^\dagger[v] = \frac{d^2}{dt^2}[(t - t^2)v] - \frac{d}{dt}\{[\alpha - \gamma + 1 + t(\beta - \alpha - 1)]v\} = 0. \quad (15.5)$$

The solution to this equation is  $v(t) = Ct^{\alpha-\gamma}(t-1)^{\gamma-\beta-1}$  (see Problem 15.5). We also need the surface term,  $Q[K, v]$ , in the Lagrange identity (see Problem 15.6 for details):  $Q[K, v](t) = C\alpha t^{\alpha-\gamma+1}(t-1)^{\gamma-\beta}(z-t)^{-\alpha-1}$ .

Finally, we need a specification of the contour. For different contours we will get different solutions. The contour chosen must, of course, have the property that  $Q[K, v]$  vanishes as a result of the integration. There are two possibilities: Either the contour is closed [ $a = b$  in (15.2)] or  $a \neq b$  but  $Q[K, v]$  takes on the same value at  $a$  and at  $b$ .

Let us consider the second of these possibilities. Clearly,  $Q[K, v](t)$  vanishes at  $t = 1$  if  $\text{Re}(\gamma) > \text{Re}(\beta)$ . Also, as  $t \rightarrow \infty$ ,

$$Q[K, v](t) \rightarrow (-1)^{-\alpha-1} C\alpha t^{\alpha-\gamma+1} t^{\gamma-\beta} t^{-\alpha-1} = (-1)^{-\alpha-1} C\alpha t^{-\beta},$$

which vanishes if  $\text{Re}(\beta) > 0$ . We thus take  $a = 1$  and  $b = \infty$ , and assume that  $\text{Re}(\gamma) > \text{Re}(\beta) > 0$ . It then follows that

$$u(z) = \int_a^b K(z, t)v(t) dt = C' \int_1^\infty (t-z)^{-\alpha} t^{\alpha-\gamma} (t-1)^{\gamma-\beta-1} dt. \tag{15.6}$$

The constant  $C'$  can be determined to be  $\Gamma(\gamma)/[\Gamma(\beta)\Gamma(\gamma-\beta)]$  (see Problem 15.7). Therefore,

$$u(z) \equiv F(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \int_1^\infty (t-z)^{-\alpha} t^{\alpha-\gamma} (t-1)^{\gamma-\beta-1} dt.$$

Euler formula for the hypergeometric function

It is customary to change the variable of integration from  $t$  to  $1/t$ . The resulting expression is called the **Euler formula** for the hypergeometric function:

$$F(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \int_0^1 (1-tz)^{-\alpha} t^{\beta-1} (1-t)^{\gamma-\beta-1} dt. \tag{15.7}$$

Note that the term  $(1-tz)^{-\alpha}$  in the integral has two branch points in the  $z$ -plane, one at  $z = 1/t$  and the other at  $z = \infty$ . Therefore, we cut the  $z$ -plane from  $z_1 = 1/t$ , a point on the positive real axis, to  $z_2 = \infty$ . Since  $0 \leq t \leq 1$ ,  $z_1$  is somewhere in the interval  $[1, \infty)$ . To ensure that the cut is applicable for all values of  $t$ , we take  $z_1 = 1$  and cut the plane along the positive real axis. It follows that Equation (15.7) is well behaved as long as

$$0 < \arg(1-z) < 2\pi. \tag{15.8}$$

We could choose a different contour, which, in general, would lead to a different solution. The following example illustrates one such choice.

**15.1.1. Example.** First note that  $Q[K, v]$  vanishes at  $t = 0$  and  $t = 1$  as long as  $\text{Re}(\gamma) > \text{Re}(\beta)$  and  $\text{Re}(\alpha) > \text{Re}(\gamma) - 1$ . Hence, we can choose the contour to start at  $t = 0$  and end

at  $t = 1$ . We then have

$$\begin{aligned} w(z) &= C'' \int_0^1 (z-t)^{-\alpha} t^{\alpha-\gamma} (1-t)^{\gamma-\beta-1} dt \\ &= C'' z^{-\alpha} \int_0^1 \left(1 - \frac{t}{z}\right)^{-\alpha} t^{\alpha-\gamma} (1-t)^{\gamma-\beta-1} dt. \end{aligned} \quad (15.9)$$

To see the relation between  $w(z)$  and the hypergeometric function, expand  $(1 - t/z)^{-\alpha}$  in the integral to get

$$w(z) = C'' z^{-\alpha} \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\Gamma(n+1)} \left(\frac{1}{z}\right)^n \int_0^1 t^{\alpha+n-\gamma} (1-t)^{\gamma-\beta-1} dt. \quad (15.10)$$

Now evaluate the integral by changing  $t$  to  $1/t$  and using Equations (11.19) and (11.17). This changes the integral to

$$\int_1^{\infty} t^{-\alpha-n-1+\beta} (t-1)^{\gamma-\beta-1} dt = \frac{\Gamma(\alpha+n+1-\gamma)\Gamma(\gamma-\beta)}{\Gamma(\alpha+n+1-\beta)}$$

Substituting this in Equation (15.10), we obtain

$$\begin{aligned} w(z) &= \frac{C''}{\Gamma(\alpha)} \Gamma(\gamma-\beta) z^{-\alpha} \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)\Gamma(\alpha+n+1-\gamma)}{\Gamma(\alpha+n+1-\beta)\Gamma(n+1)} \left(\frac{1}{z}\right)^n \\ &= \frac{C''}{\Gamma(\alpha)} \Gamma(\gamma-\beta) z^{-\alpha} \frac{\Gamma(\alpha)\Gamma(\alpha+1-\gamma)}{\Gamma(\alpha+1-\beta)} F(\alpha, \alpha-\gamma+1; \alpha-\beta+1; 1/z), \end{aligned}$$

where we have used the hypergeometric series of Chapter 14. Choosing

$$C'' = \frac{\Gamma(\alpha+1-\beta)}{\Gamma(\gamma-\beta)\Gamma(\alpha+1-\gamma)}$$

yields  $w(z) = z^{-\alpha} F(\alpha, \alpha-\gamma+1; \alpha-\beta+1; 1/z)$ , which is one of the solutions of the hypergeometric DE [Equation (14.30)].  $\blacksquare$

## 15.2 Integral Representation of the Confluent Hypergeometric Function

Having obtained the integral representation of the hypergeometric function, we can readily get the integral representation of the confluent hypergeometric function by taking the proper limit. It was shown in Chapter 14 that  $\Phi(\alpha, \gamma; z) = \lim_{\beta \rightarrow \infty} F(\alpha, \beta; \gamma; z/\beta)$ . This suggests taking the limit of Equation (15.7). The presence of the gamma functions with  $\beta$  as their arguments complicates things, but on the other hand, the symmetry of the hypergeometric function can be utilized

to our advantage. Thus, we may write

$$\begin{aligned}\Phi(\alpha, \gamma; z) &= \lim_{\beta \rightarrow \infty} F\left(\alpha, \beta; \gamma; \frac{z}{\beta}\right) = \lim_{\beta \rightarrow \infty} F\left(\beta, \alpha; \gamma; \frac{z}{\beta}\right) \\ &= \lim_{\beta \rightarrow \infty} \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \int_0^1 \left(1 - \frac{tz}{\beta}\right)^{-\beta} t^{\alpha-1} (1-t)^{\gamma-\alpha-1} dt \\ &= \Phi(\alpha, \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \int_0^1 e^{zt} t^{\alpha-1} (1-t)^{\gamma-\alpha-1} dt\end{aligned}\quad (15.11)$$

because the limit of the first term in the integral is simply  $e^{tz}$ . Note that the condition  $\text{Re}(\gamma) > \text{Re}(\alpha) > 0$  must still hold here.

Integral transforms are particularly useful in determining the asymptotic behavior of functions. We shall use them in deriving asymptotic formulas for Bessel functions later on, and Problem 15.10 derives the asymptotic formula for the confluent hypergeometric function.

### 15.3 Integral Representation of Bessel Functions

Choosing the kernel, the contour, and the function  $v(t)$  that lead to an integral representation of a function is an art, and the nineteenth century produced many masters of it. A particularly popular theme in such endeavors was the Bessel equation and Bessel functions. This section considers the integral representations of Bessel functions.

The most effective kernel for the Bessel DE is

$$K(z, t) = \left(\frac{z}{2}\right)^{\nu} \exp\left(t - \frac{z^2}{4t}\right).$$

When the Bessel DO  $\mathbf{L}_z \equiv \frac{d^2}{dz^2} + \frac{1}{z} \frac{d}{dz} + \left(1 - \frac{\nu^2}{z^2}\right)$  acts on  $K(z, t)$ , it yields

$$\mathbf{L}_z K(z, t) = \left(-\frac{\nu+1}{t} + 1 + \frac{z^2}{4t^2}\right) \left(\frac{z}{2}\right)^{\nu} e^{t-z^2/4t} = \left(\frac{d}{dt} - \frac{\nu+1}{t}\right) K(z, t).$$

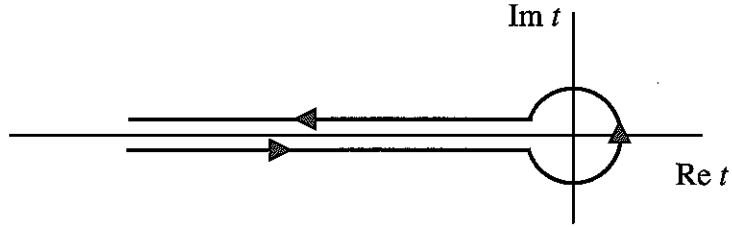
Thus,  $\mathbf{M}_t = d/dt - (\nu+1)/t$ , and Equation (13.20) gives

$$\mathbf{M}_t^{\dagger}[v(t)] = -\frac{dv}{dt} - \frac{\nu+1}{t}v = 0,$$

whose solution, including the arbitrary constant of integration  $k$ , is  $v(t) = kt^{-\nu-1}$ . When we substitute this solution and the kernel in the surface term of the Lagrange identity, Equation (13.24), we obtain

$$\mathcal{Q}[K, v](t) = p_1 K(z, t)v(t) = k \left(\frac{z}{2}\right)^{\nu} t^{-\nu-1} e^{t-z^2/(4t)}.$$





**Figure 15.1** The contour  $C$  in the  $t$ -plane used in evaluating  $J_\nu(z)$ .

A contour in the  $t$ -plane that ensures the vanishing of  $Q[K, \nu]$  for all values of  $\nu$  starts at  $t = -\infty$ , comes to the origin, orbits it on an arbitrary circle, and finally goes back to  $t = -\infty$  (see Figure 15.1). Such a contour is possible because of the factor  $e^t$  in the expression for  $Q[K, \nu]$ . We thus can write

$$J_\nu(z) = k \left(\frac{z}{2}\right)^\nu \int_C t^{-\nu-1} e^{t-z^2/(4t)} dt \quad (15.12)$$

Note that the integrand has a cut along the negative real axis due to the factor  $t^{-\nu-1}$ . If  $\nu$  is an integer, the cut shrinks to a pole at  $t = 0$ .

The constant  $k$  must be determined in such a way that the above expression for  $J_\nu(z)$  agrees with the series representation obtained in Chapter 14. It can be shown (see Problem 15.11) that  $k = 1/(2\pi i)$ . Thus, we have

integral  
representation of  
Bessel function

$$J_\nu(z) = \frac{1}{2\pi i} \left(\frac{z}{2}\right)^\nu \int_C t^{-\nu-1} e^{t-z^2/(4t)} dt$$

It is more convenient to take the factor  $(z/2)^\nu$  into the integral, introduce a new integration variable  $u = 2t/z$ , and rewrite the preceding equation as

$$J_\nu(z) = \frac{1}{2\pi i} \int_C u^{-\nu-1} e^{(z/2)(u-1/u)} du. \quad (15.13)$$

This result is valid as long as  $\text{Re}(zu) < 0$  when  $u \rightarrow -\infty$  on the negative real axis; that is,  $\text{Re}(z)$  must be positive for Equation (15.13) to work.

An interesting result can be obtained from Equation (15.13) when  $\nu$  is an integer. In that case the only singularity will be at the origin, so the contour can be taken to be a circle about the origin. This yields

$$J_n(z) = \frac{1}{2\pi i} \int_C u^{-n-1} e^{(z/2)(u-1/u)} du,$$

which is the  $n$ th coefficient of the Laurent series expansion of  $\exp[(z/2)(u-1/u)]$  about the origin. We thus have this important result:

Bessel generating  
function

$$e^{(z/2)(t-1/t)} = \sum_{n=-\infty}^{\infty} J_n(z) t^n. \quad (15.14)$$

The function  $\exp[(z/2)(t - 1/t)]$  is therefore appropriately called the **generating function** for Bessel functions of integer order (see also Problem 14.40). Equation (15.14) can be useful in deriving relations for such Bessel functions as the following example shows.

**15.3.1. Example.** Let us rewrite the LHS of (15.14) as  $e^{zt/2}e^{-z/2t}$ , expand the exponentials, and collect terms to obtain

$$\begin{aligned} e^{(z/2)(t-1/t)} &= e^{zt/2}e^{-z/2t} = \sum_{m=0}^{\infty} \frac{1}{m!} \left(\frac{zt}{2}\right)^m \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{z}{2t}\right)^n \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^n}{m!n!} \left(\frac{z}{2}\right)^{m+n} t^{m-n}. \end{aligned}$$

If we let  $m - n = k$ , change the  $m$  summation to  $k$ , and note that  $k$  goes from  $-\infty$  to  $\infty$ , we get

$$\begin{aligned} e^{(z/2)(t-1/t)} &= \sum_{k=-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+k)!n!} \left(\frac{z}{2}\right)^{2n+k} t^k \\ &= \sum_{k=-\infty}^{\infty} \left[ \left(\frac{z}{2}\right)^k \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+k+1)\Gamma(n+1)} \left(\frac{z}{2}\right)^{2n} \right] t^k. \end{aligned}$$

Comparing this equation with Equation (15.14) yields the familiar expansion for the Bessel function:

$$J_k(z) = \left(\frac{z}{2}\right)^k \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+k+1)\Gamma(n+1)} \left(\frac{z}{2}\right)^{2n}.$$

We can also obtain a recurrence relation for  $J_n(z)$ . Differentiating both sides of Equation (15.14) with respect to  $t$  yields

$$\frac{z}{2} \left(1 + \frac{1}{t^2}\right) e^{(z/2)(t-1/t)} = \sum_{n=-\infty}^{\infty} n J_n(z) t^{n-1}. \quad (15.15)$$

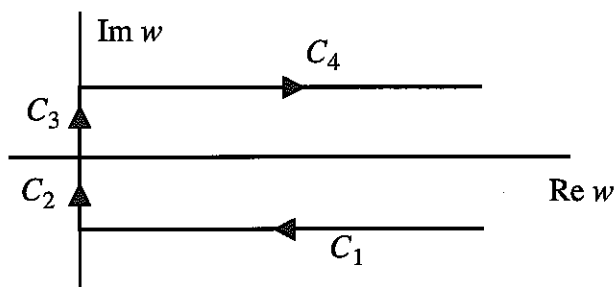
Using Equation (15.14) on the LHS gives

$$\begin{aligned} \sum_{n=-\infty}^{\infty} \left(\frac{z}{2} + \frac{z}{2t^2}\right) J_n(z) t^n &= \frac{z}{2} \sum_{n=-\infty}^{\infty} J_n(z) t^n + \frac{z}{2} \sum_{n=-\infty}^{\infty} J_n(z) t^{n-2} \\ &= \frac{z}{2} \sum_{n=-\infty}^{\infty} J_{n-1}(z) t^{n-1} + \frac{z}{2} \sum_{n=-\infty}^{\infty} J_{n+1}(z) t^{n-1}, \end{aligned} \quad (15.16)$$

where we substituted  $n - 1$  for  $n$  in the first sum and  $n + 1$  for  $n$  in the second. Equating the coefficients of equal powers of  $t$  on the LHS and the RHS of Equations (15.15) and (15.16), we get

$$n J_n(z) = \frac{z}{2} [J_{n-1}(z) + J_{n+1}(z)],$$

which was obtained by a different method in Chapter 14 [see Eq. (14.48)]. ■



**Figure 15.2** The contour  $C'$  in the  $w$ -plane used in evaluating  $J_\nu(z)$ .

We can start with Equation (15.13) and obtain other integral representations of Bessel functions by making appropriate substitutions. For instance, we can let  $u = e^w$  and assume that the circle of the contour  $C$  has unit radius. The contour  $C'$  in the  $w$ -plane is determined as follows. Write  $u = re^{i\theta}$  and  $w \equiv x + iy$ , so<sup>1</sup>  $re^{i\theta} = e^x e^{iy}$  yielding  $r = e^x$  and  $e^{i\theta} = e^{iy}$ . Along the first part of  $C$ ,  $\theta = -\pi$  and  $r$  goes from  $\infty$  to 1. Thus, along the corresponding part of  $C'$ ,  $y = -\pi$  and  $x$  goes from  $\infty$  to 0. On the circular part of  $C$ ,  $r = 1$  and  $\theta$  goes from  $-\pi$  to  $+\pi$ . Thus, along the corresponding part of  $C'$ ,  $x = 0$  and  $y$  goes from  $-\pi$  to  $+\pi$ . Finally, on the last part of  $C'$ ,  $y = \pi$  and  $x$  goes from 0 to  $\infty$ . Therefore, the contour  $C'$  in the  $w$ -plane is as shown in Figure 15.2.

Substituting  $u = e^w$  in Equation (15.13) yields

$$J_\nu(z) = \frac{1}{2\pi i} \int_{C'} e^{z \sinh w - \nu w} dw, \quad \operatorname{Re}(z) > 0, \quad (15.17)$$

which can be transformed into (see Problem 15.12)

$$J_\nu(z) = \frac{1}{\pi} \int_0^\pi \cos(\nu\theta - z \sin \theta) d\theta - \frac{\sin \nu\pi}{\pi} \int_0^\infty e^{-\nu t - z \sinh t} dt. \quad (15.18)$$

For the special case of integer  $\nu$ , we obtain

$$J_n(z) = \frac{1}{\pi} \int_0^\pi \cos(n\theta - z \sin \theta) d\theta.$$

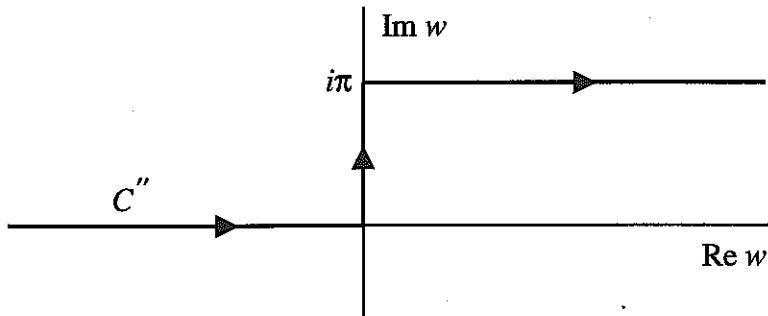
In particular,

$$J_0(z) = \frac{1}{\pi} \int_0^\pi \cos(z \sin \theta) d\theta.$$

We can use the integral representation for  $J_\nu(z)$  to find the integral representation for Bessel functions of other kinds. For instance, to obtain the integral

Integral  
representation of  
Bessel functions of  
integer order

<sup>1</sup>Do not confuse  $x$  and  $y$  with the real and imaginary parts of  $z$ .



**Figure 15.3** The contour  $C''$  in the  $w$ -plane used in evaluating  $H_v^{(1)}(z)$ .

representation for the Neumann function  $Y_\nu(z)$ , we use Equation (14.44):

$$\begin{aligned} Y_\nu(z) &= (\cot \nu\pi)J_\nu(z) - \frac{1}{\sin \nu\pi}J_{-\nu}(z) \\ &= \frac{\cot \nu\pi}{\pi} \int_0^\pi \cos(\nu\theta - z \sin \theta) d\theta - \frac{\cos \nu\pi}{\pi} \int_0^\infty e^{-\nu t - z \sinh t} dt \\ &\quad - \frac{1}{\pi \sin \nu\pi} \int_0^\pi \cos(\nu\theta + z \sin \theta) d\theta - \frac{1}{\pi} \int_0^\infty e^{\nu t - z \sinh t} dt \end{aligned}$$

with  $\operatorname{Re}(z) > 0$ . Substitute  $\pi - \theta$  for  $\theta$  in the third integral on the RHS. Then insert the resulting integrals plus Equation (15.18) in  $H_\nu^{(1)}(z) = J_\nu(z) + iY_\nu(z)$  to obtain

$$\begin{aligned} H_\nu^{(1)}(z) &= \frac{1}{\pi} \int_0^\pi e^{i(z \sin \theta - \nu\theta)} d\theta + \frac{1}{i\pi} \int_0^\infty e^{\nu t - z \sinh t} dt \\ &\quad + \frac{e^{-i\nu\pi}}{i\pi} \int_0^\infty e^{-\nu t - z \sinh t} dt, \quad \operatorname{Re}(z) > 0. \end{aligned}$$

These integrals can easily be shown to result from integrating along the contour  $C''$  of Figure 15.3. Thus, we have

$$H_\nu^{(1)}(z) = \frac{1}{i\pi} \int_{C''} e^{z \sinh w - \nu w} dw \quad \operatorname{Re}(z) > 0.$$

By changing  $i$  to  $-i$ , we can show that

$$H_\nu^{(2)}(z) = -\frac{1}{i\pi} \int_{C'''} e^{z \sinh w - \nu w} dw, \quad \operatorname{Re}(z) > 0,$$

where  $C'''$  is the mirror image of  $C''$  about the real axis.

## 15.4 Asymptotic Behavior of Bessel Functions

As mentioned before, integral representations are particularly useful for determining the asymptotic behavior of functions. For Bessel functions we can consider two kinds of limits. Assuming that both  $\nu$  and  $z = x$  are real, we can consider  $\nu \rightarrow \infty$  or  $x \rightarrow \infty$ . First, let us consider the behavior of  $J_\nu(x)$  of large order. The appropriate method for calculating the asymptotic form is the method of steepest descent discussed in Chapter 11 for which  $\nu$  takes the place of the large parameter  $\alpha$ . We use Equation (15.17) because its integrand is simpler than that of Equation (15.13). The form of the integrand in Equation (15.17) may want to suggest  $f(w) = -w$  and  $g(w) = e^{x \sinh w}$ . However, this choice does not allow setting  $f'(w)$  equal to zero. To proceed, therefore, we write the exponent as  $\nu \left( \frac{x}{\nu} \sinh w - w \right)$ , and conveniently introduce  $x/\nu \equiv 1/\cosh w_0$ , with  $w_0$  a real number, which we take to be positive. Substituting this in the equation above, we can read off

$$f(w) = \frac{\sinh w}{\cosh w_0} - w, \quad g(w) = 1.$$

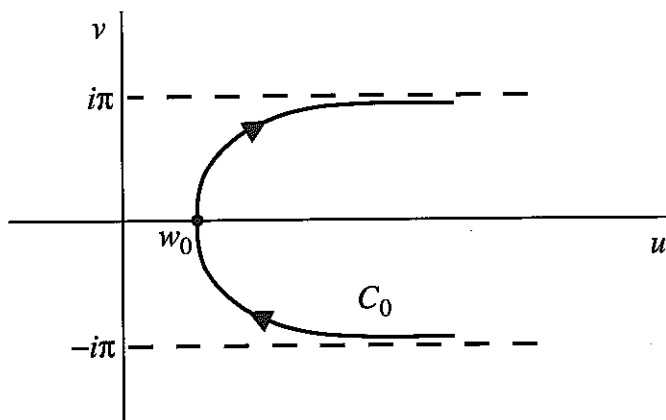
The saddle point is obtained from  $df/dw = 0$  or  $\cosh w = \cosh w_0$ . Thus,  $w = \pm w_0 + 2in\pi$ , for  $n = 0, 1, 2, \dots$ . Since the contour  $C'$  lies in the right half-plane, we choose  $w_0$  as the saddle point. The second derivative  $f''(w_0)$  is simply  $\tanh w_0$ , which is real, making  $\theta_2 = 0$ , and  $\theta_1 = \pi/2$  or  $3\pi/2$ . The convention of Chapter 11 suggests taking  $\theta_1 = \pi/2$  (see Figure 15.4). The rest is a matter of substitution. We are interested in the approximation to  $w$  up to the third order in  $t$ :  $w - w_0 = b_1 t + b_2 t^2 + b_3 t^3$ . Using Equations (11.31), (11.37), and (11.38), we can easily find the three coefficients:

$$\begin{aligned} b_1 &= \frac{\sqrt{2}}{|f''(w_0)|^{1/2}} e^{i\theta_1} = i \frac{\sqrt{2}}{\sqrt{\tanh w_0}}, \\ b_2 &= \frac{f'''(w_0)}{3|f''(w_0)|^2} e^{4i\theta_1} = \frac{\cosh^2 w_0}{3 \sinh^2 w_0}, \\ b_3 &= \left\{ \frac{5[f'''(w_0)]^2}{3[f''(w_0)]^2} - \frac{f^{(4)}(w_0)}{f''(w_0)} \right\} \frac{\sqrt{2} e^{3i\theta_1}}{12|f''(w_0)|^{3/2}} \\ &= -i \frac{\sqrt{2}}{12(\tanh w_0)^{3/2}} \left( \frac{5}{3} \coth^2 w_0 - 1 \right). \end{aligned}$$

If we substitute the above in Equation (11.36), we obtain the following asymptotic formula valid for  $\nu \rightarrow \infty$ :

$$J_\nu(x) \approx \frac{e^{x(\sinh w_0 - w_0 \cosh w_0)}}{(2\pi x \sinh w_0)^{1/2}} \left[ 1 + \frac{1}{8x \sinh w_0} \left( 1 - \frac{5}{3} \coth^2 w_0 \right) + \dots \right],$$

where  $\nu$  is related to  $w_0$  via  $\nu = x \cosh w_0$ .



**Figure 15.4** The contour  $C_0$  in the  $w$ -plane used in evaluating  $J_\nu(z)$  for large values of  $\nu$ .

Let us now consider the asymptotic behavior for large  $x$ . It is convenient to consider the Hankel functions  $H_\nu^{(1)}(x)$  and  $H_\nu^{(2)}(x)$ . The contours  $C''$  and  $C'''$  involve both the positive and the negative real axis; therefore, it is convenient, assuming that  $x > \nu$ , to write  $\nu = x \cos \beta$  so that

$$H_\nu^{(1)}(x) = \frac{1}{i\pi} \int_{C''} e^{x(\sinh w - w \cos \beta)} dw.$$

The saddle points are given by the solutions to  $\cosh w = \cos \beta$ , which are  $w_0 = \pm i\beta$ . Choosing  $w_0 = +i\beta$ , we note that the contour along which

$$\text{Im}(\sinh w - w \cos \beta) = \text{Im}(\sinh w_0 - w_0 \cos \beta)$$

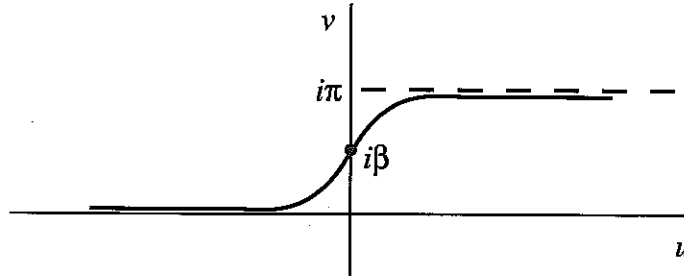
is given by  $\cosh u = [\sin \beta + (v - \beta) \cos \beta] / \sin v$ . This contour is shown in Figure 15.5. The rest of the procedure is exactly the same as for  $J_\nu(x)$  described above. In fact, to obtain the expansion for  $H_\nu^{(1)}(x)$ , we simply replace  $w_0$  by  $i\beta$ . The result is

$$H_\nu^{(1)}(x) \approx \left( \frac{2}{i\pi x \sin \beta} \right)^{1/2} e^{i(x \sin \beta - \nu\beta)} \left[ 1 + \frac{1}{8ix \sin \beta} \left( 1 + \frac{5}{3} \cot^2 \beta \right) + \dots \right].$$

When  $x$  is much larger than  $\nu$ ,  $\beta$  will be close to  $\pi/2$ , and we have

$$H_\nu^{(1)}(x) \approx \sqrt{\frac{2}{\pi x}} e^{i(x - \nu\pi/2 - \pi/4)} \left( 1 + \frac{1}{8ix} \right),$$

which, with  $1/x \rightarrow 0$ , is what we obtained in Example 11.5.2.



**Figure 15.5** The contour in the  $w$ -plane used in evaluating  $H_\nu^{(1)}(z)$  in the limit of large values of  $x$ .

The other saddle point, at  $-i\beta$ , gives the other Hankel function, with the asymptotic limit

$$H_\nu^{(2)}(x) \approx \sqrt{\frac{2}{\pi x}} e^{-i(x-\nu\pi/2-\pi/4)} \left(1 - \frac{1}{8ix}\right)$$

We can now use the expressions for the asymptotic forms of the two Hankel functions to write the asymptotic forms of  $J_\nu(x)$  and  $Y_\nu(x)$  for large  $x$ :

$$\begin{aligned} J_\nu(x) &= \frac{1}{2} [H_\nu^{(1)}(x) + H_\nu^{(2)}(x)] \\ &\approx \sqrt{\frac{2}{\pi x}} \left[ \cos\left(x - \nu\frac{\pi}{2} - \frac{\pi}{4}\right) + \frac{1}{8x} \sin\left(x - \nu\frac{\pi}{2} - \frac{\pi}{4}\right) + \dots \right], \\ Y_\nu(x) &= \frac{1}{2i} [H_\nu^{(1)}(x) - H_\nu^{(2)}(x)] \\ &\approx \sqrt{\frac{2}{\pi x}} \left[ \sin\left(x - \nu\frac{\pi}{2} - \frac{\pi}{4}\right) - \frac{1}{8x} \cos\left(x - \nu\frac{\pi}{2} - \frac{\pi}{4}\right) + \dots \right]. \end{aligned}$$

## 15.5 Problems

**15.1.** Use the change of variables  $k = \ln t$  and  $ix = \omega - \alpha$  (where  $k$  and  $x$  are the common variables used in Fourier transform equations) to show that the Fourier transform changes into a Mellin transform,

$$G(t) = \frac{1}{2\pi i} \int_{-i\infty+\alpha}^{i\infty+\alpha} F(\omega) t^{-\omega} d\omega, \quad \text{where } F(\omega) = \int_0^\infty G(t) t^{\omega-1} dt.$$

**15.2.** The Laplace transform  $L[f]$  of a function  $f(t)$  is defined as

$$L[f](s) \equiv \int_0^\infty e^{-st} f(t) dt.$$

Show that the Laplace transform of

- (a)  $f(t) = 1$  is  $\frac{1}{s}$ , where  $s > 0$ .
- (b)  $f(t) = \cosh \omega t$  is  $\frac{s}{s^2 - \omega^2}$ , where  $s^2 > \omega^2$ .
- (c)  $f(t) = \sinh \omega t$  is  $\frac{\omega}{s^2 - \omega^2}$ , where  $s^2 > \omega^2$ .
- (d)  $f(t) = \cos \omega t$  is  $\frac{s}{s^2 + \omega^2}$ .
- (e)  $f(t) = \sin \omega t$  is  $\frac{\omega}{s^2 + \omega^2}$ .
- (f)  $f(t) = e^{\omega t}$  for  $t > 0$ , is  $\frac{1}{s - \omega}$ , where  $s > \omega$ .
- (g)  $f(t) = t^n$  is  $\frac{\Gamma(n+1)}{s^{n+1}}$ , where  $s > 0$ ,  $n > -1$ .

15.3. Evaluate the integral

$$f(t) = \int_0^{\infty} \frac{\sin \omega t}{\omega} d\omega$$

by finding the Laplace transform and changing the order of integration. Express the result for both  $t > 0$  and  $t < 0$  in terms of the theta function. (You will need some results from Problem 15.2.)

15.4. Show that the Laplace transform of the derivative of a function is given by  $L[F'](s) = sL[F](s) - F(0)$ . Similarly, show that for the second derivative the transform is

$$L[F''](s) = s^2L[F](s) - sF(0) - F'(0).$$

Use these results to solve the differential equation  $u''(t) + \omega^2 u(t) = 0$  subject to the boundary conditions  $u(0) = a$ ,  $u'(0) = 0$ .

15.5. Solve the DE of Equation (15.5).

15.6. Calculate the surface term for the hypergeometric DE.

15.7. Determine the constant  $C'$  in Equation (15.6), the solution to the hypergeometric DE. Hint: Expand  $(t-z)^{-\alpha}$  inside the integral, use Equations (11.19) and (11.17), and compare the ensuing series with the hypergeometric series of Chapter 14.

15.8. Derive the Euler formula [Equation (15.7)].

15.9. Show that

$$F(\alpha, \beta; \gamma; 1) = \frac{\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)}{\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)}. \quad (15.19)$$



Hint: Use Equation (11.19). Equation (15.19) was obtained by Gauss using only hypergeometric series.

**15.10.** We determine the asymptotic behavior of  $\Phi(\alpha, \gamma; z)$  for  $z \rightarrow \infty$  in this problem. Break up the integral in Equation (15.11) into two parts, one from 0 to  $-\infty$  and the other from  $-\infty$  to 1. Substitute  $-t/z$  for  $t$  in the first integral, and  $1 - t/z$  for  $t$  in the second. Assuming that  $z \rightarrow \infty$  along the positive real axis, show that the second integral will dominate, and that

$$\Phi(\alpha, \gamma; z) \rightarrow \frac{\Gamma(\gamma)}{\Gamma(\alpha)} z^{\alpha-\gamma} e^z \quad \text{as } z \rightarrow \infty.$$

**15.11.** In this problem, we determine the constant  $k$  of Equation (15.12).

(a) Write the contour integral of Equation (15.12) for each of the three pieces of the contour. Note that  $\arg(t) = -\pi$  as  $t$  comes from  $-\infty$  and  $\arg(t) = \pi$  as  $t$  goes to  $-\infty$ . Obtain a real integral from 0 to  $\infty$ .

(b) Use the relation  $\Gamma(z)\Gamma(1-z) = \pi/\sin \pi z$ , obtained in Chapter 11, to show that

$$\Gamma(-z) = -\frac{\pi}{\Gamma(z+1)\sin \pi z}.$$

(c) Expand the function  $\exp(z^2/4t)$  in the integral of part (a), and show that the contour integral reduces to

$$-2i \sin \nu\pi \sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^{2n} \frac{\Gamma(-n-\nu)}{\Gamma(n+1)}.$$

(d) Use the result of part (c) in part (b), and compare the result with the series expansion of  $J_\nu(z)$  in Chapter 14 to arrive finally at  $k = 1/(2\pi i)$ .

**15.12.** By integrating along  $C_1, C_2, C_3,$  and  $C_4$  of Figure 15.2, derive Equation (15.18).

**15.13.** By substituting  $t = \exp(i\theta)$  in Equation (15.14), show that

$$e^{iz \sin \theta} = J_0(z) + 2 \sum_{n=1}^{\infty} J_{2n}(z) \cos(2n\theta) + 2i \sum_{n=0}^{\infty} J_{2n+1}(z) \sin[(2n+1)\theta].$$

In particular, show that

$$J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{iz \sin \theta} d\theta.$$

**15.14.** Derive the integral representations of  $H_\nu^{(1)}(x)$  and  $H_\nu^{(2)}(x)$  given in Section 15.3.

## Additional Reading

1. Dennery, P. and Krzywicki, A. *Mathematics for Physicists*, Harper and Row, 1967.
2. Mathews, J. and Walker, R. *Mathematical Methods of Physics*, 2nd ed., Benjamin, 1970.

Part V

---

# Operators on Hilbert Spaces



# An Introduction to Operator Theory

The first two parts of the book dealt almost exclusively with algebraic techniques. The third and fourth part were devoted to analytic methods. In this introductory chapter, we shall try to unite these two branches of mathematics to gain insight into the nature of some of the important equations in physics and their solutions. Let us start with a familiar problem.

## 16.1 From Abstract to Integral and Differential Operators

Let's say we want to solve an abstract vector-operator equation  $\mathbf{A}|u\rangle = |v\rangle$  in an  $N$ -dimensional vector space  $\mathcal{V}$ . To this end, we select a basis  $B = \{|a_i\rangle\}_{i=1}^N$ , write the equation in matrix form, and solve the resulting system of  $N$  linear equations. This produces the components of the solution  $|u\rangle$  in  $B$ . If components in another basis  $B'$  are desired, they can be obtained using the similarity transformation connecting the two bases (see Chapter 3).

There is a standard formal procedure for obtaining the matrix equation. It is convenient to choose an orthonormal basis  $B = \{|e_i\rangle\}_{i=1}^N$  for  $\mathcal{V}$  and refer all components to this basis. The procedure involves contracting both sides of the equation with  $\langle e_i|$  and inserting  $\mathbf{1} = \sum_{j=1}^N |e_j\rangle\langle e_j|$  between  $\mathbf{A}$  and  $|u\rangle$ :  $\sum_{j=1}^N \langle e_i|\mathbf{A}|e_j\rangle\langle e_j|u\rangle = \langle e_i|v\rangle$  for  $i = 1, 2, \dots, N$ , or

$$\sum_{j=1}^N A_{ij}u_j = v_i \quad \text{for } i = 1, 2, \dots, N, \quad (16.1)$$

where  $A_{ij} \equiv \langle e_i | \mathbf{A} | e_j \rangle$ ,  $u_j \equiv \langle e_j | u \rangle$ , and  $v_i \equiv \langle e_i | v \rangle$ . Equation (16.1) is a system of  $N$  linear equations in  $N$  unknowns  $\{u_j\}_{j=1}^N$ , which can be solved to obtain the solution(s) of the original equation in  $B$ .

A convenient basis is that in which  $\mathbf{A}$  is represented by a diagonal matrix  $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ . Then the operator equation takes the simple form  $\lambda_i u_i = v_i$ , and the solution becomes immediate.

Let us now apply the procedure just described to infinite-dimensional vector spaces, in particular, for the case of a continuous index. We want to find the solutions of  $\mathbf{K}|u\rangle = |f\rangle$ . Following the procedure used above, we obtain

$$\langle x | \underbrace{\mathbf{K} \left( \int_a^b |y\rangle w(y) \langle y| dy \right)}_{\equiv 1} |u\rangle = \int_a^b \langle x | \mathbf{K} |y\rangle w(y) \langle y | u \rangle dy = \langle x | f \rangle,$$

where we have used the results obtained in Chapter 6. Writing this in functional notation, we have

$$\int_a^b K(x, y)w(y)u(y) dy = f(x), \tag{16.2}$$

Integral operators  
and kernels

which is the continuous analogue of Equation (16.1). Here  $(a, b)$  is the interval on which the functions are defined. We note that the indices have turned into continuous arguments, and the sum has turned into an integral. The operator  $\mathbf{K}$  that leads to an equation such as (16.2) is called an **integral operator (IO)**, and the “matrix element”  $K(x, y)$  is said to be its **kernel**.

local operators

The discussion of the discrete case mentioned the possibility of the operator  $\mathbf{A}$  being diagonal in the given basis  $B$ . Let us do the same with (16.2); that is, noting that  $x$  and  $y$  are indices for  $K$ , let us assume that  $K(x, y) = 0$  for  $x \neq y$ . Such operators are called **local operators**. For local operators, the contribution to the integral comes only at the point where  $x = y$  (hence, their name). If  $K(x, y)$  is finite at this point, and the functions  $w(y)$  and  $u(y)$  are well behaved there, the LHS of (16.2) will vanish, and we will get inconsistencies. To avoid this, we need to have

$$K(x, y) = \begin{cases} 0 & \text{if } x \neq y, \\ \infty & \text{if } x = y. \end{cases}$$

Thus,  $K(x, y)$  has the behavior of a delta function. Letting  $K(x, y) \equiv L(x)\delta(x - y)/w(x)$  and substituting in Equation (16.2) yields  $L(x)u(x) = f(x)$ .

In the discrete case,  $\lambda_i$  was merely an indexed number; its continuous analogue,  $L(x)$ , may represent merely a function. However, the fact that  $x$  is a continuous variable (index) gives rise to other possibilities for  $L(x)$  that do not exist for the discrete case. For instance,  $L(x)$  could be a *differential operator*. The derivative, although defined by a limiting process involving neighboring points, is a local operator. Thus, we can speak of the derivative of a function *at a point*. For the

discrete case,  $u_i$  can only “hop” from  $i$  to  $i + 1$  and then back to  $i$ . Such a difference (as opposed to differential) process is not local; it involves not only  $i$  but also  $i + 1$ . The “point”  $i$  does not have an (infinitesimally close) neighbor.

This essential difference between discrete and continuous operators makes the latter far richer in possibilities for applications. In particular, if  $L(x)$  is considered a differential operator, the equation  $L(x)u(x) = f(x)$  leads directly to the fruitful area of differential equation theory.

## 16.2 Bounded Operators in Hilbert Spaces

The concept of an operator on a Hilbert space is extremely subtle. Even the elementary characteristics of operators, such as the operation of hermitian conjugation, cannot generally be defined on the whole Hilbert space.

In finite-dimensional vector spaces there is a one-to-one correspondence between operators and matrices. So, in some sense, the study of operators reduces to a study of matrices, which are collections of real or complex numbers. Although we have already noted an analogy between matrices and kernels, a whole new realm of questions arises when  $A_{ij}$  is replaced by  $K(x, y)$ —questions about the continuity of  $K(x, y)$  in both its arguments, about the limit of  $K(x, y)$  as  $x$  and/or  $y$  approach the “end points” of the interval on which  $K$  is defined, about the boundedness and “compactness” of  $K$ , and so on. Such subtleties are not unexpected. After all, when we tried to generalize concepts of finite-dimensional vector spaces to infinite dimensions in Chapter 5, we encountered difficulties. There we were concerned about vectors only; the generalization of operators is even more complicated.

**16.2.1. Example.** Recall that  $\mathbb{C}^\infty$  is the set of sequences  $|a\rangle = \{\alpha_i\}_{i=1}^\infty$ , or of  $\infty$ -tuples  $(\alpha_1, \alpha_2, \dots)$ , that satisfy the convergence requirement  $\sum_{j=1}^\infty |\alpha_j|^2 < \infty$  (see Example 1.1.2). It is a Hilbert space with inner product defined by  $\langle a|b\rangle = \sum_{j=1}^\infty \alpha_j^* \beta_j$ . The standard (orthonormal) basis for  $\mathbb{C}^\infty$  is  $\{|e_i\rangle\}_{i=1}^\infty$ , where  $|e_i\rangle$  has all components equal to zero except the  $i$ th one, which is 1. Then one has  $|a\rangle = \sum_{j=1}^\infty \alpha_j |e_j\rangle$ .

right-shift operator

One can introduce an operator  $\mathbf{T}_r$ , called the **right-shift operator**, by

$$\mathbf{T}_r |a\rangle = \mathbf{T}_r \left( \sum_{j=1}^\infty \alpha_j |e_j\rangle \right) = \sum_{j=1}^\infty \alpha_j |e_{j+1}\rangle.$$

In other words,  $\mathbf{T}_r$  transforms  $(\alpha_1, \alpha_2, \dots)$  to  $(0, \alpha_1, \alpha_2, \dots)$ . It is straightforward to show that  $\mathbf{T}_r$  is indeed a linear operator. ■

The first step in our study of vector spaces of infinite dimensions was getting a handle on the convergence of infinite sums. This entailed defining a norm for vectors and a distance between them. In addition, we noted that the set of linear transformations  $\mathcal{L}(\mathcal{V}, \mathcal{W})$  was a vector space in its own right. Since operators are “vectors” in this space, the study of operators requires constructing a norm in  $\mathcal{L}(\mathcal{V}, \mathcal{W})$  when  $\mathcal{V}$  and  $\mathcal{W}$  are infinite-dimensional.

**16.2.2. Definition.** Let  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be two Hilbert spaces with norms  $\|\cdot\|_1$  and  $\|\cdot\|_2$ . For any  $\mathbf{T} \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ , the number

$$\max \left\{ \frac{\|\mathbf{T}x\|_2}{\|x\|_1} \mid |x\rangle \neq 0 \right\}$$

operator norm (if it exists) is called<sup>1</sup> the **operator norm** of  $\mathbf{T}$  and is denoted by  $\|\mathbf{T}\|$ . A linear transformation whose norm is finite is called a **bounded linear transformation**. A bounded linear transformation from a Hilbert space to itself is called a **bounded operator**. The collection of all bounded linear transformations, which is a subset of  $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ , will be denoted by  $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ , and if  $\mathcal{H}_1 = \mathcal{H}_2 \equiv \mathcal{H}$ , it will be denoted by  $\mathcal{B}(\mathcal{H})$ .

Note that  $\|\cdot\|_1$  and  $\|\cdot\|_2$  are the norms induced by the inner product of  $\mathcal{H}_1$  and  $\mathcal{H}_2$ . Also note that by dividing by  $\|x\|_1$  we eliminate the possibility of dilating the norm of  $\|\mathbf{T}\|$  by choosing a “long” vector. By restricting the length of  $|x\rangle$ , one can eliminate the necessity for dividing by the length. In fact, the norm can equivalently be defined as

$$\|\mathbf{T}\| = \max \{ \|\mathbf{T}x\|_2 \mid \|x\|_1 = 1 \} = \max \{ \|\mathbf{T}x\|_2 \mid \|x\|_1 \leq 1 \}. \tag{16.3}$$

It is straightforward to show that the three definitions are equivalent and they indeed define a norm.

**16.2.3. Proposition.** An operator  $\mathbf{T}$  is bounded if and only if it maps vectors of finite norm to vectors of finite norm.

*Proof.* Clearly, if  $\mathbf{T}$  is bounded, then  $\|\mathbf{T}x\|$  has finite norm. Conversely, if  $\|\mathbf{T}x\|_2$  is finite for all  $|x\rangle$  (of unit length),  $\max\{\|\mathbf{T}x\|_2 \mid \|x\|_1 = 1\}$  is also finite, and  $\mathbf{T}$  is bounded. □

An immediate consequence of the definition is

$$\|\mathbf{T}x\|_2 \leq \|\mathbf{T}\| \|x\|_1 \quad \forall |x\rangle \in \mathcal{H}_1. \tag{16.4}$$

If we choose  $|x\rangle - |y\rangle$  instead of  $|x\rangle$ , it will follow from (16.4) that as  $|x\rangle$  approaches  $|y\rangle$ ,  $\mathbf{T}|x\rangle$  approaches  $\mathbf{T}|y\rangle$ . This is the property that characterizes continuous functions:

bounded operators are continuous **16.2.4. Proposition.** The bounded operator  $\mathbf{T} \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$  is a continuous function from  $\mathcal{H}_1$  to  $\mathcal{H}_2$ .

Another consequence of the definition is that

---

<sup>1</sup>The precise definition uses “supremum” instead of “maximum.” Rather than spending a lot of effort explaining the difference between the two concepts, we use the less precise, but more intuitively familiar, concept of “maximum.”



**16.2.5. Box.**  $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$  is a vector subspace of  $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ , and for  $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ , we have  $\mathbf{1} \in \mathcal{B}(\mathcal{H})$  and  $\|\mathbf{1}\| = 1$ .

**16.2.6. Example.** We have seen that in an inner product space, one can associate a linear operator (linear functional) to every vector. Thus, associated with the vector  $|x\rangle$  in a Hilbert space  $\mathcal{H}$  is the linear operator  $\mathbf{f}_x : \mathcal{H} \rightarrow \mathbb{C}$  defined by  $\mathbf{f}_x(|y\rangle) \equiv \langle x | y \rangle$ . We want to compare the operator norm of  $\mathbf{f}_x$  with the norm of  $|x\rangle$ . First note that by using the Schwarz inequality, we get

$$\|\mathbf{f}_x\| = \max \left\{ \frac{|\mathbf{f}_x(|y\rangle)|}{\|y\|} \mid |y\rangle \neq 0 \right\} = \max \left\{ \frac{|\langle x | y \rangle|}{\|y\|} \mid |y\rangle \neq 0 \right\} \leq \|x\|.$$

On the other hand, from  $\|x\|^2 = \mathbf{f}_x(|x\rangle)$ , we obtain

$$\|x\| = \frac{\mathbf{f}_x(|x\rangle)}{\|x\|} \leq \max \left\{ \frac{|\mathbf{f}_x(|y\rangle)|}{\|y\|} \mid |y\rangle \neq 0 \right\} = \|\mathbf{f}_x\|.$$

These two inequalities imply that  $\|\mathbf{f}_x\| = \|x\|$ . ■

derivative operator is unbounded

**16.2.7. Example.** The derivative operator  $\mathbf{D} = d/dx$  is not a bounded operator on the Hilbert space<sup>2</sup>  $\mathcal{L}^2(a, b)$  of square-integrable functions. With a function like  $f(x) = \sqrt{x - a}$ , one gets

$$\|f\|^2 = \int_a^b (x - a) dx = \frac{1}{2}(b - a)^2 \Rightarrow \|f\| = \frac{b - a}{\sqrt{2}},$$

while  $df/dx = 1/(2\sqrt{x - a})$  gives  $\|\mathbf{D}f\|^2 = \frac{1}{4} \int_a^b dx/(x - a) = \infty$ . We conclude that  $\|\mathbf{D}\| = \infty$ . ■

norm of a product is less than the product of norms.

**16.2.8. Example.** Since  $\mathcal{L}(\mathcal{H})$  is an algebra as well as a vector space, one may be interested in the relation between the product of operators and their norms. More specifically, one may want to know how  $\|\mathbf{ST}\|$  is related to  $\|\mathbf{S}\|$  and  $\|\mathbf{T}\|$ . In this example we show that

$$\|\mathbf{ST}\| \leq \|\mathbf{S}\| \|\mathbf{T}\|. \tag{16.5}$$

To do so, we use the definition of operator norm for the product  $\mathbf{ST}$ :

$$\begin{aligned} \|\mathbf{ST}\| &= \max \left\{ \frac{\|\mathbf{ST}x\|}{\|x\|} \mid |x\rangle \neq 0 \right\} \\ &= \max \left\{ \frac{\|\mathbf{ST}x\|}{\|\mathbf{T}x\|} \frac{\|\mathbf{T}x\|}{\|x\|} \mid |x\rangle \neq 0 \neq \mathbf{T}|x\rangle \right\} \\ &\leq \max \left\{ \frac{\|\mathbf{S}(\mathbf{T}|x\rangle)\|}{\|\mathbf{T}x\|} \mid \mathbf{T}|x\rangle \neq 0 \right\} \underbrace{\max \left\{ \frac{\|\mathbf{T}x\|}{\|x\|} \mid |x\rangle \neq 0 \right\}}_{=\|\mathbf{T}\|} \end{aligned}$$

<sup>2</sup>Here the two Hilbert spaces coincide, so that the derivative operator acts on a single Hilbert space.

Now note that the first term on the RHS does not scan all the vectors for maximality: It scans only the vectors in the image of  $\mathbf{T}$ . If we include all vectors, we may obtain a larger number. Therefore,

$$\max \left\{ \frac{\|\mathbf{S}(\mathbf{T}|x)\|}{\|\mathbf{T}x\|} \mid \mathbf{T}|x\rangle \neq 0 \right\} \leq \max \left\{ \frac{\|\mathbf{S}x\|}{\|x\|} \mid |x\rangle \neq 0 \right\} = \|\mathbf{S}\|,$$

and the desired inequality is established. A useful consequence of this result is  $\|\mathbf{T}^n\| \leq \|\mathbf{T}\|^n$ , which we shall use frequently.  $\blacksquare$

We can put Equation (16.5) to immediate good use.

**16.2.9. Proposition.** *Let  $\mathcal{H}$  be a Hilbert space and  $\mathbf{T} \in \mathcal{B}(\mathcal{H})$ . If  $\|\mathbf{T}\| < 1$ , then  $\mathbf{1} - \mathbf{T}$  is invertible and  $(\mathbf{1} - \mathbf{T})^{-1} = \sum_{n=0}^{\infty} \mathbf{T}^n$ .*

*Proof.* First note that the series converges, because

$$\left\| \sum_{n=0}^{\infty} \mathbf{T}^n \right\| \leq \sum_{n=0}^{\infty} \|\mathbf{T}^n\| \leq \sum_{n=0}^{\infty} \|\mathbf{T}\|^n = \frac{1}{1 - \|\mathbf{T}\|}$$

and the sum has a finite norm. Furthermore,

$$\begin{aligned} (\mathbf{1} - \mathbf{T}) \sum_{n=0}^{\infty} \mathbf{T}^n &= (\mathbf{1} - \mathbf{T}) \left( \lim_{k \rightarrow \infty} \sum_{n=0}^k \mathbf{T}^n \right) = \lim_{k \rightarrow \infty} (\mathbf{1} - \mathbf{T}) \sum_{n=0}^k \mathbf{T}^n \\ &= \lim_{k \rightarrow \infty} \left( \sum_{n=0}^k \mathbf{T}^n - \sum_{n=0}^k \mathbf{T}^{n+1} \right) = \lim_{k \rightarrow \infty} (\mathbf{1} - \mathbf{T}^{k+1}) = \mathbf{1}, \end{aligned}$$

because  $0 \leq \lim_{k \rightarrow \infty} \|\mathbf{T}^{k+1}\| \leq \lim_{k \rightarrow \infty} \|\mathbf{T}\|^{k+1} = 0$  for  $\|\mathbf{T}\| < 1$ , and the vanishing of the norm implies the vanishing of the operator itself. One can similarly show that  $(\sum_{n=0}^{\infty} \mathbf{T}^n)(\mathbf{1} - \mathbf{T}) = \mathbf{1}$ .  $\square$

A corollary of this proposition is that operators that are “close enough” to an invertible operator are invertible (see Problem 16.1). Another corollary, whose proof is left as a straightforward exercise, is the following:

**16.2.10. Corollary.** *Let  $\mathbf{T} \in \mathcal{B}(\mathcal{H})$  and  $\lambda$  a complex number such that  $\|\mathbf{T}\| < |\lambda|$ . Then  $\mathbf{T} - \lambda\mathbf{1}$  is an invertible operator, and*

$$(\mathbf{T} - \lambda\mathbf{1})^{-1} = -\frac{1}{\lambda} \sum_{n=0}^{\infty} \left( \frac{\mathbf{T}}{\lambda} \right)^n.$$

Adjoints play an important role in the study of operators. We recall that the adjoint of  $\mathbf{T}$  is defined as  $\langle y | \mathbf{T}|x\rangle^* = \langle x | \mathbf{T}^\dagger |y\rangle$  or  $\langle \mathbf{T}x | y\rangle = \langle x | \mathbf{T}^\dagger y\rangle$ . In the finite-dimensional case, we could calculate the matrix representation of the adjoint in a particular basis using this definition and generalize to all bases by similarity transformations. That is why we never raised the question of the existence of the adjoint of an operator. In the infinite-dimensional case, one must prove such an existence. We state the following theorem without proof:

**16.2.11. Theorem.** Let  $T \in \mathcal{B}(\mathcal{H})$ . Then the adjoint of  $T$ , defined by

$$\langle Tx | y \rangle = \langle x | T^\dagger y \rangle,$$

$T$  and  $T^\dagger$  have equal norms

exists. Furthermore,  $\|T\| = \|T^\dagger\|$ .

Another useful theorem that we shall use later is the following.

**16.2.12. Theorem.** Let  $\mathcal{N}(T)$  and  $\mathcal{R}(T)$  denote the null space (kernel) and the range of  $T \in \mathcal{B}(\mathcal{H})$ . We have

$$\mathcal{N}(T^\dagger) = \mathcal{R}(T)^\perp \quad \text{and} \quad \mathcal{N}(T) = \mathcal{R}(T^\dagger)^\perp.$$

*Proof.*  $|x\rangle$  is in  $\mathcal{N}(T^\dagger)$  iff  $T^\dagger|x\rangle = 0$  iff  $\langle y | T^\dagger x \rangle = 0$  for all  $|y\rangle \in \mathcal{H}$ . This holds if and only if  $\langle Ty | x \rangle = 0$  for all  $|y\rangle \in \mathcal{H}$ . This is equivalent to the statement that  $|x\rangle$  is in  $\mathcal{R}(T)^\perp$ . This chain of argument proves that  $\mathcal{N}(T^\dagger) = \mathcal{R}(T)^\perp$ . The second part of the theorem follows from the fact that  $(T^\dagger)^\dagger = T$ .  $\square$

### 16.3 Spectra of Linear Operators

One of the most important results of the theory of finite-dimensional vector spaces is the spectral decomposition theorem developed in Chapter 4. The infinite-dimensional analogue of that theorem is far more encompassing and difficult to prove. It is beyond the scope of this book to develop all the machinery needed for a thorough discussion of the infinite-dimensional spectral theory. Instead, we shall present the central results, and occasionally introduce the reader to the peripheral arguments when they seem to have their own merits.

regular point of an operator  
resolvent set and spectrum of an operator

**16.3.1. Definition.** Let  $T \in \mathcal{L}(\mathcal{H})$ . A complex number  $\lambda$  is called a **regular point** of  $T$  if the operator  $T - \lambda 1$  is bounded and invertible.<sup>3</sup> The set of all regular points of  $T$  is called the **resolvent set** of  $T$ , and is denoted by  $\rho(T)$ . The complement of  $\rho(T)$  in the complex plane is called the **spectrum** of  $T$  and is denoted by  $\sigma(T)$ .

Corollary 16.2.10 implies<sup>4</sup> that if  $T$  is bounded, then  $\rho(T)$  is not empty, and that the spectrum of a bounded linear operator on a Hilbert space is a bounded set. In fact, an immediate consequence of the corollary is that  $\lambda \leq \|T\|$  for all  $\lambda \in \sigma(T)$ .

every eigenvalue of an operator on a vector space of finite dimension is in its spectrum and vice versa

It is instructive to contrast the finite-dimensional case against the implications of the above definition. Recall that because of the dimension theorem, a linear operator on a finite-dimensional vector space  $\mathcal{V}$  is invertible if and only if it is either onto or one-to-one. Now,  $\lambda \in \sigma(T)$  if and only if  $T - \lambda 1$  is not invertible. For finite dimensions, this implies that<sup>5</sup>  $\ker(T - \lambda 1) \neq 0$ . Thus, in finite dimensions,

<sup>3</sup>If  $T$  is bounded, then  $T - \lambda 1$  is automatically bounded.

<sup>4</sup>One can simply choose a  $\lambda$  whose absolute value is greater than  $\|T\|$ .

<sup>5</sup>Note how critical finite-dimensionality is for this implication. In infinite dimensions, an operator can be one-to-one (thus having a zero kernel) without being onto.

$\lambda \in \sigma(\mathbf{T})$  if and only if there is a vector  $|a\rangle$  in  $\mathcal{V}$  such that  $(\mathbf{T} - \lambda\mathbf{1})|a\rangle = 0$ . This is the combined definition of eigenvalue and eigenvector, and is the definition we will have to use to define eigenvalues in infinite dimensions. It follows that in the finite-dimensional case,  $\sigma(\mathbf{T})$  coincides with the set of all eigenvalues of  $\mathbf{T}$ . This is not true for infinite dimensions, as the following example shows.

not all points of  $\sigma(\mathbf{T})$   
are eigenvalues

**16.3.2. Example.** Consider the right-shift operator acting on  $\mathbb{C}^\infty$ . It is easy to see that  $\|\mathbf{T}_r a\| = \|a\|$  for all  $|a\rangle$ . This yields  $\|\mathbf{T}_r\| = 1$ , so that any  $\lambda$  that belongs to  $\sigma(\mathbf{T}_r)$  must be such that  $|\lambda| \leq 1$ . We now show that the converse is also true, i.e., that if  $|\lambda| \leq 1$ , then  $\lambda \in \sigma(\mathbf{T}_r)$ . It is sufficient to show that if  $0 < |\lambda| \leq 1$ , then  $\mathbf{T}_r - \lambda\mathbf{1}$  is not invertible. To establish this, we shall show that  $\mathbf{T}_r - \lambda\mathbf{1}$  is not onto.

Suppose that  $\mathbf{T}_r - \lambda\mathbf{1}$  is onto. Then there must be a vector  $|a\rangle$  such that  $(\mathbf{T}_r - \lambda\mathbf{1})|a\rangle = |e_1\rangle$  where  $|e_1\rangle$  is the first standard basis vector of  $\mathbb{C}^\infty$ . Equating components on both sides yields the recursion relations  $\alpha_1 = -1/\lambda$ , and  $\alpha_{j-1} = \lambda\alpha_j$  for all  $j \geq 2$ . One can readily solve this recursion relation to obtain  $\alpha_j = -1/\lambda^j$  for all  $j$ . This is a contradiction, because

$$\sum_{j=1}^{\infty} |\alpha_j|^2 = \sum_{j=1}^{\infty} \frac{1}{|\lambda|^{2j}}$$

will not converge if  $0 < |\lambda| \leq 1$ , i.e.,  $|a\rangle \notin \mathbb{C}^\infty$ , and therefore  $\mathbf{T}_r - \lambda\mathbf{1}$  is not onto.

We conclude that  $\sigma(\mathbf{T}_r) = \{\lambda \in \mathbb{C} \mid 0 < |\lambda| \leq 1\}$ . If we could generalize the result of the finite-dimensional case to  $\mathbb{C}^\infty$ , we would conclude that all complex numbers whose magnitude is at most 1 are eigenvalues of  $\mathbf{T}_r$ . Quite to our surprise, the following argument shows that  $\mathbf{T}_r$  has no eigenvalues at all!

Suppose that  $\lambda$  is an eigenvalue of  $\mathbf{T}_r$ . Let  $|a\rangle$  be any eigenvector for  $\lambda$ . Since  $\mathbf{T}_r$  preserves the length of a vector, we have  $\langle a|a\rangle = \langle \mathbf{T}_r a|\mathbf{T}_r a\rangle = \langle \lambda a|\lambda a\rangle = |\lambda|^2 \langle a|a\rangle$ . It follows that  $|\lambda| = 1$ . Now write  $|a\rangle = \{\alpha_j\}_{j=1}^\infty$  and let  $\alpha_m$  be the first nonzero term of this sequence. Then  $0 = \langle \mathbf{T}_r a|e_m\rangle = \langle \lambda a|e_m\rangle = \lambda\alpha_m$ . The first equality comes about because  $\mathbf{T}_r|a\rangle$  has its first nonzero term in the  $(m+1)$ st position. Since  $\lambda \neq 0$ , we must have  $\alpha_m = 0$ , which contradicts the choice of this number. ■

## 16.4 Compact Sets

This section deals with some technical concepts, and as such will be rather formal. The central concept of this section is compactness. Although we shall be using compactness sparingly in the sequel, the notion has sufficient application in higher analysis and algebra that it warrants an introductory exposure.

Let us start with the familiar case of the real line, and the intuitive notion of “compactness.” Clearly, we do not want to call the entire real line “compact,” because intuitively, it is not. The next candidate seems to be a “finite” interval. So, first consider the *open* interval  $(a, b)$ . Can we call it compact? Intuition says “yes,” but the following argument shows that it would not be appropriate to call the open interval compact.

Consider the map  $\varphi : \mathbb{R} \rightarrow (a, b)$  given by  $\varphi(t) = \frac{b-a}{2} \tanh t + \frac{b+a}{2}$ . The reader may check that this map is continuous and bijective. Thus, we can continuously

map all of  $\mathbb{R}$  in a one-to-one manner onto  $(a, b)$ . This makes  $(a, b)$  “look” very much<sup>6</sup> like  $\mathbb{R}$ . How can we modify the interval to make it compact? We do not want to alter its finiteness. So, the obvious thing to do is to add the end points. Thus, the interval  $[a, b]$  seems to be a good candidate; and indeed it is.

The next step is to generalize the notion of a closed, finite interval and eventually come up with a definition that can be applied to all spaces. First we need some terminology.

**open ball** **16.4.1. Definition.** An *open ball*  $B_r(x)$  of radius  $r$  and center  $|x\rangle$  in a normed vector space  $\mathcal{V}$  is the set of all vectors in  $\mathcal{V}$  whose distance from  $|x\rangle$  is strictly less than  $r$ :

$$B_r(x) \equiv \{|y\rangle \in \mathcal{V} \mid \|y - x\| < r\}.$$

**open round neighborhood** We call  $B_r(x)$  an *open round neighborhood* of  $|x\rangle$ .

This is a generalization of open interval because

$$(a, b) = \left\{ y \in \mathbb{R} \mid \left| y - \frac{a+b}{2} \right| < \frac{b-a}{2} \right\}.$$

**16.4.2. Example.** A prototype of finite-dimensional normed spaces is  $\mathbb{R}^n$ . An open ball of radius  $r$  centered at  $\mathbf{x}$  is

$$B_r(\mathbf{x}) = \{y \in \mathbb{R}^n \mid (y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2 < r^2\}.$$

Thus, all points *inside* a circle form an open ball in the  $xy$ -plane, and all interior points of a solid sphere form an open ball in space. ■

**bounded subset** **16.4.3. Definition.** A *bounded subset* of a normed vector space is a subset that can be enclosed in an open ball of finite radius.

For example, any region drawn on a piece of paper is a bounded subset of  $\mathbb{R}^2$ , and any “visible” part of our environment is a bounded subset of  $\mathbb{R}^3$  because we can always find a big enough circle or sphere to enclose these subsets.

**open subset** **16.4.4. Definition.** A subset  $\mathcal{O}$  of a normed vector space  $\mathcal{V}$  is called *open* if each of its points (vectors) has an open round neighborhood lying entirely in  $\mathcal{O}$ . A *boundary point* of  $\mathcal{O}$  is a point (vector) in  $\mathcal{V}$  all of whose open round neighborhoods contain points inside and outside  $\mathcal{O}$ . A *closed subset*  $\mathcal{C}$  of  $\mathcal{V}$  is a subset that contains all of its boundary points. The *closure* of a subset  $S$  is the union of  $S$  and all of its boundary points, and is denoted by  $\bar{S}$ .

**boundary point**  
**closed subset and closure**

For example, the boundary of a region drawn on paper consists of all its boundary points. A curve drawn on paper has nothing but boundary points. Every point is also its own boundary. A boundary is always a closed set. In particular, a point is a closed set. In general, an open set cannot contain any boundary points. A frequently used property of a closed set  $\mathcal{C}$  is that a convergent sequence of points of  $\mathcal{C}$  converges to a point in  $\mathcal{C}$ .

<sup>6</sup>In mathematical jargon one says that  $(a, b)$  and  $\mathbb{R}$  are *homeomorphic*.

dense subset **16.4.5. Definition.** A subset  $W$  of a normed vector space  $\mathcal{V}$  is *dense* in  $\mathcal{V}$  if the closure of  $W$  is the entire space  $\mathcal{V}$ . Equivalently,  $W$  is dense if each vector in  $\mathcal{V}$  is infinitesimally close to at least one vector in  $W$ . In other words, given any  $|u\rangle \in \mathcal{V}$  and any  $\epsilon > 0$ , there is a  $|w\rangle \in W$  such that  $\|u - w\| < \epsilon$ , i.e., any vector in  $\mathcal{V}$  can be approximated, with arbitrary accuracy, by a vector in  $W$ .

rational numbers are dense in the real numbers

A paradigm of dense spaces is the set of rational numbers in the normed vector space of real numbers. It is a well-known fact that any real number can be approximated by a rational number with arbitrary accuracy: The decimal (or binary) representation of real numbers is precisely such an approximation. An intuitive way of imagining denseness is that the (necessarily) infinite subset is equal to *almost* all of the set, and its members are scattered “densely” everywhere in the set. The embedding of the rational numbers in the set of real numbers, and how they densely populate that set, is a good mental picture of all dense subsets. A useful property involving the concept of closure and openness has to do with continuous maps between normed vector spaces. Let  $f : \mathcal{H}_1 \rightarrow \mathcal{H}_2$  be a continuous map. Let  $\mathcal{O}_2$  be an open set in  $\mathcal{H}_2$ . Let  $f^{-1}(\mathcal{O}_2)$  denote the inverse image of  $\mathcal{O}_2$ , i.e., all points of  $\mathcal{H}_1$  that are mapped to  $\mathcal{O}_2$ . Let  $|x_1\rangle$  be a vector in  $f^{-1}(\mathcal{O}_2)$ ,  $|x_2\rangle = f(|x_1\rangle)$ , and let  $B_\epsilon(x_2)$  be a ball contained entirely in  $\mathcal{O}_2$ . Then  $f^{-1}(B_\epsilon(x_2))$  contains  $|x_1\rangle$  and lies entirely in  $f^{-1}(\mathcal{O}_2)$ . Because of the continuity of  $f$ , one can now construct an open ball centered at  $|x_1\rangle$  lying entirely in  $f^{-1}(B_\epsilon(x_2))$ , and by inclusion, in  $f^{-1}(\mathcal{O}_2)$ . This shows that every point of  $f^{-1}(\mathcal{O}_2)$  has a round open neighborhood lying entirely in  $f^{-1}(\mathcal{O}_2)$ . Thus,  $f^{-1}(\mathcal{O}_2)$  is an open subset. One can similarly show the corresponding property for closed subsets. We can summarize this in the following:

**16.4.6. Proposition.** Let  $f : \mathcal{H}_1 \rightarrow \mathcal{H}_2$  be continuous. Then the inverse image of an open (closed) subset of  $\mathcal{H}_2$  is an open (closed) subset of  $\mathcal{H}_1$ .

$\rho(\mathbf{T})$  is open, and  $\sigma(\mathbf{T})$  is closed and bounded in  $\mathbb{C}$ .

Consider the resolvent set of a bounded operator  $\mathbf{T}$ . We claim that this set is open in  $\mathbb{C}$ . To see this, note that if  $\lambda \in \rho(\mathbf{T})$ , then  $\mathbf{T} - \lambda\mathbf{1}$  is invertible. On the other hand, Problem 16.1 shows that operators close to an invertible operator are invertible. Thus, if we choose a sufficiently small positive number  $\epsilon$  and consider all complex numbers  $\mu$  within a distance  $\epsilon$  from  $\lambda$ , then all operators of the form  $\mathbf{T} - \mu\mathbf{1}$  are invertible, i.e.,  $\mu \in \rho(\mathbf{T})$ . Therefore, any  $\lambda \in \rho(\mathbf{T})$  has an open round neighborhood in the complex plane all points of which are in the resolvent. This shows that the resolvent set is open. In particular, it cannot contain any boundary points. However,  $\rho(\mathbf{T})$  and  $\sigma(\mathbf{T})$  have to be separated by a common boundary.<sup>7</sup> Since  $\rho(\mathbf{T})$  cannot contain any boundary point,  $\sigma(\mathbf{T})$  must carry the entire boundary. This shows that  $\sigma(\mathbf{T})$  is a closed subset of  $\mathbb{C}$ . Recalling that  $\sigma(\mathbf{T})$  is also bounded, we have the following result.

<sup>7</sup>The spectrum of a bounded operator need not occupy any “area” in the complex plane. It may consist of isolated points or line segments, etc., in which case the spectrum will constitute the entire boundary.

**16.4.7. Proposition.** For any  $T \in \mathcal{B}(\mathcal{H})$  the set  $\rho(T)$  is an open subset of  $\mathbb{C}$  and  $\sigma(T)$  is a closed, bounded subset of  $\mathbb{C}$ .

Let us go back to the notion of compactness. It turns out that the feature of the closed interval  $[a, b]$  most appropriate for generalization is the behavior of infinite sequences of numbers lying in the interval. More specifically, let  $\{\alpha_i\}_{i=1}^{\infty}$  be a sequence of *infinitely many* real numbers all lying in the interval  $[a, b]$ . It is intuitively clear that since there is not enough room for these points to stay away from each other, they will have to crowd around a number of points in the interval. For example, the sequence

$$\left\{ (-1)^n \frac{2n+1}{4n} \right\}_{n=1}^{\infty} = \left\{ -\frac{3}{4}, +\frac{5}{8}, -\frac{7}{12}, +\frac{9}{16}, \dots \right\}$$

in the interval  $[-1, +1]$  crowds around the two points  $-\frac{1}{2}$  and  $+\frac{1}{2}$ . In fact, the points with even  $n$  accumulate around  $+\frac{1}{2}$  and those with odd  $n$  crowd around  $-\frac{1}{2}$ . It turns out that all closed intervals of  $\mathbb{R}$  have this property, namely, all sequences crowd around some points. To see that open intervals do not share this property consider the open interval  $(0, 1)$ . The sequence  $\left\{ \frac{1}{2n+1} \right\}_{n=1}^{\infty} = \left\{ \frac{1}{3}, \frac{1}{5}, \dots \right\}$  clearly crowds only around zero, which is not a point of the interval. But we already know that open intervals are not compact.

**16.4.8. Definition.** (Bolzano–Weierstrass property) A subset  $\mathcal{K}$  of a normed vector space is called **compact** if every (infinite) sequence in  $\mathcal{K}$  has a convergent subsequence.

The reason for the introduction of a subsequence in the definition is that a sequence may have many points to which it converges. But no matter how many of these points there may exist, one can always obtain a convergent subsequence by choosing from among the points in the sequence. For instance, in the example above, one can choose the subsequence consisting of elements for which  $n$  is even. This subsequence converges to the single point  $+\frac{1}{2}$ .

An important theorem in real analysis characterizes all compact sets in  $\mathbb{R}^n$ .<sup>8</sup>

**16.4.9. Theorem.** (BWHB theorem) A subset of  $\mathbb{R}^n$  is compact if and only if it is closed and bounded.

We showed earlier that the spectrum of a bounded linear operator is closed and  $\sigma(T)$  is compact bounded. Identifying  $\mathbb{C}$  with  $\mathbb{R}^2$ , the BWHB theorem implies that

<sup>8</sup>BWHB stands for Bolzano, Weierstrass, Heine, and Borel. Bolzano and Weierstrass proved that any closed and bounded subset of  $\mathbb{R}$  has the Bolzano–Weierstrass property. Heine and Borel abstracted the notion of compactness in terms of open sets, and showed that a closed bounded subset of  $\mathbb{R}$  is compact. The BWHB theorem as applied to  $\mathbb{R}$  is usually called the Heine–Borel theorem (although some authors call it the Bolzano–Weierstrass theorem). Since the Bolzano–Weierstrass property and compactness are equivalent, we have decided to choose BWHB as the name of our theorem.

**16.4.10. Box.** *The spectrum of a bounded linear operator is a compact subset of  $\mathbb{C}$ .*

An immediate consequence of the BWHB Theorem is that every bounded subset of  $\mathbb{R}^n$  has a compact closure. Since  $\mathbb{R}^n$  is a prototype of all *finite-dimensional* (normed) vector spaces, the same statement is true for all such vector spaces. What is interesting is that the statement indeed *characterizes* the normed space:

criterion for  
finite-dimensionality

**16.4.11. Theorem.** *A normed vector space is finite-dimensional if and only if every bounded subset has a compact closure.*

This result can also be applied to *subspaces* of a normed vector space: A subspace  $\mathcal{W}$  of a normed vector space  $\mathcal{V}$  is finite-dimensional if and only if every bounded subset of  $\mathcal{W}$  has a compact closure in  $\mathcal{W}$ . A useful version of this property is stated in terms of sequences of points (vectors):

**16.4.12. Theorem.** *A subspace  $\mathcal{W}$  of a normed vector space  $\mathcal{V}$  is finite dimensional if and only if every bounded sequence in  $\mathcal{W}$  has a convergent subsequence in  $\mathcal{W}$ .*

**Karl Theodor Wilhelm Weierstrass** (1815–1897) was both the greatest analyst and the world's foremost teacher of advanced mathematics of the last third of the nineteenth century. His career was also remarkable in another way—and a consolation to all “late starters”—for he began the solid part of his professional life at the age of almost 40, when most mathematicians are long past their creative years.

His father sent him to the University of Bonn to qualify for the higher ranks of the Prussian civil service by studying law and commerce. But Karl had no interest in these subjects. He infuriated his father by rarely attending lectures, getting poor grades, and instead, becoming a champion beer drinker. He did manage to become a superb fencer, but when he returned home, he had no degree.

In order to earn his living, he made a fresh start by teaching mathematics, physics, botany, German, penmanship, and gymnastics to the children of several small Prussian towns during the day. During the nights, however, he mingled with the intellectuals of the past, particularly the great Norwegian mathematician Abel. His remarkable research on Abelian functions was carried on for years without the knowledge of another living soul; he didn't discuss it with anyone at all, or submit it for publication in the mathematical journals of the day.

All this changed in 1854 when Weierstrass at last published an account of his research on Abelian functions. This paper caught the attention of an alert professor at the University of Königsberg who persuaded his university to award Weierstrass an honorary doctor's degree. The Ministry of Education granted Weierstrass a year's leave of absence with pay





to continue his research, and the next year he was appointed to the University of Berlin, where he remained the rest of his life.

Weierstrass's great creative talents were evenly divided between his thinking and his teaching. The student notes of his lectures, and copies of these notes, and copies of copies, were passed from hand to hand throughout Europe and even America. Like Gauss he was indifferent to fame, but unlike Gauss he endeared himself to generations of students by the generosity with which he encouraged them to develop and publish, and receive credit for, ideas and theorems that he essentially originated himself. Among Weierstrass's students and followers were Cantor, Schwarz, Hölder, Mittag-Leffler, Sonja Kovalevskaya (Weierstrass's favorite student), Hilbert, Max Planck, Willard Gibbs, and many others.

In 1885 he published the famous theorem now called the **Weierstrass approximation theorem** (see Theorems 5.2.3 and 8.1.1), which was given a far-reaching generalization, with many applications, by the modern American mathematician M. H. Stone.

The quality that came to be known as "Weierstrassian rigor" was particularly visible in his contributions to the foundations of real analysis. He refused to accept any statement as "intuitively obvious," but instead demanded ironclad proof based on explicit properties of the real numbers. The careful reasoning required for these proofs was founded on a crucial property of the real numbers now known as the **BWHB theorem**.

We shall need the following proposition in our study of compact operators:

**16.4.13. Proposition.** *Let  $\mathcal{W}$  be a closed proper subspace of  $\mathcal{H}$  and  $\delta$  an arbitrary nonnegative number with  $0 \leq \delta < 1$ . Then there exists a unit vector  $|v_0\rangle \in \mathcal{H}$  such that*

$$\|x - v_0\| \geq \delta \quad \forall |x\rangle \in \mathcal{W}.$$

*Proof.* Choose a vector  $|v\rangle$  in  $\mathcal{H}$  but not in  $\mathcal{W}$  and let

$$d = \min\{\|v - x\| \mid |x\rangle \in \mathcal{W}\}.$$

We claim that  $d > 0$ . To show this, assume otherwise. Then for each (large)  $n$  and (sufficiently small)  $\epsilon$ , we could find *distinct* vectors  $\{|x_n\rangle\}$  whose distance from  $|v\rangle$  would be  $\epsilon/n$  and for which the sequence  $\{|x_n\rangle\}$  would have  $|v\rangle$  as a limit. Closure of  $\mathcal{W}$  would then imply that  $|v\rangle$  is in  $\mathcal{W}$ , a contradiction. So,  $d > 0$ .

Now, for any  $|x_0\rangle \in \mathcal{W}$ , let

$$|u\rangle \equiv |x\rangle - \frac{|v\rangle - |x_0\rangle}{\|v - x_0\|} = \frac{\overbrace{(\|v - x_0\| |x\rangle + |x_0\rangle)}^{\in \mathcal{W}} - |v\rangle}{\|v - x_0\|}$$

and note that by the definition of  $d$ , the norm of the numerator is larger than  $d$ . Therefore,  $\|u\| \geq d/\|v - x_0\|$  for every  $|x\rangle, |x_0\rangle \in \mathcal{W}$ . If we choose  $|x_0\rangle$  such that  $\|v - x_0\| < d\delta^{-1}$ , which is possible because  $d\delta^{-1} > d$ , then  $\|u\| \geq \delta$  for all  $|x\rangle \in \mathcal{W}$ . Now let  $|v_0\rangle = (|v\rangle - |x_0\rangle)/\|v - x_0\|$ . □

## 16.5 Compact Operators

It is straightforward to show that if  $\mathcal{K}$  is a compact set in  $\mathcal{H}_1$  and  $f : \mathcal{H}_1 \rightarrow \mathcal{H}_2$  is continuous, then  $f(\mathcal{K})$  (the image of  $\mathcal{K}$ ) is compact in  $\mathcal{H}_2$ . Since all bounded operators are continuous, we conclude that *all bounded operators map compact subsets onto compact subsets*. There is a special subset of  $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$  that deserves particular attention.

**compact operator** **16.5.1. Definition.** *An operator  $\mathbf{K} \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$  is called a compact operator if it maps a bounded subset of  $\mathcal{H}_1$  onto a subset of  $\mathcal{H}_2$  with compact closure.*

Since we will be dealing with function spaces, and since it is easier to deal with sequences of functions than with subsets of the space of functions, we find it more useful to have a definition of compact operators in terms of sequences rather than subsets. Thus, instead of a bounded subset, we take a subset of it consisting of a (necessarily) bounded sequence. The image of this sequence will be a sequence in a compact set, which, by definition, must have a convergent subsequence. We therefore have the following:

**16.5.2. Theorem.** *An operator  $\mathbf{K} \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$  is compact if and only if for any bounded sequence  $\{|x_n\rangle\}$  in  $\mathcal{H}_1$ , the sequence  $\{\mathbf{K}|x_n\rangle\}$  has a convergent subsequence in  $\mathcal{H}_2$ .*

**16.5.3. Example.** Consider  $\mathcal{B}(\mathcal{H})$ , the set of bounded operators on the Hilbert space  $\mathcal{H}$ . If  $\mathbf{K}$  is a compact operator and  $\mathbf{T}$  a bounded operator, then  $\mathbf{KT}$  and  $\mathbf{TK}$  are compact. This is because  $\{\mathbf{T}|x_n\rangle \equiv |y_n\rangle\}$  is a bounded sequence if  $\{|x_n\rangle\}$  is, and  $\{\mathbf{K}|y_n\rangle = \mathbf{KT}|x_n\rangle\}$  has a convergent subsequence, because  $\mathbf{K}$  is compact. For the second part, use the first definition of the compact operator and note that  $\mathbf{K}$  maps bounded sets onto compact sets, which  $\mathbf{T}$  (being continuous) maps onto a compact set. As a special case of this property we note that the product of two compact operators is compact. Similarly, one can show that any linear combination of compact operators is compact. Thus, any polynomial of a compact operator is compact. In particular,

**product of two compact operators is compact**

$$(\mathbf{1} - \mathbf{K})^n = \sum_{j=0}^n \frac{n!}{j!(n-j)!} (-\mathbf{K})^j = \mathbf{1} + \sum_{j=1}^n \frac{n!}{j!(n-j)!} (-\mathbf{K})^j \equiv \mathbf{1} - \mathbf{K}_n,$$

where  $\mathbf{K}_n$  is a compact operator. ■

**finite rank operators** **16.5.4. Definition.** *An operator  $\mathbf{T} \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$  is called a finite rank operator if its range is finite-dimensional.*

The following is clear from Theorem 16.4.12.

**16.5.5. Proposition.** *A finite rank operator is compact.*

In particular, every linear transformation of a finite-dimensional vector space is compact.

linear transformations of finite-dimensional vector spaces are compact

**16.5.6. Theorem.** *If  $\{K_n\} \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$  are compact and  $K \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$  is such that  $\|K - K_n\| \rightarrow 0$ , then  $K$  is compact.*

*Proof.* Let  $\{|x_m\rangle\}$  be a bounded sequence in  $\mathcal{H}_1$ . Let  $\{K_1|x_{m_1}\rangle\}$  be the convergent subsequence guaranteed by the compactness of  $K_1$ . Now,  $\{|x_{m_1}\rangle\}$  is a bounded sequence in  $\mathcal{H}_1$ . It therefore has a subsequence  $\{|x_{m_2}\rangle\}$  such that  $\{K_2|x_{m_2}\rangle\}$  is convergent. Note that  $\{K_1|x_{m_2}\rangle\}$  is also convergent. Continuing this process, we construct the sequence of sequences

$$\{|x_{m_1}\rangle\}, \{|x_{m_2}\rangle\}, \dots, \{|x_{m_k}\rangle\}$$

where each sequence is a subsequence of all sequences preceding it. Furthermore, all the sequences  $\{K_l|x_{m_k}\rangle\}$  for  $l = 1, \dots, k$  are convergent. In particular, if we pick the *diagonal* sequence  $\{|y_m\rangle\} \equiv \{|x_{m_m}\rangle\}$ , then for any  $l \in \mathbb{N}$ , the sequence  $\{K_l|y_m\rangle\}$  converges in  $\mathcal{H}_2$ . To show that  $K$  is compact, we shall establish that  $\{|y_m\rangle\}$  is the subsequence of  $\{|x_m\rangle\}$  such that  $\{K|y_m\rangle\}$  is convergent. Since  $\mathcal{H}_2$  is complete, it is sufficient to show that  $\{K|y_m\rangle\}$  is Cauchy. We use the so-called “ $\epsilon/3$  trick.” Write

$$K|y_m\rangle - K|y_n\rangle = K|y_m\rangle - K_l|y_m\rangle + K_l|y_m\rangle - K_l|y_n\rangle + K_l|y_n\rangle - K|y_n\rangle$$

and use the triangle inequality to obtain

$$\|K y_m - K y_n\| \leq \|K y_m - K_l y_m\| + \|K_l y_m - K_l y_n\| + \|K_l y_n - K y_n\|.$$

By choosing  $m, n$ , and  $l$  large enough, we can make each of the three terms on the RHS smaller than  $\epsilon/3$ ; the first and the third ones because  $K_l \rightarrow K$ , the second one because  $\{K_l|y_n\rangle\}$  is a convergent sequence.  $\square$

Recall that given an orthonormal basis  $\{|e_i\rangle\}_{i=1}^\infty$ , any operator  $T$  on a Hilbert space  $\mathcal{H}$  can be written as  $\sum_{i,j=1}^\infty c_{ij} |e_i\rangle \langle e_j|$ , where  $c_{ij} = \langle e_i | T |e_j\rangle$ . Now let  $K$  be a compact operator and consider the *finite rank* operators

$$K_n \equiv \sum_{i,j=1}^n c_{ij} |e_i\rangle \langle e_j|, \quad c_{ij} = \langle e_i | K |e_j\rangle.$$

Clearly,  $\|K - K_n\| \rightarrow 0$ . The hermitian adjoints  $\{K_n^\dagger\}$  are also of finite rank (therefore, compact). Barring some convergence technicality, we see that  $K^\dagger$ , which is the limit of the sequence of these compact operators, is also compact.

$K$  is compact iff  $K^\dagger$  is

**16.5.7. Theorem.**  *$K$  is a compact operator if and only if  $K^\dagger$  is.*

A particular type of operator occurs frequently in integral equation theory. These are called Hilbert–Schmidt operators and defined as follows:

Hilbert–Schmidt operators

**16.5.8. Definition.** *Let  $\mathcal{H}$  be a Hilbert space, and  $\{|e_i\rangle\}_{i=1}^\infty$  an orthonormal basis. An operator  $T \in \mathcal{L}(\mathcal{H})$  is called *Hilbert–Schmidt* if*

$$\text{tr}(\mathbf{T}^\dagger \mathbf{T}) \equiv \sum_{i=1}^{\infty} \langle e_i | \mathbf{T}^\dagger \mathbf{T} | e_i \rangle = \sum_{i=1}^{\infty} \langle \mathbf{T} e_i | \mathbf{T} e_i \rangle = \sum_{i=1}^{\infty} \|\mathbf{T} e_i\|^2 < \infty.$$

Hilbert–Schmidt operators are compact.

**16.5.9. Theorem.** *Hilbert–Schmidt operators are compact.*

For a proof, see [Rich 78, pp. 242–246].

Hilbert–Schmidt kernel

**16.5.10. Example.** It is time to give a concrete example of a compact (Hilbert–Schmidt) operator. For this, we return to Equation (16.2) with  $w(y) = 1$ , and assume that  $|u\rangle \in \mathcal{L}^2(a, b)$ . Suppose further that the function  $K(x, y)$  is continuous on the closed rectangle  $[a, b] \times [a, b]$  in the  $xy$ -plane (or  $\mathbb{R}^2$ ). Under such conditions,  $K(x, y)$  is called a **Hilbert–Schmidt kernel**. We now show that  $\mathbf{K}$  is compact. First note that due to the continuity of  $K(x, y)$ ,  $\int_a^b \int_a^b |K(x, y)|^2 dx dy < \infty$ . Next, we calculate the trace of  $\mathbf{K}^\dagger \mathbf{K}$ . Let  $\{|e_i\rangle\}_{i=1}^{\infty}$  be any orthonormal basis of  $\mathcal{L}^2(a, b)$ . Then

$$\begin{aligned} \text{tr} \mathbf{K}^\dagger \mathbf{K} &= \sum_{i=1}^{\infty} \langle e_i | \mathbf{K}^\dagger \mathbf{K} | e_i \rangle = \sum_{i=1}^{\infty} \iiint (e_i | x) \langle x | \mathbf{K}^\dagger | y \rangle \langle y | \mathbf{K} | z \rangle \langle z | e_i \rangle dx dy dz \\ &= \iiint \langle y | \mathbf{K} | x \rangle^* \langle y | \mathbf{K} | z \rangle \sum_{i=1}^{\infty} \langle z | e_i \rangle \langle e_i | x \rangle dx dy dz \\ &= \iiint \langle y | \mathbf{K} | x \rangle^* \langle y | \mathbf{K} | z \rangle \underbrace{\langle z | \left( \sum_{i=1}^{\infty} |e_i\rangle \langle e_i| \right) | x \rangle}_{=1} dx dy dz \\ &= \int_a^b \int_a^b |K(x, y)|^2 dx dy < \infty. \end{aligned}$$

**Bernard Bolzano** (1781–1848) was a Czech philosopher, mathematician, and theologian who made significant contributions to both mathematics and the theory of knowledge. He entered the Philosophy Faculty of the University of Prague in 1796, studying philosophy and mathematics. He wrote “My special pleasure in mathematics rested therefore particularly on its purely speculative parts, in other words I prized only that part of mathematics which was at the same time philosophy.”



In the autumn of 1800 he began three years of theological study while he was preparing a doctoral thesis on geometry. He received his doctorate in 1804 for a thesis in which he gave his view of mathematics and what constitutes a correct mathematical proof. In the preface he wrote:

I could not be satisfied with a completely strict proof if it were not derived from concepts which the thesis to be proved contained, but rather made use of some fortuitous, alien, intermediate concept, which is always an erroneous transition to another kind.

Two days after receiving his doctorate Bolzano was ordained a Roman Catholic priest. However, he came to realize that teaching and not ministering defined his true vocation. In the same year, Bolzano was appointed to the chair of philosophy and religion at the University of Prague. Because of his pacifist beliefs and his concern for economic justice, he was suspended from his position in 1819 after pressure from the Austrian government. Bolzano had not given up without a fight but once he was suspended on a charge of heresy he was put under house arrest and forbidden to publish.

Although some of his books had to be published outside Austria because of government censorship, he continued to write and to play an important role in the intellectual life of his country. Bolzano intended to write a series of papers on the foundations of mathematics. He wrote two, the first of which was published. Instead of publishing the second one he decided to "... make myself better known to the learned world by publishing some papers which, by their titles, would be more suited to arouse attention."

Pursuing this strategy he published *Der binomische Lehrsatz ...* (1816) and *Rein analytischer Beweis ...* (1817), which contain an attempt to free calculus from the concept of the infinitesimal. He is clear in his intention stating in the preface of the first that the work is "a sample of a new way of developing analysis." The paper gives a proof of the intermediate value theorem with Bolzano's new approach and in the work he defined what is now called a Cauchy sequence. The concept appears in Cauchy's work four years later but it is unlikely that Cauchy had read Bolzano's work.

After 1817, Bolzano published no further mathematical works for many years. Between the late 1820s and the 1840s, he worked on a major work *Grössenlehre*. This attempt to put the whole of mathematics on a logical foundation was published in parts, while Bolzano hoped that his students would finish and publish the complete work.

His work *Paradoxien des Unendlichen*, a study of paradoxes of the infinite, was published in 1851, three years after his death, by one of his students. The word "set" appears here for the first time. In this work Bolzano gives examples of 1-1 correspondences between the elements of an infinite set and the elements of a proper subset.

Bolzano's theories of mathematical infinity anticipated Georg Cantor's theory of infinite sets. It is also remarkable that he gave a function which is nowhere differentiable yet everywhere continuous.

## 16.6 Spectrum of Compact Operators

Our next task is to investigate the spectrum  $\sigma(\mathbf{K})$  of a compact operator  $\mathbf{K}$  on a Hilbert space  $\mathcal{H}$ . We are particularly interested in the set of eigenvalues and eigenvectors of compact operators. Recall that every eigenvalue of an operator on a vector space of finite dimension is in its spectrum, and that every point of the spectrum is an eigenvalue (see page 457). In general, the second statement is not true. In fact, we saw that the right-shift operator had no eigenvalue at all, yet its spectrum was the entire unit disk of the complex plane.

We first observe that  $0 \in \sigma(\mathbf{K})$ , because otherwise  $0 \in \rho(\mathbf{K})$ , which implies that  $\mathbf{K} = \mathbf{K} - 0\mathbf{1}$  is invertible with inverse  $\mathbf{K}^{-1}$ . The product of two compact operators (in fact, the product of a compact and a bounded operator) is compact (see

Example 16.5.3). This yields a contradiction<sup>9</sup> because the unit operator cannot be compact: It maps a bounded sequence to itself, not to a sequence with a convergent subsequence.

$0 \in \sigma(\mathbf{K})$  if  $\mathbf{K}$  is compact

**16.6.1. Proposition.** For any compact operator  $\mathbf{K} \in \mathcal{B}(\mathcal{H})$  on an infinite dimensional Hilbert space, we have  $0 \in \sigma(\mathbf{K})$ .

To proceed, we note that eigenvectors of  $\mathbf{K}$  corresponding to the eigenvalue  $\lambda$  belong to the null space of  $\mathbf{K} - \lambda\mathbf{1}$ . So, let<sup>10</sup>

$$\begin{aligned} \mathcal{N}_\lambda &\equiv \ker(\mathbf{K} - \lambda\mathbf{1}), & \mathcal{R}_\lambda &\equiv \text{Range}(\mathbf{K} - \lambda\mathbf{1}), \\ \mathcal{N}_\lambda^\dagger &\equiv \ker(\mathbf{K}^\dagger - \lambda^*\mathbf{1}), & \mathcal{R}_\lambda^\dagger &\equiv \text{Range}(\mathbf{K}^\dagger - \lambda^*\mathbf{1}). \end{aligned}$$

**16.6.2. Theorem.**  $\mathcal{N}_\lambda$  and  $\mathcal{N}_\lambda^\dagger$  are finite-dimensional subspaces of  $\mathcal{H}$ . Furthermore,  $\mathcal{R}_\lambda^\perp = \mathcal{N}_\lambda^\dagger$ .

*Proof.* We use Theorem 16.4.12. Let  $\{|x_n\rangle\}$  be a bounded sequence in  $\mathcal{N}_\lambda$ . Since  $\mathbf{K}$  is compact,  $\{\mathbf{K}|x_n\rangle = \lambda|x_n\rangle\}$  has a convergent subsequence. So  $\{|x_n\rangle\}$  has a convergent subsequence. This subsequence will converge to a vector in  $\mathcal{N}_\lambda$  if the latter is closed. But this follows from Proposition 16.4.6, continuity of  $\mathbf{K} - \lambda\mathbf{1}$ , the fact that  $\mathcal{N}_\lambda$  is the inverse image of the zero vector, and the fact that any single point of a space, such as the zero vector, is a closed subset. Finite-dimensionality of  $\mathcal{N}_\lambda^\dagger$  follows from the compactness of  $\mathbf{K}^\dagger$  and a similar argument as above.

To show the second statement, we observe that for any bounded operator  $\mathbf{T}$ , we have<sup>11</sup>  $|u\rangle \in \mathcal{T}(\mathcal{H})^\perp$  iff  $\langle u|v\rangle = 0$  for all  $|v\rangle \in \mathcal{T}(\mathcal{H})$  iff  $\langle u|\mathbf{T}x\rangle = 0$  for all  $|x\rangle \in \mathcal{H}$  iff  $\langle \mathbf{T}^\dagger u|x\rangle = 0$  for all  $|x\rangle \in \mathcal{H}$  iff  $\mathbf{T}^\dagger|u\rangle = 0$  iff  $|u\rangle \in \ker \mathbf{T}^\dagger$ . This shows that  $\mathcal{T}(\mathcal{H})^\perp = \ker \mathbf{T}^\dagger$ . The desired result is obtained by letting  $\mathbf{T} = \mathbf{K} - \lambda\mathbf{1}$  and noting that  $(\mathcal{W}^\perp)^\perp = \mathcal{W}$  for any subspace  $\mathcal{W}$  of a Hilbert space.  $\square$

We note that  $\mathcal{N}_\lambda$  is the eigenspace of  $\mathbf{K}$  corresponding to the eigenvalue  $\lambda$ . However, it may well happen that zero is the only number in  $\sigma(\mathbf{K})$ . In the finite-dimensional case, this corresponds to the case where the matrix representation of the operator is not diagonalizable. In such a case, the standard procedure is to look at generalized eigenvectors. We do the same in the case of compact operators.

generalized eigenvector

**16.6.3. Definition.** A vector  $|u\rangle$  is a *generalized eigenvector* of  $\mathbf{K}$  of order  $m$  if  $(\mathbf{K} - \lambda\mathbf{1})^{m-1}|u\rangle \neq 0$  but  $(\mathbf{K} - \lambda\mathbf{1})^m|u\rangle = 0$ . The set of such vectors, i.e., the null space of  $(\mathbf{K} - \lambda\mathbf{1})^m$ , will be denoted by  $\mathcal{N}_\lambda^{(m)}$ .

It is clear that

$$\{0\} \equiv \mathcal{N}_\lambda^{(0)} \subseteq \mathcal{N}_\lambda \equiv \mathcal{N}_\lambda^{(1)} \subseteq \mathcal{N}_\lambda^{(2)} \subseteq \dots \subseteq \mathcal{N}_\lambda^{(m)} \subseteq \mathcal{N}_\lambda^{(m+1)} \dots \tag{16.6}$$

<sup>9</sup>Our conclusion is valid only in infinite dimensions. In finite dimensions, all operators, including  $\mathbf{1}$ , are compact.

<sup>10</sup>In what follows, we assume that  $\lambda \neq 0$ .

<sup>11</sup>Recall that  $\mathcal{T}(\mathcal{H})$  is the range of the operator  $\mathbf{T}$ .

and each  $\mathcal{N}_\lambda^{(m)}$  is a subspace of  $\mathcal{H}$ . In general, a subspace with higher index is larger than those with lower index. If there happens to be an equality at one link of the above chain, then the equality continues all the way to the right ad infinitum. To see this, let  $p$  be the first integer for which the equality occurs, and let  $n > p$  be arbitrary. Suppose  $|u\rangle \in \mathcal{N}_\lambda^{(n+1)}$ . Then  $(\mathbf{K} - \lambda\mathbf{1})^{p+1}[(\mathbf{K} - \lambda\mathbf{1})^{n-p}|u\rangle] = (\mathbf{K} - \lambda\mathbf{1})^{n+1}|u\rangle = 0$ . It follows that  $(\mathbf{K} - \lambda\mathbf{1})^{n-p}|u\rangle$  is in  $\mathcal{N}_\lambda^{(p+1)}$ . But  $\mathcal{N}_\lambda^{(p)} = \mathcal{N}_\lambda^{(p+1)}$ . So

$$(\mathbf{K} - \lambda\mathbf{1})^n |u\rangle = (\mathbf{K} - \lambda\mathbf{1})^p [(\mathbf{K} - \lambda\mathbf{1})^{n-p} |u\rangle] = 0.$$

Thus every vector in  $\mathcal{N}_\lambda^{(n+1)}$  is also in  $\mathcal{N}_\lambda^{(n)}$ . This fact and the above chain imply that  $\mathcal{N}_\lambda^{(n)} = \mathcal{N}_\lambda^{(n+1)}$  for all  $n > p$ .

**16.6.4. Theorem.** *The subspace<sup>11</sup>  $\mathcal{N}_\lambda^{(n)}$  is finite-dimensional for each  $n$ . Moreover, there is an integer  $p$  such that*

$$\mathcal{N}_\lambda^{(n)} \neq \mathcal{N}_\lambda^{(n+1)} \quad \text{for } n = 0, 1, 2, \dots, p - 1$$

but  $\mathcal{N}_\lambda^{(n)} = \mathcal{N}_\lambda^{(n+1)}$  for all  $n \geq p$ .

*Proof.* For the first part, use the result of Example 16.5.3 to show that  $(\mathbf{K} - \lambda\mathbf{1})^n = \mathbf{K}_n - \lambda^n \mathbf{1}$  where  $\mathbf{K}_n$  is compact. Now repeat the proof of Theorem 16.6.2 for  $\mathbf{K}_n$ .

If the integer  $p$  exists, the second part of the theorem follows from the discussion preceding the theorem. To show the existence of  $p$ , suppose, to the contrary, that  $\mathcal{N}_\lambda^{(n)} \neq \mathcal{N}_\lambda^{(n+1)}$  for every positive integer  $n$ . This means that for every  $n$ , we can find a (unit) vector  $|v_n\rangle \in \mathcal{N}_\lambda^{(n+1)}$  that is not in  $\mathcal{N}_\lambda^{(n)}$  and that by Proposition 16.4.13 has the property

$$\|v_n - w\| \geq \frac{1}{2} \quad \forall |w\rangle \in \mathcal{N}_\lambda^{(n)}.$$

We thus obtain a bounded sequence  $\{|v_n\rangle\}$ . Let us apply  $\mathbf{K}$  to this sequence. If  $j > l$ , then  $(|v_j\rangle - |v_l\rangle) \in \mathcal{N}_\lambda^{(j+1)}$  by the construction of  $|v_j\rangle$  and the fact that  $\mathcal{N}_\lambda^{(l+1)} \subseteq \mathcal{N}_\lambda^{(j+1)}$ . Furthermore,

$$(\mathbf{K} - \lambda\mathbf{1})^j \{(\mathbf{K} - \lambda\mathbf{1})(|v_j\rangle - |v_l\rangle)\} = (\mathbf{K} - \lambda\mathbf{1})^{j+1} (|v_j\rangle - |v_l\rangle) = 0,$$

but

$$(\mathbf{K} - \lambda\mathbf{1})^{j-1} \{(\mathbf{K} - \lambda\mathbf{1})(|v_j\rangle - |v_l\rangle)\} = (\mathbf{K} - \lambda\mathbf{1})^j (|v_j\rangle - |v_l\rangle) \neq 0$$

<sup>11</sup>For infinite dimensions, the fact that linear combinations of a subset belong to the subset is not sufficient to make that subset into a subspace. The subset must also be closed. We normally leave out the rather technical proof of closure.

by the definition of  $\mathcal{N}_\lambda^{(j+1)}$ . Therefore,  $(\mathbf{K} - \lambda\mathbf{1})(|v_j\rangle - |v_l\rangle) \in \mathcal{N}_\lambda^{(j)}$ . Now note that

$$\mathbf{K}|v_j\rangle - \mathbf{K}|v_l\rangle = \lambda \underbrace{\left\{ \frac{1}{\lambda}(\mathbf{K} - \lambda\mathbf{1})|v_j\rangle - \frac{1}{\lambda}(\mathbf{K} - \lambda\mathbf{1})|v_l\rangle \right\}}_{\in \mathcal{N}_\lambda^{(j+1)}, \notin \mathcal{N}_\lambda^{(l+1)}} + |v_j\rangle - |v_l\rangle.$$

It follows from Proposition 16.4.13 that the norm of the vector in curly brackets is larger than  $\frac{1}{2}$ . Hence,  $\|\mathbf{K}|v_j\rangle - \mathbf{K}|v_l\rangle\| \geq |\lambda|/2$ , i.e., since  $j$  and  $l$  are arbitrary, the sequence  $\{\mathbf{K}|v_n\rangle\}$  does not have a convergent subsequence. This contradicts the fact that  $\mathbf{K}$  is compact. So, there must exist a  $p$  such that  $\mathcal{N}_\lambda^{(p)} = \mathcal{N}_\lambda^{(p+1)}$ .  $\square$

We also need the range of various powers of  $\mathbf{K} - \lambda\mathbf{1}$ . Thus, let  $\mathcal{R}_\lambda^{(n)} = \text{Range}(\mathbf{K} - \lambda\mathbf{1})^n$ . One can show that

$$\mathcal{H} = \mathcal{R}_\lambda^{(0)} \supseteq \mathcal{R}_\lambda^{(1)} \supseteq \dots \supseteq \mathcal{R}_\lambda^{(n)} \supseteq \mathcal{R}_\lambda^{(n+1)} \supseteq \dots$$

**16.6.5. Theorem.** *Each  $\mathcal{R}_\lambda^{(n)}$  is a subspace of  $\mathcal{H}$ . Moreover, there is an integer  $q$  such that  $\mathcal{R}_\lambda^{(n)} \neq \mathcal{R}_\lambda^{(n+1)}$  for  $n = 0, 1, \dots, q - 1$ , but  $\mathcal{R}_\lambda^{(n)} = \mathcal{R}_\lambda^{(n+1)}$  for all  $n \geq q$ .*

*Proof.* The proof is similar to that of Theorem 16.6.4. The only extra step needed is to show that  $\mathcal{R}_\lambda^{(n)}$  is closed. We shall not reproduce this step.  $\square$

**16.6.6. Theorem.** *Let  $q$  be the integer of Theorem 16.6.5. Then*

1.  $\mathcal{H} = \mathcal{N}_\lambda^{(q)} \oplus \mathcal{R}_\lambda^{(q)}$ .
2.  $\mathcal{N}_\lambda^{(q)}$  and  $\mathcal{R}_\lambda^{(q)}$  are invariant subspaces of  $\mathbf{K}$ .
3. The only vector in  $\mathcal{R}_\lambda^{(q)}$  that  $\mathbf{K} - \lambda\mathbf{1}$  maps to zero is the zero vector. In fact, when restricted to  $\mathcal{R}_\lambda^{(q)}$ , the operator  $\mathbf{K} - \lambda\mathbf{1}$  is invertible.

*Proof.* (1) Recall that  $\mathcal{H} = \mathcal{N}_\lambda^{(q)} \oplus \mathcal{R}_\lambda^{(q)}$  means that every vector of  $\mathcal{H}$  can be written as the sum of a vector in  $\mathcal{N}_\lambda^{(q)}$  and a vector in  $\mathcal{R}_\lambda^{(q)}$ , and the only vector common to both subspaces is zero. We show the latter property first. In fact, we show that  $\mathcal{N}_\lambda^{(m)} \cap \mathcal{R}_\lambda^{(q)} = \{0\}$  for any integer  $m$ . Suppose  $|x\rangle$  is in this intersection. For each  $n \geq q$ , there must be a vector  $|x_n\rangle$  in  $\mathcal{H}$  such that  $|x\rangle = (\mathbf{K} - \lambda\mathbf{1})^n |x_n\rangle$  because  $\mathcal{R}_\lambda^{(n)} = \mathcal{R}_\lambda^{(q)}$  for  $n \geq q$ . If  $|x\rangle \neq 0$ , then  $|x_n\rangle \notin \mathcal{N}_\lambda^{(n)}$  for each  $n$ . Now let  $r$  be the larger of the two integers  $(p, q)$  where  $p$  is the integer of Theorem 16.6.4. Then

$$|x_r\rangle \notin \mathcal{N}_\lambda^{(r)}. \tag{16.7}$$



From

$$0 = (\mathbf{K} - \lambda \mathbf{1})^m |x\rangle = (\mathbf{K} - \lambda \mathbf{1})^{m+r} |x_r\rangle$$

and

$$0 \neq (\mathbf{K} - \lambda \mathbf{1})^{m-1} |x\rangle = (\mathbf{K} - \lambda \mathbf{1})^{m+r-1} |x_n\rangle$$

it follows that  $|x_r\rangle \in \mathcal{N}_\lambda^{(m+r)}$ . But  $\mathcal{N}_\lambda^{(m+r)} = \mathcal{N}_\lambda^{(r)} = \mathcal{N}_\lambda^{(p)}$ , contradicting Equation (16.7). We conclude that  $|x\rangle$  must be zero. By the definition of  $\mathcal{R}_\lambda^{(q)}$ , for any vector  $|z\rangle$  in  $\mathcal{H}$ , we have that  $(\mathbf{K} - \lambda \mathbf{1})^q |z\rangle \in \mathcal{R}_\lambda^{(q)}$ . Since  $\mathcal{R}_\lambda^{(q)} = \mathcal{R}_\lambda^{(2q)}$ , there must be a vector  $|y\rangle \in \mathcal{H}$  such that  $(\mathbf{K} - \lambda \mathbf{1})^q |z\rangle = (\mathbf{K} - \lambda \mathbf{1})^{2q} |y\rangle$  or  $(\mathbf{K} - \lambda \mathbf{1})^q [|z\rangle - (\mathbf{K} - \lambda \mathbf{1})^q |y\rangle] = 0$ . This shows that  $|z\rangle - (\mathbf{K} - \lambda \mathbf{1})^q |y\rangle$  is in  $\mathcal{N}_\lambda^{(q)}$ . On the other hand,

$$|z\rangle = \underbrace{[|z\rangle - (\mathbf{K} - \lambda \mathbf{1})^q |y\rangle]}_{\in \mathcal{N}_\lambda^{(q)}} + \underbrace{(\mathbf{K} - \lambda \mathbf{1})^q |y\rangle}_{\in \mathcal{R}_\lambda^{(q)}}$$

and the first part of the theorem is done.

(2) For the second part, we simply note that  $(\mathbf{K} - \lambda \mathbf{1})\mathcal{N}_\lambda^{(k)} \subseteq \mathcal{N}_\lambda^{(k-1)} \subseteq \mathcal{N}_\lambda^{(k)}$ , and that

$$\mathbf{K}(\mathcal{N}_\lambda^{(q)}) = (\mathbf{K} - \lambda \mathbf{1} + \lambda \mathbf{1})(\mathcal{N}_\lambda^{(q)}) = \underbrace{(\mathbf{K} - \lambda \mathbf{1})(\mathcal{N}_\lambda^{(q)})}_{\subset \mathcal{N}_\lambda^{(q)}} + \underbrace{\lambda \mathbf{1}(\mathcal{N}_\lambda^{(q)})}_{\subset \mathcal{N}_\lambda^{(q)}} \subset \mathcal{N}_\lambda^{(q)}.$$

Similarly,

$$\mathbf{K}(\mathcal{R}_\lambda^{(q)}) = (\mathbf{K} - \lambda \mathbf{1} + \lambda \mathbf{1})(\mathcal{R}_\lambda^{(q)}) = \underbrace{(\mathbf{K} - \lambda \mathbf{1})(\mathcal{R}_\lambda^{(q)})}_{\subset \mathcal{R}_\lambda^{(q+1)} \subset \mathcal{R}_\lambda^{(q)}} + \underbrace{\lambda \mathbf{1}(\mathcal{R}_\lambda^{(q)})}_{\subset \mathcal{R}_\lambda^{(q)}} \subset \mathcal{R}_\lambda^{(q)}.$$

(3) Suppose  $|z\rangle \in \mathcal{R}_\lambda^{(q)}$  and  $(\mathbf{K} - \lambda \mathbf{1})|z\rangle = 0$ . Then  $|z\rangle = (\mathbf{K} - \lambda \mathbf{1})^q |y\rangle$  for some  $|y\rangle$  in  $\mathcal{H}$ , and  $0 = (\mathbf{K} - \lambda \mathbf{1})|z\rangle = (\mathbf{K} - \lambda \mathbf{1})^{q+1} |y\rangle$ , or  $|y\rangle \in \mathcal{N}_\lambda^{(q+1)}$ . From part (1)—with  $m = q + 1$ —we conclude that  $|z\rangle = 0$ . It follows that  $\mathbf{K} - \lambda \mathbf{1}$  is injective (or 1-1). We also have

$$\begin{aligned} (\mathbf{K} - \lambda \mathbf{1})\mathcal{R}_\lambda^{(q)} &= (\mathbf{K} - \lambda \mathbf{1})(\mathbf{K} - \lambda \mathbf{1})^q (\mathcal{H}) \\ &= (\mathbf{K} - \lambda \mathbf{1})^{q+1} (\mathcal{H}) = \mathcal{R}_\lambda^{(q+1)} = \mathcal{R}_\lambda^{(q)}. \end{aligned}$$

Therefore, when restricted to  $\mathcal{R}_\lambda^{(q)}$ , the operator  $\mathbf{K} - \lambda \mathbf{1}$  is surjective (or onto) as well. Thus,  $(\mathbf{K} - \lambda \mathbf{1}) : \mathcal{R}_\lambda^{(q)} \rightarrow \mathcal{R}_\lambda^{(q)}$  is bijective, and therefore has an inverse.  $\square$

**16.6.7. Corollary.** *The two integers  $p$  and  $q$  introduced in Theorems 16.6.4 and 16.6.5 are equal.*

*Proof.* The proof is left as a problem for the reader (see Problem 16.5).  $\square$

The next theorem characterizes the spectrum of a compact operator completely. In order to prove it, we need the following lemma.

**16.6.8. Lemma.** *Let  $\mathbf{K}_\lambda : \mathcal{R}_\lambda^{(q)} \rightarrow \mathcal{R}_\lambda^{(q)}$  be the restriction of  $\mathbf{K}$  to  $\mathcal{R}_\lambda^{(q)}$ . Then:*

1. *Each nonzero point of  $\sigma(\mathbf{K})$  is an eigenvalue of  $\mathbf{K}$  whose eigenspace is finite-dimensional.*
2.  $\sigma(\mathbf{K}_\lambda) = \sigma(\mathbf{K})$ .
3. *Every infinite sequence in  $\sigma(\mathbf{K})$  converges to zero.*

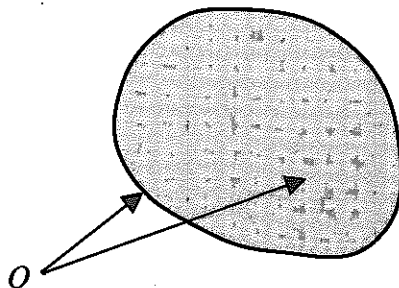
*Proof.* (1) If  $\lambda \neq 0$  is not an eigenvalue of  $\mathbf{K}$ , the null space of  $\mathbf{K} - \lambda\mathbf{1}$  is zero. This says that  $\{0\} = \mathcal{N}_\lambda^{(0)} = \mathcal{N}_\lambda^{(1)} = \dots$ , i.e.,  $p = q = 0$ . From Theorem 16.6.6, we conclude that  $\mathcal{H} = \mathcal{N}_\lambda^{(0)} \oplus \mathcal{R}_\lambda^{(0)} = \mathcal{R}_\lambda^{(1)}$ . Therefore,  $\mathbf{K} - \lambda\mathbf{1}$  is onto. Part (3) of Theorem 16.6.6 shows that  $\mathbf{K} - \lambda\mathbf{1}$  is one-to-one. Thus,  $\mathbf{K} - \lambda\mathbf{1}$  is invertible, and  $\lambda \in \rho(\mathbf{K})$ . So,  $\lambda \notin \sigma(\mathbf{K})$ .

(2) Clearly  $\sigma(\mathbf{K}_\lambda) \subseteq \sigma(\mathbf{K})$ . To show the reverse inclusion, first note that  $\mathcal{R}_\lambda^{(q)}$  is infinite-dimensional because  $\mathcal{N}_\lambda^{(q)}$  has finite dimension. Thus by Proposition 16.6.1,  $0 \in \sigma(\mathbf{K}_\lambda)$ . Now let  $\mu$ —nonzero and distinct from  $\lambda$ —be in  $\sigma(\mathbf{K})$ . By part (1)  $\mu$  is an eigenvalue of  $\mathbf{K}$ , so there is a vector  $|u\rangle \in \mathcal{H}$  such that  $\mathbf{K}|u\rangle = \mu|u\rangle$ . We also have  $(\mathbf{K} - \lambda\mathbf{1})|u\rangle = (\mu - \lambda)|u\rangle$ , or  $(\mathbf{K} - \lambda\mathbf{1})^q|u\rangle = (\mu - \lambda)^q|u\rangle$ . Thus,  $(\mu - \lambda)^q|u\rangle$  (and, therefore  $|u\rangle$ ) is in  $\mathcal{R}_\lambda^{(q)}$ . Therefore, we can restrict  $\mathbf{K}$  to  $\mathcal{R}_\lambda^{(q)}$ , i.e., we can write  $\mathbf{K}|u\rangle = \mu|u\rangle$  as  $\mathbf{K}_\lambda|u\rangle = \mu|u\rangle$ , or  $(\mathbf{K}_\lambda - \mu\mathbf{1})|u\rangle = 0$ . Hence,  $\mu \in \sigma(\mathbf{K}_\lambda)$ . We conclude that every point of  $\sigma(\mathbf{K})$  is a point of  $\sigma(\mathbf{K}_\lambda)$  and  $\sigma(\mathbf{K}) \subseteq \sigma(\mathbf{K}_\lambda)$ .

(3) Let  $\lambda$  be the limit of an infinite sequence in  $\sigma(\mathbf{K}) = \sigma(\mathbf{K}_\lambda)$ . If  $\lambda \neq 0$ ,  $\mathbf{K}_\lambda - \lambda\mathbf{1}$  will be invertible (Theorem 16.6.6 part 3), indicating that  $\lambda \in \rho(\mathbf{K}_\lambda)$ . Since  $\rho(\mathbf{K}_\lambda)$  is open, we can find an open round neighborhood of  $\lambda$  entirely in  $\rho(\mathbf{K}_\lambda)$ . This contradicts the property of a limit of an infinite sequence whereby any neighborhood of the limit contains (infinitely many) other points of the sequence. Therefore, we must conclude that no nonzero  $\lambda$  can be the limit of an infinite sequence in  $\sigma(\mathbf{K})$ .  $\square$

**16.6.9. Theorem.** *Let  $\mathbf{K}$  be a compact operator on an infinite-dimensional Hilbert space  $\mathcal{H}$ . Then*

1.  $0 \in \sigma(\mathbf{K})$ .
2. *Each nonzero point of  $\sigma(\mathbf{K})$  is an eigenvalue of  $\mathbf{K}$  whose eigenspace is finite-dimensional.*
3.  $\sigma(\mathbf{K})$  *is either a finite set or it is a sequence that converges to zero.*



**Figure 16.1** The shaded area represents a convex subset of the vector space. It consists of vectors whose tips lie in the shaded region. It is clear that there is a (unique) vector belonging to the subset whose length is minimum.

*Proof.* (1) was proved in Proposition 16.6.1. (2) was shown in the lemma above.

(3) Let  $\sigma_n(\mathbf{K}) \equiv \{\lambda \in \sigma(\mathbf{K}) \mid |\lambda| \geq 1/n\}$ . Clearly,  $\sigma_n(\mathbf{K})$  must be a finite set, because otherwise the infinite set would constitute a sequence that by compactness of  $\sigma_n(\mathbf{K})$  would have to have (at least) a limit point. By part (2), this limit must be zero, which is not included in  $\sigma_n(\mathbf{K})$ . Let  $\sigma_1(\mathbf{K}) = \{\lambda_i\}_{i=1}^k$ , arranged in order of decreasing absolute value. Next, let  $\lambda_{k+1}, \lambda_{k+2}, \dots$  label the elements of  $\sigma_2(\mathbf{K})$  not accounted for in  $\sigma_1(\mathbf{K})$ , again arranged in decreasing absolute value. If this process stops after a finite number of steps,  $\sigma(\mathbf{K})$  is finite. Otherwise, continue the process to construct a sequence whose limit by necessity is zero.  $\square$

## 16.7 Spectral Theorem for Compact Operators

The finite-dimensional spectral decomposition theorem of Chapter 4 was based on the existence of eigenvalues, eigenspaces, and projection operators. Such existence was guaranteed by the existence of an inner product for any finite-dimensional vector space. The task of establishing spectral decomposition for infinite-dimensional vector spaces is complicated not only by the possibility of the absence of an inner product, but also by the questions of completeness, closure, and convergence. One can eliminate the first two hindrances by restricting oneself to a Hilbert space. However, even so, one has to deal with other complications of infinite dimensions.

As an example, consider the relation  $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$ , which is trivially true for any subspace  $\mathcal{W}$  in finite dimensions once an orthonormal basis is chosen. Recall that the procedure for establishing this relation is to complement a basis of  $\mathcal{W}$  to produce a basis for the whole space. In an infinite-dimensional Hilbert space, we do not know a priori how to complement the basis of a subspace (which may be infinite-dimensional). Thus, one has to prove the existence of the orthogonal complement of a subspace. Without going into details, we sketch the proof. First a definition:

convex subset **16.7.1. Definition.** A convex subset  $E$  of a vector space is a collection of vectors such that if  $|u\rangle$  and  $|v\rangle$  are in  $E$ , then  $|u\rangle - t(|u\rangle - |v\rangle)$  is also in  $E$  for all  $0 \leq t \leq 1$ .

Intuitively, any two points of a convex subset can be connected by a straight line segment lying entirely in the subset.

Let  $E$  be a convex subset (not a subspace) of a Hilbert space  $\mathcal{H}$ . One can show that there exists a unique vector in  $E$  with minimal norm (see Figure 16.1). Now let  $\mathcal{M}$  be a subspace of  $\mathcal{H}$ . For an arbitrary vector  $|u\rangle$  in  $\mathcal{H}$ , consider the subset  $E = |u\rangle - \mathcal{M}$ , i.e., all vectors of the form  $|u\rangle - |m\rangle$  with  $|m\rangle \in \mathcal{M}$ . Denote the unique vector of minimal norm of  $|u\rangle - \mathcal{M}$  by  $|u\rangle - |Pu\rangle$  with  $|Pu\rangle \in \mathcal{M}$ . One can show that  $|u\rangle - |Pu\rangle$  is orthogonal to  $\mathcal{M}$ , i.e.,  $(|u\rangle - |Pu\rangle) \in \mathcal{M}^\perp$  (see Figure 16.2). Obviously, only the zero vector can be simultaneously in  $\mathcal{M}$  and  $\mathcal{M}^\perp$ . Furthermore, any vector  $|u\rangle$  in  $\mathcal{H}$  can be written as  $|u\rangle = |Pu\rangle + (|u\rangle - |Pu\rangle)$  with  $|Pu\rangle \in \mathcal{M}$  and  $(|u\rangle - |Pu\rangle) \in \mathcal{M}^\perp$ . This shows that  $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ . In words, a Hilbert space is the direct sum of any one of its subspaces and the orthogonal complement of that subspace. The vector  $|Pu\rangle$  so constructed is the projection of  $|u\rangle$  in  $\mathcal{M}$ .

A projection operator  $\mathbf{P}$  can be defined as a linear operator with the property that  $\mathbf{P}^2 = \mathbf{P}$ . One can then show the following.

**16.7.2. Theorem.** The kernel  $\ker \mathbf{P}$  of a projection operator is the orthogonal complement of the range  $\mathbf{P}(\mathcal{H})$  of  $\mathbf{P}$  in  $\mathcal{H}$  iff  $\mathbf{P}$  is hermitian.

This is the reason for demanding hermiticity of the projection operators in our treatment of the finite-dimensional case.

We now concentrate on the compact operators, and first look at hermitian compact operators. We need two lemmas:

**16.7.3. Lemma.** Let  $\mathbf{H} \in \mathcal{B}(\mathcal{H})$  be a bounded hermitian operator on the Hilbert space  $\mathcal{H}$ . Then  $\|\mathbf{H}\| = \max\{|\langle \mathbf{H}x|x \rangle| \mid \|x\| = 1\}$ .

*Proof.* Let  $M$  denote the positive number on the RHS. From the definition of the norm of an operator, we easily obtain  $|\langle \mathbf{H}x|x \rangle| \leq \|\mathbf{H}\| \|x\|^2 = \|\mathbf{H}\|$ , or  $M \leq \|\mathbf{H}\|$ . For the reverse inequality, see Problem 16.7. □

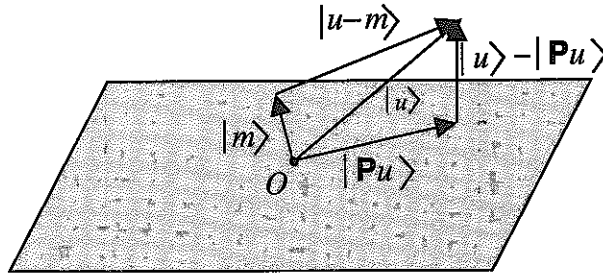
**16.7.4. Lemma.** Let  $\mathbf{K} \in \mathcal{B}(\mathcal{H})$  be a hermitian compact operator. Then there is an eigenvalue  $\lambda$  of  $\mathbf{K}$  such that  $|\lambda| = \|\mathbf{K}\|$ .

*Proof.* Let  $\{|x_n\rangle\}$  be a sequence of unit vectors such that

$$\|\mathbf{K}\| = \lim_n |\langle \mathbf{K}x_n|x_n \rangle|.$$

This is always possible, as the following argument shows. Let  $\epsilon$  be a small positive number. There must exist a unit vector  $|x_1\rangle \in \mathcal{H}$  such that

$$\|\mathbf{K}\| - \epsilon = |\langle \mathbf{K}x_1|x_1 \rangle|,$$



**Figure 16.2** The shaded area represents the subspace  $\mathcal{M}$  of the vector space. The convex subset  $E$  consists of all vectors connecting points of  $\mathcal{M}$  to the tip of  $|u\rangle$ . It is clear that there is a (unique) vector belonging to  $E$  whose length is minimum. The figure shows that this vector is orthogonal to  $\mathcal{M}$ .

because otherwise,  $\|\mathbf{K}\| - \epsilon$  would be greater than or equal to the norm of the operator (see Lemma 16.7.3). Similarly, there must exist another (different) unit vector  $|x_2\rangle \in \mathcal{H}$  such that  $\|\mathbf{K}\| - \epsilon/2 = |\langle \mathbf{K}x_2 | x_2 \rangle|$ . Continuing this way, we construct an infinite sequence of unit vectors  $\{|x_n\rangle\}$  with the property  $\|\mathbf{K}\| - \epsilon/n = |\langle \mathbf{K}x_n | x_n \rangle|$ . This construction clearly produces the desired sequence. Note that the argument holds for *any* hermitian bounded operator; compactness is not necessary.

Now define  $\alpha_n \equiv \langle \mathbf{K}x_n | x_n \rangle$  and let  $\alpha = \lim \alpha_n$ , so that  $|\alpha| = \|\mathbf{K}\|$ . Compactness of  $\mathbf{K}$  implies that  $\{\mathbf{K}x_n\}$  converges. Let  $|y\rangle \in \mathcal{H}$  be the limit of  $\{\mathbf{K}x_n\}$ . Then  $\|y\rangle = \lim \|\mathbf{K}x_n\| \leq \|\mathbf{K}\| \|x_n\| = \|\mathbf{K}\|$ . On the other hand,

$$0 \leq \|\mathbf{K}x_n - \alpha x_n\|^2 = \|\mathbf{K}x_n\|^2 - 2\alpha \langle \mathbf{K}x_n | x_n \rangle + |\alpha|^2.$$

Taking the limit and noting that  $\alpha_n$  and  $\alpha$  are real, we get

$$0 \leq \lim \|\mathbf{K}x_n\|^2 - 2\alpha \lim \langle \mathbf{K}x_n | x_n \rangle + |\alpha|^2 = \|y\rangle^2 - 2\alpha^2 + \alpha^2 \Rightarrow \|y\rangle^2 \geq \|\mathbf{K}\|^2.$$

It follows from these two inequalities that  $\|y\rangle = \|\mathbf{K}\|$  and that  $\lim |x_n\rangle = |y\rangle / \alpha$ . Furthermore,

$$(\mathbf{K} - \alpha \mathbf{1})(|y\rangle / \alpha) = (\mathbf{K} - \alpha \mathbf{1})(\lim |x_n\rangle) = \lim (\mathbf{K} - \alpha \mathbf{1}) |x_n\rangle = 0$$

Therefore,  $\alpha$  is an eigenvalue of  $\mathbf{K}$  with eigenvector  $|y\rangle / \alpha$ . □

Let us order all the eigenvalues of Theorem 16.6.9 in decreasing absolute value. Let  $\mathcal{M}_n$  denote the (finite-dimensional) eigenspace corresponding to eigenvalue  $\lambda_n$ , and  $\mathbf{P}_n$  the projection to  $\mathcal{M}_n$ . The eigenspaces are pairwise orthogonal and  $\mathbf{P}_n \mathbf{P}_m = 0$  for  $m \neq n$ . This follows in exact analogy with the finite-dimensional case.

First assume that  $\mathbf{K}$  has only finitely many eigenvalues,

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_r| > 0.$$

Let  $\mathcal{M} \equiv \mathcal{M}_1 \oplus \mathcal{M}_2 \oplus \cdots \oplus \mathcal{M}_r = \sum_{j=1}^r \oplus \mathcal{M}_j$ , and let  $\mathcal{M}_0$  be the orthogonal complement of  $\mathcal{M}$ . Since each eigenspace is invariant under  $\mathbf{K}$ , so is  $\mathcal{M}$ . Therefore, by Theorem 4.2.3—which holds for finite- as well as infinite-dimensional vector spaces—and the fact that  $\mathbf{K}$  is hermitian,  $\mathcal{M}_0$  is also invariant. Let  $\mathbf{K}_0$  be the restriction of  $\mathbf{K}$  to  $\mathcal{M}_0$ . By Lemma 16.7.4,  $\mathbf{K}_0$  has an eigenvalue  $\lambda$  such that  $|\lambda| = \|\mathbf{K}_0\|$ . If  $\lambda \neq 0$ , it must be one of the eigenvalues already accounted for, because any eigenvalue of  $\mathbf{K}_0$  is also an eigenvalue of  $\mathbf{K}$ . This is impossible, because  $\mathcal{M}_0$  is orthogonal to all the eigenspaces. So,  $\lambda = 0$ , or  $|\lambda| = \|\mathbf{K}_0\| = 0$ , or  $\mathbf{K}_0 = \mathbf{0}$ , i.e.,  $\mathbf{K}$  acts as the zero operator on  $\mathcal{M}_0$ .

Let  $\mathbf{P}_0$  be the orthogonal projection on  $\mathcal{M}_0$ . Then  $\mathcal{H} = \sum_{j=0}^r \oplus \mathcal{M}_j$ , and we have  $\mathbf{1} = \sum_{j=0}^r \mathbf{P}_j$ , and for an arbitrary  $|x\rangle \in \mathcal{H}$ , we have

$$\mathbf{K}|x\rangle = \mathbf{K} \left( \sum_{j=0}^r \mathbf{P}_j |x\rangle \right) = \sum_{j=0}^r \mathbf{K}(\mathbf{P}_j |x\rangle) = \sum_{j=1}^r \lambda_j (\mathbf{P}_j |x\rangle).$$

It follows that  $\mathbf{K} = \sum_{j=1}^r \lambda_j \mathbf{P}_j$ . Notice that the range of  $\mathbf{K}$  is  $\sum_{j=1}^r \oplus \mathcal{M}_j$ , which is finite-dimensional. Thus,  $\mathbf{K}$  has finite rank. Barring some technical details, which we shall not reproduce here, the case of a compact hermitian operator with infinitely many eigenvalues goes through in the same way (see [DeVi 90, pp. 179–180]):

spectral theorem for compact hermitian operators

**16.7.5. Theorem.** (spectral theorem: compact hermitian operators) *Let  $\mathbf{K}$  be a compact hermitian operator on a Hilbert space  $\mathcal{H}$ . Let  $\{\lambda_j\}_{j=1}^N$  be the distinct nonzero eigenvalues of  $\mathbf{K}$  arranged in decreasing order of absolute values. For each  $j$  let  $\mathcal{M}_j$  be the eigenspace of  $\mathbf{K}$  corresponding to eigenvalue  $\lambda_j$  and  $\mathbf{P}_j$  its projection operator with the property  $\mathbf{P}_i \mathbf{P}_j = \mathbf{0}$  for  $i \neq j$ . Then:*

1. *If  $N < \infty$ , then  $\mathbf{K}$  is an operator of finite rank,  $\mathbf{K} = \sum_{j=1}^N \lambda_j \mathbf{P}_j$ , and  $\mathcal{H} = \mathcal{M}_0 \oplus \mathcal{M}_1 \oplus \cdots \oplus \mathcal{M}_N$ , or  $\mathbf{1} = \sum_{j=0}^N \mathbf{P}_j$ , where  $\mathcal{M}_0$  is infinite-dimensional.*
2. *If  $N = \infty$ , then  $\lambda_j \rightarrow 0$  as  $j \rightarrow \infty$ ,  $\mathbf{K} = \sum_{j=1}^{\infty} \lambda_j \mathbf{P}_j$ , and  $\mathcal{H} = \mathcal{M}_0 \oplus \sum_{j=1}^{\infty} \oplus \mathcal{M}_j$ , or  $\mathbf{1} = \sum_{j=0}^{\infty} \mathbf{P}_j$ , where  $\mathcal{M}_0$  could be finite- or infinite-dimensional. Furthermore,*

$$\left\| \mathbf{K} - \sum_{j=1}^m \lambda_j \mathbf{P}_j \right\| = |\lambda_{m+1}| \quad \forall m,$$

*which shows that the infinite series above converges for an operator norm.*

The eigenspaces of a compact hermitian operator are orthogonal and, by (2) of Theorem 16.7.5, span the entire space. By the Gram–Schmidt process, one can select an orthonormal basis for each eigenspace. We therefore have the following corollary.

**16.7.6. Corollary.** *If  $\mathbf{K}$  is a compact hermitian operator on a Hilbert space  $\mathcal{H}$ , then the eigenvectors of  $\mathbf{K}$  constitute an orthonormal basis for  $\mathcal{H}$ .*

**16.7.7. Theorem.** Let  $\mathbf{K}$  be a compact hermitian operator on a Hilbert space  $\mathcal{H}$  and let  $\mathbf{K} = \sum_{j=1}^N \lambda_j \mathbf{P}_j$ , where  $N$  could be infinite. A bounded linear operator on  $\mathcal{H}$  commutes with  $\mathbf{K}$  if and only if it commutes with every  $\mathbf{P}_j$ .

*Proof.* The “if” part is straightforward. So assume that the bounded operator  $\mathbf{T}$  commutes with  $\mathbf{K}$ . For  $|x\rangle \in \mathcal{M}_j$ , we have  $(\mathbf{K} - \lambda_j)\mathbf{T}|x\rangle = \mathbf{T}(\mathbf{K} - \lambda_j)|x\rangle = 0$ . Similarly,  $(\mathbf{K} - \lambda_j)\mathbf{T}^\dagger|x\rangle = \mathbf{T}^\dagger(\mathbf{K} - \lambda_j)|x\rangle = 0$ , because  $0 = [\mathbf{T}, \mathbf{K}]^\dagger = [\mathbf{T}^\dagger, \mathbf{K}]$ . These equations show that both  $\mathbf{T}$  and  $\mathbf{T}^\dagger$  leave  $\mathcal{M}_j$  invariant. This means that  $\mathcal{M}_j$  reduces  $\mathbf{T}$ , and by Theorem 4.2.5,  $\mathbf{TP}_j = \mathbf{P}_j\mathbf{T}$ .  $\square$

Next we prove the spectral theorem for a normal operator. Recall that any operator  $\mathbf{T}$  can be written as  $\mathbf{T} = \mathbf{T}_r + i\mathbf{T}_i$  where  $\mathbf{T}_r = \frac{1}{2}(\mathbf{T} + \mathbf{T}^\dagger)$  and  $\mathbf{T}_i = \frac{1}{2i}(\mathbf{T} - \mathbf{T}^\dagger)$  are hermitian, and since both  $\mathbf{T}$  and  $\mathbf{T}^\dagger$  are compact,  $\mathbf{T}_r$  and  $\mathbf{T}_i$  are compact as well. For normal operators, we have the extra condition that  $[\mathbf{T}_r, \mathbf{T}_i] = [\mathbf{T}, \mathbf{T}^\dagger] = 0$ . Let  $\mathbf{T}_r = \sum_{j=1}^N \lambda_j \mathbf{P}_j$  and  $\mathbf{T}_i = \sum_{k=1}^N \mu_k \mathbf{Q}_k$  be the spectral decompositions of  $\mathbf{T}_r$  and  $\mathbf{T}_i$ . Using Theorem 16.7.7, it is straightforward to show that if  $[\mathbf{T}_r, \mathbf{T}_i] = 0$  then  $[\mathbf{P}_j, \mathbf{Q}_k] = 0$ . Now, since  $\mathcal{H} = \sum_{j=0}^N \oplus \mathcal{M}_j = \sum_{k=0}^N \oplus \mathcal{N}_k$ , where  $\mathcal{M}_j$  are the eigenspaces of  $\mathbf{T}_r$  and  $\mathcal{N}_k$  those of  $\mathbf{T}_i$ , we have, for any  $|x\rangle \in \mathcal{H}$ ,

$$\mathbf{T}_r|x\rangle = \left( \sum_{j=1}^N \lambda_j \mathbf{P}_j \right) \left( \sum_{k=0}^N \mathbf{Q}_k|x\rangle \right) = \sum_{j=1}^N \sum_{k=0}^N \lambda_j \mathbf{P}_j \mathbf{Q}_k|x\rangle.$$

Similarly,  $\mathbf{T}_i|x\rangle = \mathbf{T}_i(\sum_{j=0}^N \mathbf{P}_j|x\rangle) = \sum_{k=1}^N \sum_{j=0}^N \mu_k \mathbf{Q}_k \mathbf{P}_j|x\rangle$ . Combining these two relations and noting that  $\mathbf{Q}_k \mathbf{P}_j = \mathbf{P}_j \mathbf{Q}_k$  gives

$$\mathbf{T}|x\rangle = (\mathbf{T}_r + i\mathbf{T}_i)|x\rangle = \sum_{j=0}^N \sum_{k=0}^N (\lambda_j + i\mu_k) \mathbf{P}_j \mathbf{Q}_k|x\rangle.$$

The projection operators  $\mathbf{P}_j \mathbf{Q}_k$  project onto the intersection of  $\mathcal{M}_j$  and  $\mathcal{N}_k$ . Therefore,  $\mathcal{M}_j \cap \mathcal{N}_k$  are the eigenspaces of  $\mathbf{T}$ . Only those terms in the sum for which  $\mathcal{M}_j \cap \mathcal{N}_k \neq \emptyset$  contribute. As before, we can order the eigenvalues according to their absolute values.

spectral theorem for compact normal operators

**16.7.8. Theorem.** (spectral theorem: compact normal operators) Let  $\mathbf{T}$  be a compact normal operator on a Hilbert space  $\mathcal{H}$ . Let  $\{\lambda_j\}_{j=1}^N$  (where  $N$  can be  $\infty$ ) be the distinct nonzero eigenvalues of  $\mathbf{T}$  arranged in decreasing order of absolute values. For each  $n$  let  $\mathcal{M}_n$  be the eigenspace of  $\mathbf{T}$  corresponding to eigenvalue  $\lambda_n$  and  $\mathbf{P}_n$  its projection operator with the property  $\mathbf{P}_m \mathbf{P}_n = 0$  for  $m \neq n$ . Then:

1. If  $N < \infty$ , then  $\mathbf{T}$  is an operator of finite rank  $\mathbf{T} = \sum_{n=1}^N \lambda_n \mathbf{P}_n$ , and  $\mathcal{H} = \mathcal{M}_0 \oplus \mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_N$ , or  $\mathbf{1} = \sum_{j=0}^N \mathbf{P}_j$ , where  $\mathcal{M}_0$  is infinite-dimensional.
2. If  $N = \infty$ , then  $\lambda_n \rightarrow 0$  as  $n \rightarrow \infty$ ,  $\mathbf{T} = \sum_{n=1}^\infty \lambda_n \mathbf{P}_n$ , and  $\mathcal{H} = \mathcal{M}_0 \oplus \sum_{n=1}^\infty \mathcal{M}_n$ , or  $\mathbf{1} = \sum_{j=0}^\infty \mathbf{P}_j$ , where  $\mathcal{M}_0$  could be finite- or infinite-dimensional.

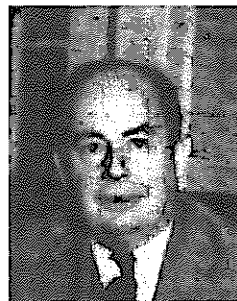
As in the case of a compact hermitian operator, by the Gram–Schmidt process, one can select an orthonormal basis for each eigenspace of a normal operator, in which case we have the following:

**16.7.9. Corollary.** *If  $\mathbf{T}$  is a compact normal operator on a Hilbert space  $\mathcal{H}$ , then the eigenvectors of  $\mathbf{T}$  constitute an orthonormal basis for  $\mathcal{H}$ .*

One can use Theorem 16.7.8 to write any function of a normal operator  $\mathbf{T}$  as an expansion in terms of the projection operators of  $\mathbf{T}$ . First we note that  $\mathbf{T}^k$  has  $\lambda_n^k$  as its expansion coefficients. Next, we add various powers of  $\mathbf{T}$  in the form of a polynomial and conclude that the expansion coefficients for a polynomial  $p(\mathbf{T})$  are  $p(\lambda_n)$ . Finally, for any function  $f(\mathbf{T})$  we have

$$f(\mathbf{T}) = \sum_{n=1}^{\infty} f(\lambda_n) \mathbf{P}_n. \quad (16.8)$$

**Johann (John) von Neumann**, (1903–1957), the eldest of three sons of Max von Neumann, a well-to-do Jewish banker, was privately educated until he entered the gymnasium in 1914. His unusual mathematical abilities soon came to the attention of his teachers, who pointed out to his father that teaching him conventional school mathematics would be a waste of time; he was therefore tutored in mathematics under the guidance of university professors, and by the age of nineteen he was already recognized as a professional mathematician and had published his first paper.



Von Neumann was Privatdozent at Berlin from 1927 to 1929 and at Hamburg in 1929–1930, then went to Princeton University for three years; in 1933 he was invited to join the newly opened Institute for Advanced Study, of which he was the youngest permanent member at that time. At the outbreak of World War II, von Neumann was called upon to participate in various scientific projects related to the war effort: In particular, from 1943 he was a consultant on the construction of the atomic bomb at Los Alamos. After the war he retained his membership on numerous government boards and committees, and in 1954 he became a member of the Atomic Energy Commission. His health began to fail in 1955, and he died of cancer two years later.

It is only in comparison with the greatest mathematical geniuses of history that von Neumann's scope in pure mathematics may appear somewhat restricted; it was far beyond the range of most of his contemporaries, and his extraordinary work in applied mathematics, in which he certainly equals Gauss, Cauchy, or Poincaré, more than compensates for its limitations. Von Neumann's work in pure mathematics was accomplished between 1925 and 1940, when he seemed to be advancing at a breathless speed on all fronts of logic and analysis at once, not to speak of mathematical physics. The dominant theme in von Neumann's work is by far his work on the *spectral theory of operators* in Hilbert spaces. For twenty years he was the undisputed master in this area, which contains what is now considered his most profound and most original creation, the theory of rings of operators.



The first papers (1927) in which Hilbert space theory appears are those on the *foundations of quantum mechanics*. These investigations later led von Neumann to a systematic study of unbounded hermitian operators.

Von Neumann's most famous work in theoretical physics is his *axiomatization of quantum mechanics*. When he began work in that field in 1927, the methods used by its founders were hard to formulate in precise mathematical terms: "Operators" on "functions" were handled without much consideration of their domain of definition or their topological properties, and it was blithely assumed that such "operators," when self-adjoint, could always be "diagonalized" (as in the finite dimensional case), at the expense of introducing Dirac delta functions as "eigenvectors." Von Neumann showed that mathematical rigor could be restored by taking as basic axioms the assumptions that the states of a physical system were points of a Hilbert space and that the measurable quantities were Hermitian (generally unbounded) operators densely defined in that space.

After 1927 von Neumann also devoted much effort to more specific problems of quantum mechanics, such as the problem of measurement and the foundation of quantum statistics and quantum thermodynamics, proving in particular an ergodic theorem for quantum systems. All this work was developed and expanded in *Mathematische Grundlagen der Quantenmechanik* (1932), in which he also discussed the much-debated question of "causality" versus "indeterminacy" and concluded that no introduction of "hidden parameters" could keep the basic structure of quantum theory and restore "causality."

Von Neumann's uncommon grasp of applied mathematics, treated as a whole without divorcing theory from experimental realization, was nowhere more apparent than in his work on computers. He became interested in *numerical computations* in connection with the need for quick estimates and approximate results that developed with the technology used for the war effort—particularly the complex problems of hydrodynamics—and the completely new problems presented by the harnessing of nuclear energy, for which no ready-made theoretical solutions were available. Von Neumann's extraordinary ability for rapid mental calculation was legendary. The story is told of a friend who brought him a simple kinematics problem. Two trains, a certain given distance apart, move toward each other at a given speed. A fly, initially on the windshield of one of the trains, flies back and forth between them, again at a known constant speed. When the trains collide, how far has the fly traveled? One way to solve the problem is to add up all the successively smaller distances in each individual flight. (The easy way is to multiply the fly's speed by the time elapsed until the crash.) After a few seconds of thought, von Neumann quickly gave the correct answer.

"That's strange," remarked his friend, "Most people try to sum the infinite series."

"What's strange about that?" von Neumann replied. "That's what I did."

In closing this section, let us remark that the paradigm of compact operators, namely the Hilbert–Schmidt operator, is such because it is defined on the *finite* rectangle  $[a, b] \times [a, b]$ . If this rectangle grows beyond limit, or equivalently, if the Hilbert space is  $\mathcal{L}^2(R_\infty)$ , where  $R_\infty$  is some infinite region of the real line, then the compactness property breaks down, as the following example illustrates.

**16.7.10. Example.** Consider the two kernels

$$K_1(x, t) = e^{-|x-t|} \quad \text{and} \quad K_2(x, t) = \sin xt$$

where the first one acts on  $\mathcal{L}^2(-\infty, \infty)$  and the second one on  $\mathcal{L}^2(0, \infty)$ . One can show (see Problem 16.8) that these two kernels have, respectively, the two eigenfunctions

$$e^{iat}, \quad \alpha \in \mathbb{R}, \quad \text{and} \quad \sqrt{\frac{\pi}{2}} e^{at} + \frac{t}{a^2 + t^2}, \quad a > 0,$$

corresponding to the two eigenvalues

$$\lambda = \frac{2}{1 + \alpha^2}, \quad \alpha \in \mathbb{R}, \quad \text{and} \quad \lambda = \sqrt{\frac{\pi}{2}}.$$

We see that in the first case, all real numbers between 0 and 2 are eigenvalues, rendering this set uncountable. In the second case, there are infinitely (in fact, uncountably) many eigenvectors (one for each  $a$ ) corresponding to the single eigenvalue  $\sqrt{\pi/2}$ . Note, however, that in the first case the eigenfunctions and in the second case the kernel have infinite norms. ■

## 16.8 Resolvents

The discussion of the preceding section showed that the spectrum of a normal compact operator is countable. Removing the compactness property in general will remove countability, as shown in Example 16.7.10. We have also seen that the right-shift operator, a bounded operator, has uncountably many points in its spectrum. We therefore expect that the sums in Theorem 16.7.8 should be replaced by integrals in the spectral decomposition theorem for (noncompact) bounded operators. We shall not discuss the spectral theorem for general operators. However, one special class of noncompact operators is essential for the treatment of Sturm–Liouville theory (to be studied in Chapters 18 and 19). For these operators, the concept of resolvent will be used, which we develop in this section. This concept also makes a connection between the countable (algebraic) and the uncountable (analytic) cases.

resolvent of an operator

**16.8.1. Definition.** Let  $\mathbf{T}$  be an operator and  $\lambda \in \rho(\mathbf{T})$ . The operator  $\mathbf{R}_\lambda(\mathbf{T}) \equiv (\mathbf{T} - \lambda\mathbf{1})^{-1}$  is called the *resolvent of  $\mathbf{T}$  at  $\lambda$* .

There are two important properties of the resolvent that are useful in analyzing the spectrum of operators. Let us assume that  $\lambda, \mu \in \rho(\mathbf{T})$ ,  $\lambda \neq \mu$ , and take the difference between their resolvents. Problem 16.9 shows how to obtain the following relation:

$$\mathbf{R}_\lambda(\mathbf{T}) - \mathbf{R}_\mu(\mathbf{T}) = (\lambda - \mu)\mathbf{R}_\lambda(\mathbf{T})\mathbf{R}_\mu(\mathbf{T}). \tag{16.9}$$

To obtain the second property of the resolvent, we formally (and indefinitely) differentiate  $\mathbf{R}_\lambda(\mathbf{T})$  with respect to  $\lambda$  and evaluate the result at  $\lambda = \mu$ :

$$\frac{d}{d\lambda} \mathbf{R}_\lambda(\mathbf{T}) = \frac{d}{d\lambda} [(\mathbf{T} - \lambda\mathbf{1})^{-1}] = (\mathbf{T} - \lambda\mathbf{1})^{-2} = \mathbf{R}_\lambda^2(\mathbf{T}).$$

Differentiating both sides of this equation, we get  $2R_\lambda^3(T)$ , and in general,

$$\frac{d^n}{d\lambda^n} R_\lambda(T) = n! R_\lambda^{n+1}(T) \Rightarrow \left. \frac{d^n}{d\lambda^n} R_\lambda(T) \right|_{\lambda=\mu} = n! R_\mu^{n+1}(T).$$

Assuming that the Taylor series expansion exists, we may write

$$R_\lambda(T) = \sum_{n=0}^{\infty} \frac{(\lambda - \mu)^n}{n!} \left. \frac{d^n}{d\lambda^n} R_\lambda(T) \right|_{\lambda=\mu} = \sum_{n=0}^{\infty} (\lambda - \mu)^n R_\mu^{n+1}(T), \tag{16.10}$$

which is the second property of the resolvent.

We now look into the spectral decomposition from an analytical viewpoint. For convenience, we concentrate on the finite-dimensional case and let  $A$  be an arbitrary (not necessarily hermitian)  $N \times N$  matrix. Let  $\lambda$  be a complex number that is larger (in absolute value) than any of the eigenvalues of  $A$ . Since all operators on finite-dimensional vector spaces are compact, Lemma 16.7.4 assures us that  $|\lambda| > \|T\|$ , and it is then possible to expand  $R_\lambda(T)$  in a *convergent* power series as follows:

$$R_\lambda(A) = (A - \lambda 1)^{-1} = -\frac{1}{\lambda} \sum_{n=0}^{\infty} \left(\frac{A}{\lambda}\right)^n. \tag{16.11}$$

This is the Laurent expansion of  $R_\lambda(A)$ . We can immediately read off the residue of  $R_\lambda(A)$  (the coefficient of  $1/\lambda$ ):

$$\text{Res}[R_\lambda(A)] = -1 \Rightarrow -\frac{1}{2\pi i} \oint_{\Gamma} R_\lambda(A) d\lambda = 1,$$

where  $\Gamma$  is a circle with its center at the origin and a radius large enough to encompass all the eigenvalues of  $A$  [see Figure 16.3(a)]. A similar argument shows that

$$-\frac{1}{2\pi i} \oint_{\Gamma} \lambda R_\lambda(A) d\lambda = A,$$

and in general,

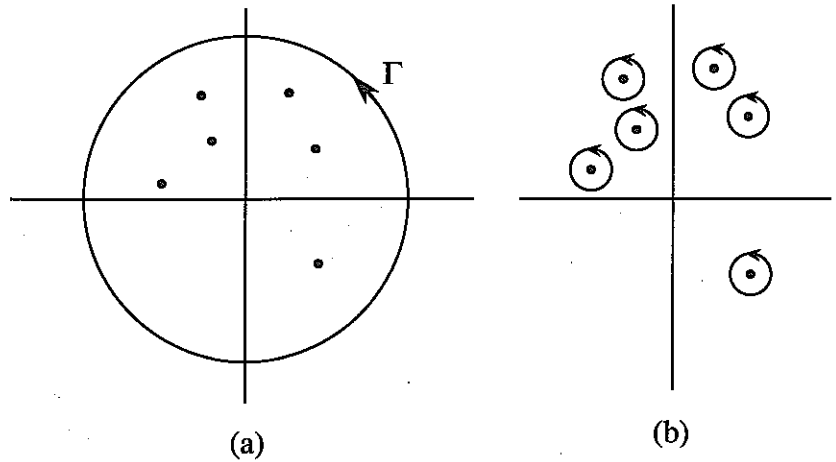
$$-\frac{1}{2\pi i} \oint_{\Gamma} \lambda^n R_\lambda(A) d\lambda = A^n \quad \text{for } n = 0, 1, \dots$$

Using this and assuming that we can expand the function  $f(A)$  in a power series, we get

$$-\frac{1}{2\pi i} \oint_{\Gamma} f(\lambda) R_\lambda(A) d\lambda = f(A). \tag{16.12}$$

Writing this equation in the form

$$\frac{1}{2\pi i} \oint_{\Gamma} \frac{f(\lambda)}{\lambda 1 - A} d\lambda = f(A)$$



**Figure 16.3** (a) The large circle encompassing all eigenvalues. (b) the deformed contour consisting of small circles orbiting the eigenvalues.

makes it recognizable as the generalization of the Cauchy integral formula to operator-valued functions. To use any of the above integral formulas, we must know the analytic behavior of  $R_\lambda(A)$ . From the formula of the inverse of a matrix given in Chapter 3, we have

$$[R_\lambda(A)]_{jk} = [(A - \lambda I)^{-1}]_{jk} = \frac{C_{jk}(\lambda)}{\det(A - \lambda I)} = \frac{C_{jk}(\lambda)}{p(\lambda)},$$

where  $C_{jk}(\lambda)$  is the cofactor of the  $ij$ th element of the matrix  $A - \lambda I$  and  $p(\lambda)$  is the characteristic polynomial of  $A$ . Clearly,  $C_{jk}(\lambda)$  is also a polynomial. Thus,  $[R_\lambda(A)]_{jk}$  is a rational function of  $\lambda$ . It follows that  $R_\lambda(A)$  has only poles as singularities (see Example 10.2.2). The poles are simply the zeros of the denominator, i.e., the eigenvalues of  $A$ . We can deform the contour  $\Gamma$  in such a way that it consists of small circles  $\gamma_j$  that encircle the isolated eigenvalues  $\lambda_j$  [see Figure 16.3(b)]. Then, with  $f(A) = 1$ , Equation (16.12) yields

$$1 = -\frac{1}{2\pi i} \sum_{j=1}^r \oint_{\gamma_j} R_\lambda(A) d\lambda = \sum_{j=1}^r P_j, \quad P_j \equiv -\frac{1}{2\pi i} \oint_{\gamma_j} R_\lambda(A) d\lambda. \quad (16.13)$$

It can be shown (see Example 16.8.2 below) that  $\{P_j\}$  is a set of orthogonal projection operators. Thus, Equation (16.13) is a resolution of identity, as specified in the spectral decomposition theorem in Chapter 4.

**16.8.2. Example.** We want to show that the  $P_j$  are projection operators. First let  $i = j$ . Then<sup>12</sup>

$$P_j^2 = \left(-\frac{1}{2\pi i}\right)^2 \oint_{\gamma_j} R_\lambda(A) d\lambda \oint_{\gamma_j} R_\mu(A) d\mu.$$

Note that  $\lambda$  need not be equal to  $\mu$ . In fact, we are free to choose  $|\lambda - \lambda_j| > |\mu - \lambda_j|$ , i.e., let the circle corresponding to  $\lambda$  integration be outside that of  $\mu$  integration.<sup>13</sup> We can then rewrite the above double integral as

$$\begin{aligned} P_j^2 &= \left(-\frac{1}{2\pi i}\right)^2 \oint_{\gamma_j^{(\lambda)}} \oint_{\gamma_j^{(\mu)}} R_\lambda(A) R_\mu(A) d\lambda d\mu \\ &= \left(-\frac{1}{2\pi i}\right)^2 \oint_{\gamma_j^{(\lambda)}} \oint_{\gamma_j^{(\mu)}} \left[ \frac{R_\lambda(A)}{\lambda - \mu} - \frac{R_\mu(A)}{\lambda - \mu} \right] d\lambda d\mu \\ &= \left(-\frac{1}{2\pi i}\right)^2 \left\{ \oint_{\gamma_j^{(\lambda)}} R_\lambda(A) d\lambda \oint_{\gamma_j^{(\mu)}} \frac{d\mu}{\lambda - \mu} - \oint_{\gamma_j^{(\mu)}} R_\mu(A) d\mu \oint_{\gamma_j^{(\lambda)}} \frac{d\lambda}{\lambda - \mu} \right\}, \end{aligned}$$

where we used Equation (16.9) to go to the second line. Now note that

$$\oint_{\gamma_j^{(\mu)}} \frac{d\mu}{\lambda - \mu} = 0 \quad \text{and} \quad \oint_{\gamma_j^{(\lambda)}} \frac{d\lambda}{\lambda - \mu} = 2\pi i$$

because  $\lambda$  lies outside  $\gamma_j^{(\mu)}$  and  $\mu$  lies inside  $\gamma_j^{(\lambda)}$ . Hence,

$$P_j^2 = \left(-\frac{1}{2\pi i}\right)^2 \left\{ 0 - 2\pi i \oint_{\gamma_j^{(\mu)}} R_\mu(A) d\mu \right\} = -\frac{1}{2\pi i} \oint_{\gamma_j^{(\mu)}} R_\mu(A) d\mu = P_j.$$

The remaining part, namely  $P_j P_k = 0$  for  $k \neq j$ , can be done similarly (see Problem 16.10). ■

Now we let  $f(A) = A$  in Equation (16.12), deform the contour as above, and write

$$\begin{aligned} A &= -\frac{1}{2\pi i} \sum_{j=1}^r \oint_{\gamma_j} \lambda R_\lambda(A) d\lambda \\ &= -\frac{1}{2\pi i} \sum_{j=1}^r \left[ \lambda_j \oint_{\gamma_j} R_\lambda(A) d\lambda + \oint_{\gamma_j} (\lambda - \lambda_j) R_\lambda(A) d\lambda \right] \\ &\equiv \sum_{j=1}^r (\lambda_j P_j + D_j), \quad D_j \equiv \oint_{\gamma_j} (\lambda - \lambda_j) R_\lambda(A) d\lambda. \end{aligned} \tag{16.14}$$

<sup>12</sup>We have not discussed multiple integrals of complex functions. A rigorous study of such integrals involves the theory of functions of several complex variables—a subject we have to avoid due to lack of space. However, in the simple case at hand, the theory of real multiple integrals is an honest guide.

<sup>13</sup>This is possible because the poles are isolated.

It can be shown (see Problem 16.11) that

$$D_j^n = \oint_{\gamma_j} (\lambda - \lambda_j)^n R_\lambda(A) d\lambda.$$

In particular, since  $R_\lambda(A)$  has only poles as singularities, there exists a positive integer  $m$  such that  $D_j^m = 0$ . We have not yet made any assumptions about  $A$ . If we assume that  $A$  is hermitian, for example, then  $R_\lambda(A)$  will have simple poles (see Problem 16.12). It follows that  $(\lambda - \lambda_j)R_\lambda(A)$  will be analytic at  $\lambda_j$  for all  $j = 1, 2, \dots, r$ , and  $D_j = 0$  in Equation (16.14). We thus have

$$A = \sum_{j=1}^r \lambda_j P_j,$$

which is the spectral decomposition discussed in Chapter 4. Problem 16.13 shows that the  $P_j$  are hermitian.

**16.8.3. Example.** The most general  $2 \times 2$  hermitian matrix is of the form

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12}^* & a_{22} \end{pmatrix},$$

where  $a_{11}$  and  $a_{22}$  are real numbers. Thus,

$$\det(A - \lambda I) = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - |a_{12}|^2$$

which has roots

$$\begin{aligned} \lambda_1 &= \frac{1}{2}[a_{11} + a_{22} - \sqrt{(a_{11} - a_{22})^2 + 4|a_{12}|^2}], \\ \lambda_2 &= \frac{1}{2}[a_{11} + a_{22} + \sqrt{(a_{11} - a_{22})^2 + 4|a_{12}|^2}]. \end{aligned}$$

The inverse of  $A - \lambda I$  can immediately be written:

$$\begin{aligned} R_\lambda(A) &= (A - \lambda I)^{-1} = \frac{1}{\det(A - \lambda I)} \begin{pmatrix} a_{22} - \lambda & -a_{12} \\ -a_{12}^* & a_{11} - \lambda \end{pmatrix} \\ &= \frac{1}{(\lambda - \lambda_1)(\lambda - \lambda_2)} \begin{pmatrix} a_{22} - \lambda & -a_{12} \\ -a_{12}^* & a_{11} - \lambda \end{pmatrix}. \end{aligned}$$

We want to verify that  $R_\lambda(A)$  has only simple poles. Two cases arise:

1. If  $\lambda_1 \neq \lambda_2$ , then it is clear that  $R_\lambda(A)$  has simple poles.
2. If  $\lambda_1 = \lambda_2$ , it appears that  $R_\lambda(A)$  may have a pole of order 2. However, note that if  $\lambda_1 = \lambda_2$ , then the square roots in the above equations must vanish. This happens iff  $a_{11} = a_{22} \equiv a$  and  $a_{12} = 0$ . It then follows that  $\lambda_1 = \lambda_2 \equiv a$ , and

$$R_\lambda(A) = \frac{1}{(\lambda - a)^2} \begin{pmatrix} a - \lambda & 0 \\ 0 & a - \lambda \end{pmatrix}.$$

This clearly shows that  $R_\lambda(A)$  has only simple poles in this case. ■

Jordan canonical  
form

If  $A$  is not hermitian,  $D_j \neq 0$ ; however,  $D_j$  is nevertheless *nilpotent*. That is,  $D_j^m = 0$  for some positive integer  $m$ . This property and Equation (16.14) can be used to show that  $A$  can be cast into a *Jordan canonical form* via a similarity transformation. That is, there exists an  $N \times N$  matrix  $S$  such that

$$SAS^{-1} = J = \begin{pmatrix} J_1 & 0 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & J_m \end{pmatrix}$$

where  $J_k$  is a matrix of the form

$$J_k = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda & 1 \end{pmatrix}$$

in which  $\lambda$  is one of the eigenvalues of  $A$ . Different  $J_k$  may contain the same eigenvalues of  $A$ . For a discussion of the Jordan canonical form of a matrix, see [Birk 77], [Denn 67], or [Halm 58].

## 16.9 Problems

**16.1.** Suppose that  $S$  is a bounded operator,  $T$  an invertible operator, and that

$$\|T - S\| < \frac{1}{\|T^{-1}\|}.$$

Show that  $S$  is invertible. Hint: Show that  $T^{-1}S$  is invertible. Thus, an operator that is “sufficiently close” to an invertible operator is invertible.

**16.2.** Let  $\mathcal{V}$  and  $\mathcal{W}$  be finite-dimensional vector spaces. Show that  $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$  is necessarily bounded.

**16.3.** Let  $\mathcal{H}$  be a Hilbert space, and  $T \in \mathcal{L}(\mathcal{H})$  an isometry, i.e., a linear operator that does not change the norm of any vector. Show that  $\|T\| = 1$ .

**16.4.** Show that (a) the unit operator is not compact, and that (b) the inverse of a compact operator cannot be bounded. Hint: For (b) use the results of Example 16.5.3.

**16.5.** Prove Corollary 16.6.7. Hint: Let  $|x\rangle \in \mathcal{N}_\lambda^{(q+1)}$  and write it as  $|x\rangle = |n\rangle + |r\rangle$  with  $|n\rangle \in \mathcal{N}_\lambda^{(q)}$  and  $|r\rangle \in \mathcal{R}_\lambda^{(q)}$ . Apply  $(K - \lambda I)^{q+1}$  to  $|r\rangle$ , and invoke part (3) of Theorem 16.6.6 to show that  $|r\rangle \in \mathcal{N}_\lambda^{(q)}$ . Conclude that  $|r\rangle = 0$ ,  $\mathcal{N}_\lambda^{(q+1)} = \mathcal{N}_\lambda^{(q)}$ ,

and  $q \geq p$ . To establish the reverse inequality, apply  $(\mathbf{K} - \lambda \mathbf{1})^p$  to both sides of the direct sum of part (1) of Theorem 16.6.6, and notice that the LHS is  $\mathcal{R}_\lambda^{(p)}$ , the second term of the RHS is zero, and the first term is  $\mathcal{R}_\lambda^{(q+p)}$ . Now conclude that  $p \geq q$ .

**16.6.** Let  $|u\rangle \in \mathcal{H}$  and let  $\mathcal{M}$  be a subspace of  $\mathcal{H}$ . Show that the subset  $E = |u\rangle - \mathcal{M}$  is convex. Show that  $E$  is not necessarily a subspace of  $\mathcal{H}$ .

**16.7.** Show that for any hermitian operator  $\mathbf{H}$ , we have

$$4 \langle \mathbf{H}x | y \rangle = \langle \mathbf{H}(x+y) | x+y \rangle - \langle \mathbf{H}(x-y) | x-y \rangle \\ + i[\langle \mathbf{H}(x+iy) | x+iy \rangle - \langle \mathbf{H}(x-iy) | x-iy \rangle].$$

Now let  $|x\rangle = \lambda |z\rangle$  and  $|y\rangle = |\mathbf{H}z\rangle / \lambda$ , where  $\lambda = (\|\mathbf{H}z\| / \|z\|)^{1/2}$ , and show that

$$\|\mathbf{H}z\|^2 = \langle \mathbf{H}x | y \rangle \leq M \|z\| \|\mathbf{H}z\|,$$

where  $M = \max\{|\langle \mathbf{H}z | z \rangle| / \|z\|^2\}$ . Now conclude that  $\|\mathbf{H}\| \leq M$ .

**16.8.** Show that the two kernels  $K_1(x, t) = e^{-|x-t|}$  and  $K_2(x, t) = \sin xt$ , where the first one acts on  $\mathcal{L}^2(-\infty, \infty)$  and the second one on  $\mathcal{L}^2(0, \infty)$ , have the two eigenfunctions

$$e^{iat}, \quad \alpha \in \mathbb{R}, \quad \text{and} \quad \sqrt{\frac{\pi}{2}} e^{at} + \frac{t}{a^2 + t^2}, \quad a > 0,$$

respectively, corresponding to the two eigenvalues

$$\lambda = \frac{2}{1 + \alpha^2}, \quad \alpha \in \mathbb{R}, \quad \text{and} \quad \lambda = \sqrt{\frac{\pi}{2}}.$$

**16.9.** Derive Equation (16.9). Hint: Multiply  $\mathbf{R}_\lambda(\mathbf{T})$  by  $\mathbf{1} = \mathbf{R}_\mu(\mathbf{T})(\mathbf{T} - \mu \mathbf{1})$  and  $\mathbf{R}_\mu(\mathbf{T})$  by  $\mathbf{1} = \mathbf{R}_\lambda(\mathbf{T})(\mathbf{T} - \lambda \mathbf{1})$ .

**16.10.** Finish Example 16.8.2 by showing that  $\mathbf{P}_j \mathbf{P}_k = 0$  for  $k \neq j$ .

**16.11.** Show that  $\mathbf{D}_j^n = \oint_{\gamma_j} (\lambda - \lambda_j)^n \mathbf{R}_\lambda(\mathbf{A}) d\lambda$ . Hint: Use mathematical induction and the technique used in Example 16.8.2.

**16.12.** (a) Take the inner product of  $|u\rangle = (\mathbf{A} - \lambda \mathbf{1}) |v\rangle$  with  $|v\rangle$  and show that for a hermitian  $\mathbf{A}$ ,  $\text{Im} \langle v | u \rangle = -(\text{Im} \lambda) \|v\|^2$ . Now use the Schwarz inequality to obtain

$$\|v\| \leq \frac{\|u\|}{|\text{Im} \lambda|} \Rightarrow \|\mathbf{R}_\lambda(\mathbf{A}) |u\rangle\| \leq \frac{\|u\|}{|\text{Im} \lambda|}.$$

(b) Use this result to show that

$$\|(\lambda - \lambda_j) \mathbf{R}_\lambda(\mathbf{A}) |u\rangle\| \leq \left(1 + \left| \frac{\text{Re}(\lambda - \lambda_j)}{\text{Im}(\lambda - \lambda_j)} \right| \right) \|u\| = (1 + |\cot \theta|) \|u\|,$$



where  $\theta$  is the angle that  $\lambda - \lambda_j$  makes with the real axis and  $\lambda$  is chosen to have an imaginary part. From this result conclude that  $\mathbf{R}_\lambda(\mathbf{A})$  has a simple pole when  $\mathbf{A}$  is hermitian.

**16.13.** (a) Show that when  $\mathbf{A}$  is hermitian,  $[\mathbf{R}_\lambda(\mathbf{A})]^\dagger = \mathbf{R}_{\lambda^*}(\mathbf{A})$ .

(b) Write  $\lambda - \lambda_j = r_j e^{i\theta}$  in the definition of  $\mathbf{P}_j$  in Equation (16.13). Take the hermitian conjugate of both sides and use (a) to show that  $\mathbf{P}_j$  is hermitian. Hint: You will have to change the variable of integration a number of times.

## Additional Readings

1. DeVito, C. *Functional Analysis and Linear Operator Theory*, Addison-Wesley, 1990. Our treatment of compact operators follows this reference's discussion.
2. Glimm, J. and Jaffe, A. *Quantum Physics*, 2nd ed., Springer-Verlag, 1987. One of the most mathematical treatments of the subject, and therefore a good introduction to operator theory (see the appendix to Part I).
3. Reed, M. and Simon, B. *Fourier Analysis, Self-Adjointness*, Academic Press, 1980.
4. Richtmyer, R. *Principles of Advanced Mathematical Physics*, Springer-Verlag, 1978. Discusses resolvents in detail.
5. Zeidler, E. *Applied Functional Analysis*, Springer-Verlag, 1995.

# Integral Equations

The beginning of Chapter 16 showed that to solve a vector-operator equation one transforms it into an equation involving a sum over a discrete index [the matrix equation of Equation (16.1)], or an equation involving an integral over a continuous index [Equation (16.2)]. The latter is called an **integral equation**, which we shall investigate here using the machinery of Chapter 16.

## 17.1 Classification

Volterra and  
Fredholm equations  
of first and second  
kind

Integral equations can be divided into two major groups. Those that have a variable limit of integration are called **Volterra equations**; those that have constant limits of integration are called **Fredholm equations**. If the unknown function appears only inside the integral, the integral equation is said to be **of the first kind**. Integral equations having the unknown function outside the integral as well as inside are said to be **of the second kind**. The four kinds of equations can be written as follows.

$$\int_a^x K(x, t)u(t) dt = v(x), \quad \text{Volterra equation of the 1st kind,}$$

$$\int_a^b K(x, t)u(t) dt = v(x), \quad \text{Fredholm equation of the 1st kind,}$$

$$u(x) = v(x) + \int_a^x K(x, t)u(t) dt, \quad \text{Volterra equation of the 2nd kind,}$$

$$u(x) = v(x) + \int_a^b K(x, t)u(t) dt, \quad \text{Fredholm equation of the 2nd kind.}$$

In all these equations,  $K(x, t)$  is called the **kernel** of the integral equation.

In the theory of integral equations of the second kind, one usually multiplies the integral by a nonzero complex number  $\lambda$ . Thus, the Fredholm equation of the second kind becomes

$$u(x) = v(x) + \lambda \int_a^b K(x, t)u(t) dt, \tag{17.1}$$

and for the Volterra equation of the second kind one obtains

$$u(x) = v(x) + \lambda \int_a^x K(x, t)u(t) dt. \tag{17.2}$$

characteristic value  
of an integral  
equation

A  $\lambda$  that satisfies (17.2) with  $v(x) = 0$  is called a **characteristic value** of the integral equation. In the abstract operator language both equations are written as

$$|u\rangle = |v\rangle + \lambda \mathbf{K}|u\rangle \Rightarrow (\mathbf{K} - \lambda^{-1})|u\rangle = -\lambda^{-1}|v\rangle. \tag{17.3}$$

Thus  $\lambda$  is a characteristic value for (17.1) if and only if  $\lambda^{-1}$  is an eigenvalue of  $\mathbf{K}$ . Recall that when the interval of integration  $(a, b)$  is finite,  $K(x, t)$  is called a *Hilbert–Schmidt* kernel. Example 16.5.10 showed that  $\mathbf{K}$  is a compact operator, and by Theorem 16.6.9, the eigenvalues of  $\mathbf{K}$  either form a finite set or a sequence that converges to zero.

**17.1.1. Theorem.** *The characteristic values of a Fredholm equation of the second kind either form a finite set or a sequence of complex numbers increasing beyond limit in absolute value.*

Our main task in this chapter is to study methods of solving integral equations of the second kind. We treat the Volterra equation first because it is easier to solve. Let us introduce the notation

$$K[u](x) \equiv \int_a^x K(x, t)u(t) dt \quad \text{and} \quad K^n[u](x) = K[K^{n-1}[u]](x) \tag{17.4}$$

whereby  $K[u]$  denotes a function whose value at  $x$  is given by the integral on the RHS of the first equation in (17.4). One can show with little difficulty that the associated operator  $\mathbf{K}$  is compact. Let  $M = \max\{|K(x, t)| \mid a \leq t \leq x \leq b\}$  and note that

$$|\lambda K[u](x)| = \left| \lambda \int_a^x K(x, t)u(t) dt \right| \leq |\lambda| |M| \|u\|_\infty (x - a),$$

where  $\|u\|_\infty \equiv \max\{|u(x)| \mid x \in (a, b)\}$ .

Using mathematical induction, one can show that (see Problem 17.1)

$$|(\lambda K)^n[u](x)| \leq |\lambda|^n |M|^n \|u\|_\infty \frac{(x - a)^n}{n!}. \tag{17.5}$$

Since  $b \geq x$ , we can replace  $x$  with  $b$  and still satisfy the inequality. Then the inequality of Equation (17.5) will hold for all  $x$ , and we can write the equation as an operator norm inequality:  $\|(\lambda \mathbf{K})^n\| \leq |\lambda|^n |M|^n \|u\|_\infty (b-a)^n / n!$ . Therefore,

$$\left\| \sum_{n=0}^{\infty} (\lambda \mathbf{K})^n \right\| \leq \sum_{n=0}^{\infty} \|(\lambda \mathbf{K})^n\| \leq \sum_{n=0}^{\infty} \frac{|\lambda|^n |M|^n (b-a)^n}{n!} = e^{M|\lambda|(b-a)},$$

and the series  $\sum_{n=0}^{\infty} (\lambda \mathbf{K})^n$  converges for all  $\lambda$ . In fact, a direct calculation shows that the series converges to the inverse of  $\mathbf{1} - \lambda \mathbf{K}$ . Thus, the latter is invertible and the spectrum of  $\mathbf{K}$  has no nonzero points. We have just shown the following.

**17.1.2. Theorem.** *The Volterra equation of the second kind has no nonzero characteristic value. In particular, the operator  $\mathbf{1} - \lambda \mathbf{K}$  is invertible, and the Volterra equation of the second kind always has a unique solution given by the convergent infinite series  $u(x) = \sum_{j=0}^{\infty} \lambda^j \int_a^x K^j(x, t)v(t) dt$  where  $K^j(x, t)$  is defined inductively in Equation (17.4).*

Volterra equation of the second kind has a unique solution and no nonzero characteristic value

Vito Volterra (1860–1940) was only 11 when he became interested in mathematics while reading Legendre's *Geometry*. At the age of 13 he began to study the three body problem and made some progress.

His family were extremely poor (his father had died when Vito was two years old) but after attending lectures at Florence he was able to proceed to Pisa in 1878. At Pisa he studied under Betti, graduating as a doctor of physics in 1882. His thesis on hydrodynamics included some results of Stokes, discovered later but independently by Volterra.

He became Professor of Mechanics at Pisa in 1883, and upon Betti's death, he occupied the chair of mathematical physics. After spending some time at Turin as the chair of mechanics, he was awarded the chair of mathematical physics at the University of Rome in 1900.

Volterra conceived the idea of a theory of functions that depend on a continuous set of values of another function in 1883. Hadamard was later to introduce the word "functional," which replaced Volterra's original terminology. In 1890 Volterra used his functional calculus to show that the theory of Hamilton and Jacobi for the integration of the differential equations of dynamics could be extended to other problems of mathematical physics.

His most famous work was done on **integral equations**. He began this study in 1884, and in 1896 he published several papers on what is now called the Volterra integral equation. He continued to study functional analysis applications to integral equations producing a large number of papers on composition and permutable functions.

During the First World War Volterra joined the Air Force. He made many journeys to France and England to promote scientific collaboration. After the war he returned to the University of Rome, and his interests moved to mathematical biology. He studied the Verhulst equation and the logistic curve. He also wrote on predator-prey equations.

In 1922 Fascism seized Italy, and Volterra fought against it in the Italian Parliament. However, by 1930 the Parliament was abolished, and when Volterra refused to take an oath



of allegiance to the Fascist government in 1931, he was forced to leave the University of Rome. From the following year he lived mostly abroad, mainly in Paris, but also in Spain and other countries.

**17.1.3. Example.** Differential equations can be transformed into integral equations. For instance, consider the SOLDE

$$\frac{d^2u}{dx^2} + p_1(x)\frac{du}{dx} + p_0(x)u = r(x), \quad u(a) = c_1, \quad u'(a) = c_2.$$

By integrating the DE once, we obtain

$$\frac{du}{dx} = -\int_a^x p_1(t)u'(t) dt - \int_a^x p_0(t)u(t) dt + \int_a^x r(t) dt + c_2.$$

Integrating the first integral by parts gives

$$u'(x) = -p_1(x)u(x) + \underbrace{\int_a^x [p_1'(t) - p_0(t)]u(t) dt}_{\equiv f(x)} + \underbrace{\int_a^x r(t) dt}_{\equiv g(x)} + p_1(a)c_1 + c_2.$$

Integrating once more yields

$$\begin{aligned} u(x) &= -\int_a^x p_1(t)u(t) dt + \int_a^x f(s) ds + \int_a^x g(s) ds + (x-a)[p_1(a)c_1 + c_2] \\ &= -\int_a^x p_1(t)u(t) dt + \int_a^x ds \int_a^s [p_1'(t) - p_0(t)]u(t) dt \\ &\quad + \int_a^x ds \int_a^s r(t) dt + (x-a)[p_1(a)c_1 + c_2] + c_1 \\ &= \int_a^x \{(x-t)[p_1'(t) - p_0(t)] - p_1(t)\} u(t) dt \\ &\quad + \int_a^x (x-t)r(t) dt + (x-a)[p_1(a)c_1 + c_2] + c_1, \end{aligned} \tag{17.6}$$

where we have used the formula

$$\int_a^x ds \int_a^s f(t) dt = \int_a^x (x-t)f(t) dt,$$

which the reader may verify by interchanging the order of integration on the LHS. Equation (17.6) is a Volterra equation of the second kind with kernel

$$K(x, t) \equiv (x-t)[p_1'(t) - p_0(t)] - p_1(t)$$

and  $v(x) \equiv \int_a^x (x-t)r(t) dt + (x-a)[p_1(a)c_1 + c_2] + c_1$ . ■

Neumann series  
solution

We now outline a systematic approach to obtaining the infinite series of Theorem 17.1.2, which also works for the Fredholm equation of the second kind as we shall see in the next section. In the latter case, the series is guaranteed to converge

only if  $|\lambda||\mathbf{K}| < 1$ . This approach has the advantage that in each successive step, we obtain a better approximation to the solution. Writing the equation as

$$|u\rangle = |v\rangle + \lambda\mathbf{K}|u\rangle, \quad (17.7)$$

we can interpret it as follows. The difference between  $|u\rangle$  and  $|v\rangle$  is  $\lambda\mathbf{K}|u\rangle$ . If  $\lambda\mathbf{K}$  were absent, the two vectors  $|u\rangle$  and  $|v\rangle$  would be equal. The effect of  $\lambda\mathbf{K}$  is to change  $|u\rangle$  in such a way that when the result is added to  $|v\rangle$ , it gives  $|u\rangle$ . As our initial approximation, therefore, we take  $|u\rangle$  to be equal to  $|v\rangle$  and write  $|u_0\rangle = |v\rangle$ , where the index reminds us of the order (in this case zeroth, because  $\lambda\mathbf{K} = 0$ ) of the approximation. To find a better approximation, we always substitute the latest approximation for  $|u\rangle$  in the RHS of Equation (17.7). At this stage, we have  $|u_1\rangle = |v\rangle + \lambda\mathbf{K}|u_0\rangle = |v\rangle + \lambda\mathbf{K}|v\rangle$ . Still a better approximation is achieved if we substitute this expression in (17.7):

$$|u_2\rangle = |v\rangle + \lambda\mathbf{K}|u_1\rangle = |v\rangle + \lambda\mathbf{K}(|v\rangle + \lambda\mathbf{K}|v\rangle) = |v\rangle + \lambda\mathbf{K}|v\rangle + \lambda^2\mathbf{K}^2|v\rangle.$$

The procedure is now clear. Once  $|u_n\rangle$ , the  $n$ th approximation, is obtained, we can get  $|u_{n+1}\rangle$  by substituting in the RHS of (17.7).

Before continuing, let us write the above equations in integral form. In what follows, we shall concentrate on the Fredholm equation. To obtain the result for the Volterra equation, one simply replaces  $b$ , the upper limit of integration, with  $x$ . The first approximation can be obtained by substituting  $v(t)$  for  $u(t)$  on the RHS of Equation (17.1). This yields

$$u_1(x) = v(x) + \lambda \int_a^b K(x, t)v(t) dt.$$

Substituting this back in Equation (17.1) gives

$$\begin{aligned} u_2(x) &= v(x) + \lambda \int_a^b ds K(x, s)u_1(s) \\ &= v(x) + \lambda \int_a^b ds K(x, s)v(s) + \lambda^2 \int_a^b dt \left[ \int_a^b K(x, s)K(s, t) ds \right] v(t) \\ &= v(x) + \lambda \int_a^b dt K(x, t)v(t) + \lambda^2 \int_a^b dt K^2(x, t)v(t), \end{aligned}$$

where  $K^2(x, t) \equiv \int_a^b K(x, s)K(s, t) ds$ . Similar expressions can be derived for  $u_3(x)$ ,  $u_4(x)$ , and so forth. The integrals expressing various "powers" of  $K$  can be obtained using Dirac notation and vectors with continuous indices, as discussed

in Chapter 6. Thus, for instance,

$$\begin{aligned} K^3(x, t) &\equiv \langle x | \mathbf{K} \left( \underbrace{\int_a^b |s_1\rangle \langle s_1| ds_1}_{=1} \right) \mathbf{K} \left( \underbrace{\int_a^b |s_2\rangle \langle s_2| ds_2}_{=1} \right) \mathbf{K} | t \rangle \\ &= \int_a^b ds_1 \int_a^b ds_2 \langle x | \mathbf{K} | s_1 \rangle \langle s_1 | \mathbf{K} | s_2 \rangle \langle s_2 | \mathbf{K} | t \rangle \\ &= \int_a^b ds_1 \int_a^b ds_2 K(x, s_1) K(s_1, s_2) K(s_2, t). \end{aligned}$$

We can always use this technique to convert an equation in kets into an equation in functions and integrals. Therefore, we can concentrate on the abstract operator equation and its various approximations.

Continuing to the  $n$ th-order approximation, we easily obtain

$$|u_n\rangle = |v\rangle + \lambda \mathbf{K} |v\rangle + \cdots + \lambda^n \mathbf{K}^n |v\rangle = \sum_{j=0}^n (\lambda \mathbf{K})^j |v\rangle, \quad (17.8)$$

whose integral form is

$$u_n(x) = \sum_{j=0}^n \lambda^j \int_a^b K^j(x, t) v(t) dt. \quad (17.9)$$

Here  $K^j(x, t)$  is defined inductively by

$$\begin{aligned} K^0(x, t) &= \langle x | \mathbf{K}^0 | t \rangle = \langle x | \mathbf{1} | t \rangle = \langle x | t \rangle = \delta(x - t), \\ K^j(x, t) &= \langle x | \mathbf{K} \mathbf{K}^{j-1} | t \rangle = \langle x | \mathbf{K} \left( \int_a^b |s\rangle \langle s| ds \right) \mathbf{K}^{j-1} | t \rangle \\ &= \int_a^b K(x, s) K^{j-1}(s, t) ds. \end{aligned}$$

The limit of  $u_n(x)$  as  $n \rightarrow \infty$  gives

$$u(x) = \sum_{j=0}^{\infty} \lambda^j \int_a^b K^j(x, t) v(t) dt. \quad (17.10)$$

The convergence of this series, called the **Neumann series**, is always guaranteed for the Volterra equation. For the Fredholm equation, we need to impose the extra condition  $|\lambda| \|\mathbf{K}\| < 1$ .

**17.1.4. Example.** As an example, let us find the solution of  $u(x) = 1 + \lambda \int_0^x u(t) dt$ , a Volterra equation of the second kind. Here,  $v(x) = 1$  and  $K(x, t) = 1$ , and it is straightforward to calculate approximations to  $u(x)$ :

$$\begin{aligned} u_0(x) &= v(x) = 1, & u_1(x) &= 1 + \lambda \int_0^x K(x, t) u_0(t) dt = 1 + \lambda x, \\ u_2(x) &= 1 + \lambda \int_0^x K(x, t) u_1(t) dt = 1 + \lambda \int_0^x (1 + \lambda t) dt = 1 + \lambda x + \frac{\lambda^2 x^2}{2}. \end{aligned}$$

It is clear that the  $n$ th term will look like

$$u_n(x) = 1 + \lambda x + \frac{\lambda^2 x^2}{2} + \cdots + \frac{\lambda^n x^n}{n!} = \sum_{j=0}^n \frac{\lambda^j x^j}{j!}.$$

As  $n \rightarrow \infty$ , we obtain  $u(x) = e^{\lambda x}$ . By direct substitution, it is readily checked that this is indeed a solution of the original integral equation. ■

## 17.2 Fredholm Integral Equations

We can use our knowledge of compact operators gained in the previous chapter to study Fredholm equations of the second kind. With  $\lambda \neq 0$  a complex number, we consider the characteristic equation

$$(\mathbf{1} - \lambda \mathbf{K})|u\rangle = |v\rangle, \quad \text{or} \quad u(x) - \lambda K[u](x) = v(x), \quad (17.11)$$

where all functions are square-integrable on  $[a, b]$ , and  $K(x, t)$ , the Hilbert-Schmidt kernel, is square-integrable on the rectangle  $[a, b] \times [a, b]$ .

Using Proposition 16.2.9, we immediately see that Equation (17.11) has a unique solution if  $|\lambda| \|\mathbf{K}\| < 1$ , and the solution is of the form

$$|u\rangle = (\mathbf{1} - \lambda \mathbf{K})^{-1} |v\rangle = \sum_{n=0}^{\infty} \lambda^n \mathbf{K}^n |v\rangle, \quad (17.12)$$

or  $u(x) = \sum_{n=0}^{\infty} \lambda^n K^n[v](x)$ , where  $K^n[v](x)$  is defined as in Equation (17.4) except that now  $b$  replaces  $x$  as the upper limit of integration.

**17.2.1. Example.** Consider the integral equation

$$u(x) - \int_0^1 K(x, t)u(t) dt = x, \quad \text{where} \quad K(x, t) = \begin{cases} x & \text{if } 0 \leq x < t, \\ t & \text{if } t < x \leq 1. \end{cases}$$

Here  $\lambda = 1$ ; therefore, a Neumann series solution exists if  $\|\mathbf{K}\| < 1$ . It is convenient to write  $K$  in terms of the theta function:<sup>1</sup>

$$K(x, t) = x\theta(t - x) + t\theta(x - t). \quad (17.13)$$

This gives  $|K(x, t)|^2 = x^2\theta(t - x) + t^2\theta(x - t)$  because  $\theta^2(x - t) = \theta(x - t)$  and  $\theta(x - t)\theta(t - x) = 0$ . Thus, we have

$$\begin{aligned} \|\mathbf{K}\|^2 &= \int_0^1 dx \int_0^1 dt |K(x, t)|^2 \\ &= \int_0^1 dx \int_0^1 x^2 \theta(t - x) dt + \int_0^1 dx \int_0^1 t^2 \theta(x - t) dt \\ &= \int_0^1 dt \int_0^t x^2 dx + \int_0^1 dx \int_0^x t^2 dt = \int_0^1 dt \left(\frac{t^3}{3}\right) + \int_0^1 dx \left(\frac{x^3}{3}\right) = \frac{1}{6}. \end{aligned}$$

<sup>1</sup>Recall that the theta function is defined to be 1 if its argument is positive, and 0 if it is negative.



Since this is less than 1, the Neumann series converges, and we have<sup>2</sup>

$$u(x) = \sum_{j=0}^{\infty} \lambda^j \int_a^b K^j(x, t)v(t) dt = \sum_{j=0}^{\infty} \int_0^1 K^j(x, t) dt \equiv \sum_{j=0}^{\infty} f_j(x).$$

The first few terms are evaluated as follows:

$$\begin{aligned} f_0(x) &= \int_0^1 K^0(x, t)t dt = \int_0^1 \delta(x, t)t dt = x \\ f_1(x) &= \int_0^1 K(x, t)t dt = \int_0^1 [x\theta(t-x) + t\theta(x-t)]t dt \\ &= x \int_x^1 t dt + \int_0^x t^2 dt = \frac{x}{2} - \frac{x^3}{6}. \end{aligned}$$

The next term is trickier than the first two because of the product of the theta functions. We first substitute Equation (17.13) in the integral for the second-order term, and simplify

$$\begin{aligned} f_2(x) &= \int_0^1 K^2(x, t)t dt = \int_0^1 t dt \int_0^1 K(x, s)K(s, t) ds \\ &= \int_0^1 t dt \int_0^1 [x\theta(s-x) + s\theta(x-s)][s\theta(t-s) + t\theta(s-t)] ds \\ &= x \int_0^1 t dt \int_0^1 s\theta(s-x)\theta(t-s) ds + x \int_0^1 t^2 dt \int_0^1 \theta(s-x)\theta(s-t) ds \\ &\quad + \int_0^1 t dt \int_0^1 s^2\theta(x-s)\theta(t-s) ds + \int_0^1 t^2 dt \int_0^1 s\theta(x-s)\theta(s-t) ds. \end{aligned}$$

It is convenient to switch the order of integration at this point. This is because of the presence of  $\theta(x-s)$  and  $\theta(s-x)$ , which do not involve  $t$  and are best integrated last. Thus, we have

$$\begin{aligned} f_2(x) &= x \int_0^1 s\theta(s-x) ds \int_s^1 t dt + x \int_0^1 \theta(s-x) ds \int_0^s t^2 dt \\ &\quad + \int_0^1 s^2\theta(x-s) ds \int_s^1 t dt + \int_0^1 s\theta(x-s) ds \int_0^s t^2 dt \\ &= x \int_x^1 s ds \left( \frac{1}{2} - \frac{s^2}{2} \right) + x \int_x^1 ds \frac{s^3}{3} + \int_0^x s^2 ds \left( \frac{1}{2} - \frac{s^2}{2} \right) + \int_0^x s ds \frac{s^3}{3} \\ &= \frac{5}{24}x - \frac{1}{12}x^3 + \frac{1}{120}x^5. \end{aligned}$$

As a test of his/her knowledge of  $\theta$ -function manipulation, the reader is urged to perform the integration in reverse order. Adding all the terms, we obtain an approximation for  $u(x)$  that is valid for  $0 \leq x \leq 1$ :

$$u(x) \approx f_0(x) + f_1(x) + f_2(x) = \frac{41}{24}x - \frac{1}{4}x^3 + \frac{1}{120}x^5. \quad \blacksquare$$

<sup>2</sup>Note that in this case (Fredholm equation), we can calculate the  $j$ th term in isolation. In the Volterra case, it was more natural to calculate the solution up to a given order.

We have seen that the Volterra equation of the second kind has a unique solution which can be written as an infinite series (see Theorem 17.1.2). The case of the Fredholm equation of the second kind is more complicated because of the existence of eigenvalues. The general solution of Equation (17.11) is discussed in the following:

**Fredholm alternative** **17.2.2. Theorem.** (Fredholm Alternative) *Let  $\mathbf{K}$  be a Hilbert–Schmidt operator and  $\lambda$  a complex number. Then either*

1.  $\lambda$  is a regular value of Equation (17.11)—or  $\lambda^{-1}$  is a regular point of the operator  $\mathbf{K}$ —in which case the equation has the unique solution  $|u\rangle = (\mathbf{1} - \lambda\mathbf{K})^{-1} |v\rangle$ , or
2.  $\lambda$  is a characteristic value of Equation (17.11) ( $\lambda^{-1}$  is an eigenvalue of the operator  $\mathbf{K}$ ), in which case the equation has a solution if and only if  $|v\rangle$  is in the orthogonal complement of the (finite-dimensional) null space of  $\mathbf{1} - \lambda^*\mathbf{K}^\dagger$ .

*Proof.* The first part is trivial if we recall that by definition, regular points of  $\mathbf{K}$  are those complex numbers  $\mu$  which make the operator  $\mathbf{K} - \mu\mathbf{1}$  invertible.

For part (2), we first show that the null space of  $\mathbf{1} - \lambda^*\mathbf{K}^\dagger$  is finite-dimensional. We note that  $\mathbf{1} - \lambda\mathbf{K}$  is invertible if and only if its adjoint  $\mathbf{1} - \lambda^*\mathbf{K}^\dagger$  is invertible, and  $\lambda \in \rho(\mathbf{K})$  iff  $\lambda^* \in \rho(\mathbf{K}^\dagger)$ . Since the spectrum of an operator is composed of all points that are not regular, we conclude that  $\lambda$  is in the spectrum of  $\mathbf{K}$  if and only if  $\lambda^*$  is in the spectrum of  $\mathbf{K}^\dagger$ . For compact operators, all nonzero points of the spectrum are eigenvalues. Therefore, the nonzero points of the spectrum of  $\mathbf{K}^\dagger$ , a compact operator by Theorem 16.5.7, are all eigenvalues of  $\mathbf{K}^\dagger$ , and the null space of  $\mathbf{1} - \lambda^*\mathbf{K}^\dagger$  is finite-dimensional (Theorem 16.6.2). Next, we note that the equation itself requires that  $|v\rangle$  be in the range of the operator  $\mathbf{1} - \lambda\mathbf{K}$ , which, by Theorem 16.6.2, is the orthogonal complement of the null space of  $\mathbf{1} - \lambda^*\mathbf{K}^\dagger$ .  $\square$

**Erik Ivar Fredholm** (1866–1927) was born in Stockholm, the son of a well-to-do merchant family. He received the best education possible and soon showed great promise in mathematics, leaning especially toward the applied mathematics of practical mechanics in a year of study at Stockholm's Polytechnic Institute. Fredholm finished his education at the University of Uppsala, obtaining his doctorate in 1898. He also studied at the University of Stockholm during this same period and eventually received an appointment to the faculty there. Fredholm remained there the rest of his professional life.



His first contribution to mathematics was contained in his doctoral thesis, in which he studied a first-order partial differential equation in three variables, a problem that arises in the deformation of anisotropic media. Several years later

he completed this work by finding the fundamental solution to a general elliptic partial differential equation with constant coefficients.

Fredholm is perhaps best known for his studies of the integral equation that bears his name. Such equations occur frequently in physics. Fredholm's genius led him to note the similarity between his equation and a relatively familiar matrix-vector equation, resulting in his identification of a quantity that plays the same role in his equation as the determinant plays in the matrix-vector equation. He thus obtained a method for determining the existence of a solution and later used an analogous expression to derive a solution to his equation akin to the Cramer's rule solution to the matrix-vector equation. He further showed that the solution could be expressed as a power series in a complex variable. This latter result was considered important enough that Poincaré assumed it without proof (in fact he was unable to prove it) in a study of related partial differential equations.

Fredholm then considered the homogeneous form of his equation. He showed that under certain conditions, the vector space of solutions is finite-dimensional. David Hilbert later extended Fredholm's work to a complete eigenvalue theory of the Fredholm equation, which ultimately led to the discovery of Hilbert spaces.

### 17.2.1 Hermitian Kernel

Of special interest are integral equations in which the kernel is hermitian, which occurs exactly when the operator is hermitian. Such a kernel has the property that<sup>3</sup>  $\langle x | \mathbf{K} | t \rangle^* = \langle t | \mathbf{K} | x \rangle$  or  $[K(x, t)]^* = K(t, x)$ . For such kernels we can use the spectral theorem for compact hermitian operators to find a series solution for the integral equation. First we recall that

$$\mathbf{K} = \sum_{j=1}^N \lambda_j^{-1} \mathbf{P}_j = \sum_{j=1}^N \lambda_j^{-1} \sum_{k=1}^{m_k} |e_k^{(j)}\rangle \langle e_k^{(j)}|,$$

where we have used  $\lambda_j^{-1}$  to denote the eigenvalue of the operator<sup>4</sup> and expanded the projection operator in terms of orthonormal basis vectors of the corresponding finite-dimensional eigenspace. Recall that  $N$  can be infinity. Instead of the double sum, we can sum once over all the basis vectors and write  $\mathbf{K} = \sum_{n=1}^{\infty} \lambda_n^{-1} |u_n\rangle \langle u_n|$ . Here  $n$  counts all the orthonormal eigenvectors of the Hilbert space, and  $\lambda_n^{-1}$  is the eigenvalue corresponding to the eigenvector  $|u_n\rangle$ . Therefore,  $\lambda_n^{-1}$  may be repeated in the sum. The action of  $\mathbf{K}$  on a vector  $|u\rangle$  is given by

$$\mathbf{K} |u\rangle = \sum_{n=1}^{\infty} \lambda_n^{-1} \langle u_n | u \rangle |u_n\rangle. \quad (17.14)$$

<sup>3</sup>Since we are dealing mainly with real functions, hermiticity of  $\mathbf{K}$  implies the symmetry of  $K$ , i.e.,  $K(x, t) = K(t, x)$ .

<sup>4</sup> $\lambda_j$  is the characteristic value of the integral equation, or the inverse of the eigenvalue of the corresponding operator.

If the Hilbert space is  $\mathcal{L}^2[a, b]$ , we may be interested in the functional form of this equation. We obtain such a form by multiplying both sides by  $\langle x |$ :

$$K[u](x) \equiv \langle x | \mathbf{K} | u \rangle = \sum_{n=1}^{\infty} \lambda_n^{-1} \langle u_n | u \rangle \langle x | u_n \rangle = \sum_{n=1}^{\infty} \lambda_n^{-1} \langle u_n | u \rangle u_n(x).$$

Hilbert-Schmidt theorem

That this series converges *uniformly* in the interval  $[a, b]$  is known as the **Hilbert-Schmidt theorem**.

**17.2.3. Example.** Let us solve  $u(x) = x + \lambda \int_a^b K(x, t)u(t) dt$ , where  $K(x, t) \equiv xt$  is a symmetric (hermitian) kernel, by the Neumann series method. We note that

$$\begin{aligned} \|\mathbf{K}\|^2 &= \int_a^b \int_a^b |K(x, t)|^2 dx dt = \int_a^b \int_a^b x^2 t^2 dx dt \\ &= \int_a^b x^2 dx \int_a^b t^2 dt = \left( \int_a^b x^2 dx \right)^2 = \frac{1}{9}(b^3 - a^3)^2, \end{aligned}$$

or

$$\|\mathbf{K}\| = \int_a^b x^2 dx = \frac{1}{3}(b^3 - a^3),$$

and the Neumann series converges if  $|\lambda| (b^3 - a^3) < 3$ . Assuming that this condition holds, we have

$$u(x) = x + \sum_{j=1}^{\infty} \lambda^j \int_a^b K^j(x, t) dt.$$

The special form of the kernel allows us to calculate  $K^j(x, t)$  directly:

$$\begin{aligned} K^j(x, t) &= \int_a^b \int_a^b \cdots \int_a^b K(x, s_1)K(s_1, s_2) \cdots K(s_{j-1}, t) ds_1 ds_2 \cdots ds_{j-1} \\ &= \int_a^b \int_a^b \cdots \int_a^b x s_1^2 s_2^2 \cdots s_{j-1}^2 t ds_1 ds_2 \cdots ds_{j-1} \\ &= xt \left( \int_a^b s^2 ds \right)^{j-1} = xt \|\mathbf{K}\|^{j-1}. \end{aligned}$$

It follows that  $\int_a^b K^j(x, t) dt = x \|\mathbf{K}\|^{j-1} \frac{1}{3}(b^3 - a^3) = x \|\mathbf{K}\|^j$ . Substituting this in the expression for  $u(x)$  yields

$$\begin{aligned} u(x) &= x + \sum_{j=1}^{\infty} \lambda^j x \|\mathbf{K}\|^j = x + x \lambda \|\mathbf{K}\| \sum_{j=1}^{\infty} \lambda^{j-1} \|\mathbf{K}\|^{j-1} \\ &= x \left( 1 + \lambda \|\mathbf{K}\| \frac{1}{1 - \lambda \|\mathbf{K}\|} \right) = \frac{x}{1 - \lambda \|\mathbf{K}\|} = \frac{3x}{3 - \lambda(b^3 - a^3)}. \end{aligned}$$

Because of the simplicity of the kernel, we can solve the integral equation exactly. First we write

$$u(x) = x + \lambda \int_a^b x t u(t) dt = x + \lambda x \int_a^b t u(t) dt \equiv x(1 + \lambda A), \tag{17.15}$$

where  $A = \int_a^b tu(t) dt$ . Multiplying both sides by  $x$  and integrating, we obtain

$$A = \int_a^b xu(x) dx = (1 + \lambda A) \int_a^b x^2 dx = (1 + \lambda A)\|K\| \Rightarrow A = \frac{\|K\|}{1 - \lambda\|K\|}.$$

Substituting  $A$  in Equation (17.15) gives

$$u(x) = x \left( 1 + \lambda \frac{\|K\|}{1 - \lambda\|K\|} \right) = \frac{x}{1 - \lambda\|K\|}.$$

This solution is the same as the first one we obtained. However, no series was involved here, and therefore no assumption is necessary concerning  $|\lambda| \|K\|$ . ■

If one can calculate the eigenvectors  $|u_n\rangle$  and the eigenvalues  $\lambda_n^{-1}$ , then one can obtain a solution for the integral equation in terms of these eigenfunctions as follows: Substitute (17.14) in the Fredholm equation [Equation (17.3)] to get

$$|u\rangle = |v\rangle + \lambda \sum_{n=1}^{\infty} \lambda_n^{-1} \langle u_n | u \rangle |u_n\rangle. \tag{17.16}$$

Multiply both sides by  $\langle u_m |$ :

$$\langle u_m | u \rangle = \langle u_m | v \rangle + \lambda \sum_{n=1}^{\infty} \lambda_n^{-1} \langle u_n | u \rangle \underbrace{\langle u_m | u_n \rangle}_{=\delta_{mn}} = \langle u_m | v \rangle + \lambda \lambda_m^{-1} \langle u_m | u \rangle, \tag{17.17}$$

or, if  $\lambda$  is not one of the eigenvalues,

$$\left( 1 - \frac{\lambda}{\lambda_m} \right) \langle u_m | u \rangle = \langle u_m | v \rangle \Rightarrow \langle u_m | u \rangle = \frac{\lambda_m \langle u_m | v \rangle}{\lambda_m - \lambda}.$$

Substituting this in Equation (17.16) gives

$$|u\rangle = |v\rangle + \lambda \sum_{n=1}^{\infty} \frac{\langle u_n | v \rangle}{\lambda_n - \lambda} |u_n\rangle, \tag{17.18}$$

and in the functional form,

$$u(x) = v(x) + \lambda \sum_{n=1}^{\infty} \frac{\langle u_n | v \rangle}{\lambda_n - \lambda} u_n(x), \quad \lambda \neq \lambda_n \forall n. \tag{17.19}$$

In case  $\lambda = \lambda_m$  for some  $m$ , the Fredholm alternative (Theorem 17.2.2) says that we will have a solution only if  $|v\rangle$  is in the orthogonal complement of the null space of  $1 - \lambda_m K$ . Moreover, Equation (17.17) shows that  $\langle u_m | u \rangle$ , the expansion coefficients of the basis vectors of the eigenspace  $\mathcal{M}_m$ , cannot be specified.

<sup>5</sup>Remember that  $K$  is hermitian; therefore,  $\lambda_m$  is real.

However, Equation (17.17) does determine the rest of the coefficients as before. In this case, the solution can be written as

$$|u\rangle = |v\rangle + \sum_{k=1}^r c_k |u_m^{(k)}\rangle + \lambda \sum_{\substack{n=1 \\ n \neq m}}^{\infty} \frac{\langle u_n | v \rangle}{\lambda_n - \lambda} |u_n\rangle, \quad (17.20)$$

where  $r$  is the (finite) dimension of  $\mathcal{M}_m$ ,  $k$  labels the orthonormal basis  $\{|u_m^{(k)}\rangle\}$  of  $\mathcal{M}_m$ , and  $\{c_k\}_{k=1}^r$  are arbitrary constants. In functional form, this equation becomes

$$u(x) = v(x) + \sum_{k=1}^r c_k u_m^{(k)}(x) + \lambda \sum_{\substack{n=1 \\ n \neq m}}^{\infty} \frac{\langle u_n | v \rangle}{\lambda_n - \lambda} u_n(x). \quad (17.21)$$

**17.2.4. Example.** We now give an example of the application of Equation (17.19). We want to solve  $u(x) = 3 \int_{-1}^1 K(x, t) u(t) dt + x^2$  where

$$K(x, t) = \sum_{k=0}^{\infty} \frac{u_k(x) u_k(t)}{2^{k/2}}, \quad u_k(x) = \sqrt{\frac{2k+1}{2}} P_k(x),$$

and  $P_k(x)$  is a Legendre polynomial.

We first note that  $\{u_k\}$  is an orthonormal set of functions, that  $K(x, t)$  is real and symmetric (therefore, hermitian), and that

$$\begin{aligned} \int_{-1}^1 dt \int_{-1}^1 dx |K(x, t)|^2 &= \int_{-1}^1 dt \int_{-1}^1 dx \sum_{k,l=0}^{\infty} \frac{u_k(x) u_k(t)}{2^{k/2}} \frac{u_l(x) u_l(t)}{2^{l/2}} \\ &= \sum_{k,l=0}^{\infty} \frac{1}{2^{k/2}} \frac{1}{2^{l/2}} \underbrace{\int_{-1}^1 u_k(x) u_l(x) dx}_{=\delta_{kl}} \underbrace{\int_{-1}^1 u_k(t) u_l(t) dt}_{=\delta_{kl}} \\ &= \sum_{k=0}^{\infty} \frac{1}{2^k} \delta_{kk} = \sum_{k=0}^{\infty} \frac{1}{2^k} = 2 < \infty. \end{aligned}$$

Thus,  $K(x, t)$  is a Hilbert–Schmidt kernel.

Now note that

$$\begin{aligned} \int_{-1}^1 K(x, t) u_k(t) dt &= \int_{-1}^1 \sum_{l=0}^{\infty} \frac{u_l(x) u_l(t)}{2^{l/2}} u_k(t) dt \\ &= \sum_{l=0}^{\infty} \frac{u_l(x)}{2^{l/2}} \underbrace{\int_{-1}^1 u_l(t) u_k(t) dt}_{=\delta_{kl}} = \frac{1}{2^{k/2}} u_k(x). \end{aligned}$$

This shows that  $u_k$  is an eigenfunction of  $K(x, t)$  with eigenvalue  $1/2^{k/2}$ . Since  $3 \neq 1/2^{k/2}$  for any integer  $k$ , we can use Equation (17.19) to write

$$u(x) = x^2 + 3 \sum_{k=0}^{\infty} \frac{\int_{-1}^1 u_k(s) s^2 ds}{2^{k/2} - 3} u_k(x).$$

But  $\int_{-1}^1 u_k(s)s^2 ds = 0$  for  $k \geq 3$ . For  $k \leq 2$ , we use the first three Legendre polynomials to get

$$\int_{-1}^1 u_0(s)s^2 ds = \frac{\sqrt{2}}{3}, \quad \int_{-1}^1 u_1(s)s^2 ds = 0, \quad \int_{-1}^1 u_2(s)s^2 ds = \frac{2\sqrt{2}}{3\sqrt{5}}.$$

This gives  $u(x) = \frac{1}{2} - 2x^2$ . The reader is urged to substitute this solution in the original integral equation and verify that it works. ■

## 17.2.2 Degenerate Kernels

degenerate or  
separable kernel

The preceding example involves the simplest kind of degenerate, or separable, kernels. A kernel is called **degenerate**, or **separable**, if it can be written as a finite sum of products of functions of one variable:

$$K(x, t) = \sum_{j=1}^n \phi_j(x)\psi_j^*(t), \quad (17.22)$$

where  $\phi_j$  and  $\psi_j$  are assumed to be square-integrable. Substituting (17.22) in the Fredholm integral equation of the second kind, we obtain

$$u(x) - \lambda \sum_{j=1}^n \phi_j(x) \int_a^b \psi_j^*(t)u(t) dt = v(x).$$

If we define  $\mu_j = \int_a^b \psi_j^*(t)u(t) dt$ , the preceding equation becomes

$$u(x) - \lambda \sum_{j=1}^n \mu_j \phi_j(x) = v(x). \quad (17.23)$$

Multiply this equation by  $\psi_i^*(x)$  and integrate over  $x$  to get

$$\mu_i - \lambda \sum_{j=1}^n \mu_j A_{ij} = v_i \quad \text{for } i = 1, 2, \dots, n, \quad (17.24)$$

where  $A_{ij} = \int_a^b \psi_i^*(t)\phi_j(t) dt$  and  $v_i = \int_a^b \psi_i^*(t)v(t) dt$ . With  $\mu_i$ ,  $v_i$ , and  $A_{ij}$  as components of column vectors  $u$ ,  $v$ , and a matrix  $A$ , we can write the above linear system of equations as

$$u - \lambda Au = v, \quad \text{or} \quad (1 - \lambda A)u = v. \quad (17.25)$$

We can now determine the  $\mu_i$  by solving the system of linear equations given by (17.24). Once the  $\mu_i$  are determined, Equation (17.23) gives  $u(x)$ . Thus, for a degenerate kernel the Fredholm problem reduces to a system of linear equations.

**17.2.5. Example.** As a concrete example of an integral equation with degenerate kernel, we solve  $u(x) - \lambda \int_0^1 (1+xt)u(t) dt = x$  for two different values of  $\lambda$ . The kernel,  $K(x, t) = 1 + xt$ , is separable with  $\phi_1(x) = 1$ ,  $\psi_1(t) = 1$ ,  $\phi_2(x) = x$ , and  $\psi_2(t) = t$ . This gives the matrix

$$A = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}.$$

For convenience, we define the matrix  $B \equiv 1 - \lambda A$ .

(a) First assume that  $\lambda = 1$ . In that case  $B$  has a nonzero determinant. Thus,  $B^{-1}$  exists, and can be calculated to be

$$B^{-1} = \begin{pmatrix} -\frac{8}{3} & -2 \\ -2 & 0 \end{pmatrix}.$$

With

$$v_1 = \int_0^1 \psi_1^*(t)v(t) dt = \int_0^1 t dt = \frac{1}{2} \quad \text{and} \quad v_2 = \int_0^1 \psi_2^*(t)v(t) dt = \int_0^1 t^2 dt = \frac{1}{3}$$

we obtain

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = B^{-1}v = \begin{pmatrix} -\frac{8}{3} & -2 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix}.$$

Equation (17.23) then gives  $u(x) = \mu_1\phi_1(x) + \mu_2\phi_2(x) + x = -2$ .

(b) Now, for the purpose of illustrating the other alternative of Theorem 17.2.2, let us take  $\lambda = 8 + 2\sqrt{13}$ . Then

$$B = 1 - \lambda A = - \begin{pmatrix} 7 + 2\sqrt{13} & 4 + \sqrt{13} \\ 4 + \sqrt{13} & (5 + 2\sqrt{13})/3 \end{pmatrix},$$

and  $\det B = 0$ . This shows that  $8 + 2\sqrt{13}$  is a characteristic value of the equation. We thus have a solution only if  $v(x) \equiv x$  is orthogonal to the null space of  $1 - \lambda^*A^\dagger = B^\dagger$ . To determine a basis for this null space, we have to find vectors  $|z\rangle$  such that  $B^\dagger|z\rangle = 0$ . Since  $\lambda$  is real, and  $B$  is real and symmetric,  $B^\dagger = B$ , and we must solve

$$\begin{pmatrix} 7 + 2\sqrt{13} & 4 + \sqrt{13} \\ 4 + \sqrt{13} & (5 + 2\sqrt{13})/3 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = 0.$$

The solution to this equation is a multiple of  $|z\rangle \equiv \begin{pmatrix} 3 \\ -2 - \sqrt{13} \end{pmatrix}$ . If the integral equation is to have a solution, the column vector  $v$  (whose corresponding ket we denote by  $|v\rangle$ ) must be orthogonal to  $|z\rangle$ . But

$$\langle z|v\rangle = (3 \quad -2 - \sqrt{13}) \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \end{pmatrix} \neq 0.$$

Therefore, the integral equation has no solution. ■



The reader may feel uneasy that the functions  $\phi_j(x)$  and  $\psi_j(t)$  appearing in a degenerate kernel are arbitrary to within a multiplicative function. After all, we can multiply  $\phi_j(x)$  by a nonzero function, and divide  $\psi_j(t)$  by the same function, and get the same kernel. Such a change clearly alters the matrices A and B and therefore seems likely to change the solution,  $u(x)$ . That this is not the case is demonstrated in Problem 17.2. In fact, it can be shown quite generally that the transformations described above do not change the solution.

As the alert reader may have noticed, we have been avoiding the problem of solving the eigenvalue (characteristic) problem for integral operators. Such a problem is nontrivial, and the analogue of the finite-dimensional case, where one works with determinants and characteristic polynomials, does not exist. An exception is a degenerate hermitian<sup>6</sup> kernel, i.e., a kernel of the form  $K(x, t) = \sum_{i=1}^n h_i(x)h_i^*(t)$ . Substituting this in the characteristic-value equation

$$u(x) = \lambda \int_a^b K(x, t)u(t) dt, \quad \lambda \neq 0,$$

we obtain  $u(x) = \lambda \sum_{i=1}^n h_i(x) \int_a^b h_i^*(t)u(t) dt$ . Defining  $\mu_i \equiv \int_a^b h_i^*(t)u(t) dt$  and substituting it back in the equation gives

$$u(x) = \lambda \sum_{i=1}^n h_i(x)\mu_i. \quad (17.26)$$

Multiplying this equation by  $\lambda^{-1}h_k^*(x)$  and integrating over  $x$  yields

$$\lambda^{-1}\mu_k = \sum_{i=1}^n \left[ \int_a^b h_k^*(x)h_i(x) dx \right] \mu_i \equiv \sum_{i=1}^n m_{ki}\mu_i.$$

This is an eigenvalue equation for the hermitian  $n \times n$  matrix M with elements  $m_{ij}$ , which, by spectral theorem for hermitian operators, can be solved. In fact, the matrix need not be hermitian; as long as it is *normal*, the eigenvalue problem can be solved. Once the eigenvectors and the eigenvalues are found, we can substitute them in Equation (17.26) and obtain  $u(x)$ . We expect to find a finite number of eigenfunctions and eigenvalues. Our analysis of compact operators included such a case. That analysis also showed that the entire (infinite-dimensional) Hilbert space could be written as the direct sum of eigenspaces that are finite-dimensional for nonzero eigenvalues. Therefore, we expect the eigenspace corresponding to the zero eigenvalue (or infinite characteristic value) to be infinite-dimensional. The following example illustrates these points.

**17.2.6. Example.** Let us find the nonzero characteristic values and corresponding eigenfunctions of the kernel  $K(x, t) = 1 + \sin(x + t)$  for  $-\pi \leq x, t \leq \pi$ .

<sup>6</sup>Actually, the problem of a degenerate kernel that leads to a normal matrix, as described below, can also be solved.

We are seeking functions  $u$  and scalars  $\lambda$  satisfying  $u(x) = \lambda K[u](x)$ , or

$$u(x) = \lambda \int_{-\pi}^{\pi} [1 + \sin(x+t)]u(t) dt.$$

Expanding  $\sin(x+t)$ , we obtain

$$u(x) = \lambda \int_{-\pi}^{\pi} [1 + \sin x \cos t + \cos x \sin t]u(t) dt, \quad (17.27)$$

or

$$\lambda^{-1}u(x) = \mu_1 + \mu_2 \sin x + \mu_3 \cos x, \quad (17.28)$$

where  $\mu_1 = \int_{-\pi}^{\pi} u(t) dt$ ,  $\mu_2 = \int_{-\pi}^{\pi} u(t) \cos t dt$ , and  $\mu_3 = \int_{-\pi}^{\pi} u(t) \sin t dt$ . Integrate both sides of Equation (17.28) with respect to  $x$  from  $-\pi$  to  $\pi$  to obtain  $\lambda^{-1}\mu_1 = 2\pi\mu_1$ . Similarly, multiplying by  $\sin x$  and  $\cos x$  and integrating yields

$$\lambda^{-1}\mu_2 = \pi\mu_3 \quad \text{and} \quad \lambda^{-1}\mu_3 = \pi\mu_2. \quad (17.29)$$

If  $\mu_1 \neq 0$ , we get  $\lambda^{-1} = 2\pi$ , which, when substituted in (17.29), yields  $\mu_2 = \mu_3 = 0$ . We thus have, as a first solution,  $\lambda_1^{-1} = 2\pi$  and  $|u_1\rangle = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ , where  $\alpha$  is an arbitrary constant. Equation (17.28) now gives  $\lambda_1^{-1}u_1(x) = \mu_1$ , or  $u_1(x) = c_1$ , where  $c_1$  is an arbitrary constant to be determined.

On the other hand,  $\mu_1 = 0$  if  $\lambda^{-1} \neq 2\pi$ . Then Equation (17.29) yields  $\lambda^{-1} = \pm\pi$  and  $\mu_2 = \pm\mu_3$ . For  $\lambda^{-1} \equiv \lambda_{\pm}^{-1} = \pi$ , Equation (17.28) gives

$$u(x) \equiv u_{+}(x) = c_{+}(\sin x + \cos x),$$

and for  $\lambda^{-1} \equiv \lambda_{-}^{-1} = -\pi$ , it yields  $u(x) \equiv u_{-}(x) = c_{-}(\sin x - \cos x)$ , where  $c_{\pm}$  are arbitrary constants to be determined by normalization of eigenfunctions. The normalized eigenfunctions are

$$u_1 = \frac{1}{\sqrt{2\pi}}, \quad u_{\pm}(x) = \frac{1}{\sqrt{2\pi}}(\sin x \pm \cos x).$$

Direct substitution in the original integral equation easily verifies that  $u_1$ ,  $u_{+}$ , and  $u_{-}$  are eigenfunctions of the integral equation with the eigenvalues calculated above.

Let us now consider the zero eigenvalue (or infinite characteristic value). Divide both sides of Equation (17.27) by  $\lambda$  and take the limit of  $\lambda \rightarrow \infty$ . Then the integral equation becomes

$$\int_{-\pi}^{\pi} [1 + \sin x \cos t + \cos x \sin t]u(t) dt = 0.$$

The solutions  $u(t)$  to this equation would span the eigenspace corresponding to the zero eigenvalue, or infinite characteristic value. We pointed out above that this eigenspace is expected to be infinite-dimensional. This expectation is borne out once we note that all functions of the form  $\sin nt$  or  $\cos nt$  with  $n \geq 2$  make the above integral zero; and there are infinitely many such functions. ■

## 17.3 Problems

17.1. Use mathematical induction to derive Equation (17.5).

17.2. Repeat part (a) of Example 17.2.5 using

$$\phi_1(x) = \frac{1}{2}, \quad \psi_1(t) = 2, \quad \phi_2(x) = x, \quad \psi_2(t) = t$$

so that we still have  $K(x, t) = \phi_1(x)\psi_1(t) + \phi_2(x)\psi_2(t)$ .

17.3. Use the spectral theorem for compact hermitian operators to show that if the kernel of a Hilbert–Schmidt operator has a finite number of nonzero eigenvalues, then the kernel is separable. Hint: See the discussion at the beginning of Section 17.2.1.

17.4. Use the method of successive approximations to solve the Volterra equation  $u(x) = \lambda \int_0^x u(t) dt$ . Then derive a DE equivalent to the Volterra equation (make sure to include the initial condition), and solve it.

17.5. Regard the Fourier transform,

$$\mathbf{F}[f](x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixy} f(y) dy$$

as an integral operator.

(a) Show that  $\mathbf{F}^2[f](x) = f(-x)$ .

(b) Deduce, therefore, that the only eigenvalues of this operator are  $\lambda = \pm 1, \pm i$ .

(c) Let  $f(x)$  be any even function of  $x$ . Show that an appropriate choice of  $\alpha$  can make  $u = f + \alpha \mathbf{F}[f]$  an eigenfunction of  $\mathbf{F}$ . (This shows that the eigenvalues of  $\mathbf{F}$  have infinite multiplicity.)

17.6. For what values of  $\lambda$  does the following integral equation have a solution?

$$u(x) = \lambda \int_0^\pi \sin(x+t)u(t) dt + x.$$

What is that solution? Redo the problem using a Neumann series expansion. Under what condition is the series convergent?

17.7. It is possible to multiply the functions  $\phi_j(x)$  by  $\gamma_j(x)$  and  $\psi_j(t)$  by  $1/\gamma_j(t)$  and still get the same degenerate kernel,  $K(x, t) = \sum_{j=1}^n \phi_j(x)\psi_j(t)$ . Show that such arbitrariness, although affecting the matrices  $\mathbf{A}$  and  $\mathbf{B}$ , does not change the solution of the Fredholm problem

$$u(x) - \lambda \int_a^b K(x, t)u(t) dt = f(x).$$

17.8. Show, by direct substitution, that the solution found in Example 17.2.4 does satisfy its integral equation.

17.9. Solve  $u(x) = \frac{1}{2} \int_{-1}^1 (x+t)u(t) dt + x$ .

17.10. Solve  $u(x) = \lambda \int_0^1 xt u(t) dt + x$  using the Neumann series method. For what values of  $\lambda$  is the series convergent? Now find the eigenvalues and eigenfunctions of the kernel and solve the problem using these eigenvalues and eigenfunctions.

17.11. Solve  $u(x) = \lambda \int_0^\infty K(x, t)u(t) dt + x^\alpha$ , where  $\alpha$  is any real number except a negative integer, and  $K(x, t) = e^{-(x+t)}$ . For what values of  $\lambda$  does the integral equation have a solution?

17.12. Solve the integral equations

$$(a) u(x) = e^x + \lambda \int_0^1 xt u(t) dt. \quad (b) u(x) = \lambda \int_0^\pi \sin(x-t)u(t) dt.$$

$$(c) u(x) = x^2 + \int_0^1 xt u(t) dt. \quad (d) u(x) = x + \int_0^x u(t) dt.$$

17.13. Solve the integral equation  $u(x) = x + \lambda \int_0^1 (x+t)tu(t) dt$ , keeping terms up to  $\lambda^2$ .

17.14. Solve the integral equation  $u(x) = e^{-|x|} + \lambda \int_{-\infty}^\infty e^{-|x-t|}u(t) dt$ , assuming that  $f$  remains finite as  $x \rightarrow \pm\infty$ .

17.15. Solve the integral equation  $u(x) = e^{-|x|} + \lambda \int_0^\infty u(t) \cos xt dt$ , assuming that  $f$  remains finite as  $x \rightarrow \pm\infty$ .

## Additional Reading

1. DeVito, C. *Functional Analysis and Linear Operator Theory*, Addison-Wesley, 1990.
2. Jörgen, K. *Linear Integral Operators*, Pitman, 1982. Translated from its original German, this is a thorough (but formal) introduction to integral operators and equations.
3. Zeidler, E. *Applied Functional Analysis*, Springer-Verlag, 1995.

# Sturm–Liouville Systems: Formalism

The linear operators discussed in the last two chapters were exclusively *integral* operators. Most applications of physical interest, however, involve *differential* operators (DO). Unfortunately, differential operators are unbounded. We noted that complications arise when one abandons the compactness property of the operator, e.g., sums turn into integrals and one loses one's grip over the eigenvalues of noncompact operators. The transition to unbounded operators further complicates matters. Fortunately, the formalism of one type of DOs that occur most frequently in physics can be studied in the context of compact operators. Such a study is our aim for this chapter.

## 18.1 Unbounded Operators with Compact Resolvent

As was pointed out in Example 16.2.7, the derivative operator cannot be defined for *all* functions in  $\mathcal{L}^2(a, b)$ . This motivates the following:

**18.1.1. Definition.** Let  $\mathcal{D}$  be a linear manifold<sup>1</sup> in the Hilbert space  $\mathcal{H}$ . A linear map  $\mathbf{T} : \mathcal{D} \rightarrow \mathcal{H}$  will be called a **linear operator in**<sup>2</sup>  $\mathcal{H}$ .  $\mathcal{D}$  is called the **domain of  $\mathbf{T}$**  and often denoted by  $\mathcal{D}(\mathbf{T})$ .

domain of a linear operator

**18.1.2. Example.** The domain of the derivative operator  $\mathbf{D}$ , as an operator on  $\mathcal{L}^2(a, b)$ , cannot be the entire space. On the other hand,  $\mathbf{D}$  is defined on the linear manifold  $\mathcal{M}$  in  $\mathcal{L}^2(a, b)$  spanned by  $\{e^{i2n\pi x/L}\}$  with  $L = b - a$ . As we saw in Chapter 8,  $\mathcal{M}$  is dense

---

<sup>1</sup>A linear manifold of an infinite-dimensional normed vector space  $\mathcal{V}$  is a proper subset that is a vector space in its own right, but is not necessarily closed.

<sup>2</sup>As opposed to *on*  $\mathcal{H}$ .

(see Definition 16.4.5 and the discussion following it) in  $\mathcal{L}^2(a, b)$ . This is the essence of Fourier series: That every function in  $\mathcal{L}^2(a, b)$  can be expanded in (i.e., approximated by) a Fourier series. It turns out that many unbounded operators on a Hilbert space share the same property, namely that their domains are dense in the Hilbert space.

Another important property of Fourier expansion is the fact that if the function is differentiable, then one can differentiate both sides, i.e., one can differentiate a Fourier expansion term by term if such an operation makes sense for the original function. Define the sequence  $\{f_m\}$  by

$$f_m(x) = \sum_{n=-m}^m a_n e^{i2\pi nx/L}, \quad a_n = \frac{1}{\sqrt{L}} \int_a^b f(x) e^{-i2\pi nx/L} dx.$$

Then we can state the property above as follows: Suppose  $\{f_m\}$  is in  $\mathcal{M}$ . If  $\lim f_m = f$  and  $\lim f'_m = g$ , then  $f' = g$  and  $f \in \mathcal{M}$ . Many unbounded operators share this property. ■

**18.1.3. Definition.** Let  $\mathcal{D}$  be a linear manifold in the Hilbert space  $\mathcal{H}$ . Let  $\mathbf{T} : \mathcal{D} \rightarrow \mathcal{H}$  be a linear operator in  $\mathcal{H}$ . Suppose that for any sequence  $\{|u_n\}$  in  $\mathcal{D}$ , both  $\{|u_n\}$  and  $\{\mathbf{T}|u_n\}$  converge in  $\mathcal{H}$ , i.e.,

$$\lim |u_n\rangle = |u\rangle \quad \text{and} \quad \lim \mathbf{T}|u_n\rangle = |v\rangle.$$

closed operator We say that  $\mathbf{T}$  is closed if  $|v\rangle \in \mathcal{D}$  and  $\mathbf{T}|u\rangle = |v\rangle$ .

Notice that we cannot demand that  $|v\rangle$  be in  $\mathcal{D}$  for a general operator. This, as we saw in the preceding example, will not be appropriate for unbounded operators.

The restriction of the domain of an unbounded operator is necessitated by the fact that the action of the operator on a vector in the Hilbert space in general takes that vector out of the space. The following theorem (see [DeVi 90, pp. 251–252] for a proof) shows why this is necessary:

**18.1.4. Theorem.** A closed linear operator in  $\mathcal{H}$  that is defined at every point of  $\mathcal{H}$  (so that  $\mathcal{D} = \mathcal{H}$ ) is bounded.

Thus, if we are interested in unbounded operators (for instance, differential operators), we have to restrict their domains. In particular, we have to accept the possibility of an operator whose adjoint has a different domain.<sup>3</sup>

difference between hermitian and self-adjoint operators

**18.1.5. Definition.** Let  $\mathbf{T}$  be a linear operator in  $\mathcal{H}$ . We shall say that  $\mathbf{T}$  is hermitian if  $\mathbf{T}^\dagger$  is an extension of  $\mathbf{T}$ , i.e.,  $\mathcal{D}(\mathbf{T}) \subset \mathcal{D}(\mathbf{T}^\dagger)$  and  $\mathbf{T}^\dagger|u\rangle = \mathbf{T}|u\rangle$  for all  $|u\rangle \in \mathcal{D}(\mathbf{T})$ .  $\mathbf{T}$  is called self-adjoint if  $\mathcal{D}(\mathbf{T}) = \mathcal{D}(\mathbf{T}^\dagger)$ .

As we shall see shortly, certain types of Sturm–Liouville operators, although unbounded, lend themselves to a study within the context of compact operators.

operators with compact resolvent

**18.1.6. Definition.** A hermitian linear operator  $\mathbf{T}$  in a Hilbert space  $\mathcal{H}$  is said to have a compact resolvent if there is a  $\mu \in \rho(\mathbf{T})$  for which the resolvent  $\mathbf{R}_\mu(\mathbf{T})$  is compact.

<sup>3</sup>This subtle difference between hermitian and self-adjoint is stated here merely to warn the reader and will be confined to the present discussion. The two qualifiers will be (ab)used interchangeably in the rest of the book.

An immediate consequence of this definition is that  $\mathbf{R}_\lambda(\mathbf{T})$  is compact for all  $\lambda \in \rho(\mathbf{T})$ . To see this, note that  $\mathbf{R}_\lambda(\mathbf{T})$  is bounded by Definition 16.3.1. Now use Equation (16.9) and write

$$\mathbf{R}_\lambda(\mathbf{T}) = [1 + (\lambda - \mu)\mathbf{R}_\lambda(\mathbf{T})]\mathbf{R}_\mu(\mathbf{T}).$$

The RHS is a product of a bounded<sup>4</sup> and a compact operator, and therefore must be compact. The compactness of the resolvent characterizes its spectrum by Theorem 16.7.8. As the following theorem shows, this in turn characterizes the spectrum of the operators with compact resolvent.

**18.1.7. Theorem.** *Let  $\mathbf{T}$  be an operator with compact resolvent  $\mathbf{R}_\lambda(\mathbf{T})$  where  $\lambda \in \rho(\mathbf{T})$ . Then  $0 \neq \mu \in \rho(\mathbf{R}_\lambda(\mathbf{T}))$  if and only if  $(\lambda + 1/\mu) \in \rho(\mathbf{T})$ . Similarly,  $\mu \neq 0$  is an eigenvalue of  $\mathbf{R}_\lambda(\mathbf{T})$  if and only if  $(\lambda + 1/\mu)$  is an eigenvalue of  $\mathbf{T}$ . Furthermore, the eigenvectors of  $\mathbf{R}_\lambda(\mathbf{T})$  corresponding to  $\mu$  coincide with those of  $\mathbf{T}$  corresponding to  $(\lambda + 1/\mu)$ .*

*Proof.* The proof consists of a series of two-sided implications involving definitions. We give the proof of the first part, the second part being very similar:

$$\begin{aligned} \mu \in \rho(\mathbf{R}_\lambda(\mathbf{T})) & \quad \text{iff} & \quad \mathbf{R}_\lambda(\mathbf{T}) - \mu\mathbf{1} & \quad \text{is invertible.} \\ \mathbf{R}_\lambda(\mathbf{T}) - \mu\mathbf{1} & \quad \text{is invertible iff} & \quad (\mathbf{T} - \lambda\mathbf{1})^{-1} - \mu\mathbf{1} & \quad \text{is invertible.} \\ (\mathbf{T} - \lambda\mathbf{1})^{-1} - \mu\mathbf{1} & \quad \text{is invertible iff} & \quad \mathbf{1} - \mu(\mathbf{T} - \lambda\mathbf{1}) & \quad \text{is invertible.} \\ \mathbf{1} - \mu(\mathbf{T} - \lambda\mathbf{1}) & \quad \text{is invertible iff} & \quad \frac{1}{\mu}\mathbf{1} - \mathbf{T} + \lambda\mathbf{1} & \quad \text{is invertible.} \\ \left(\frac{1}{\mu} + \lambda\right)\mathbf{1} - \mathbf{T} & \quad \text{is invertible iff} & \quad \left(\frac{1}{\mu} + \lambda\right) \in \rho(\mathbf{T}). \end{aligned}$$

Comparing the LHS of the first line with the RHS of the last line, we obtain the first part of the theorem. □

A consequence of this theorem is that the eigenspaces of an (unbounded) operator with compact resolvent are finite-dimensional, i.e., such an operator has only finitely many eigenvectors corresponding to each of its eigenvalues. Moreover, arranging the eigenvalues  $\mu_n$  of the resolvent in decreasing order (as done in Theorem 16.7.8), we conclude that the eigenvalues of  $\mathbf{T}$  can be arranged in a sequence in increasing order of their absolute values and the limit of this sequence is infinity.

**18.1.8. Example.** Consider the operator  $\mathbf{T}$  in  $\mathcal{L}^2(0, 1)$  defined by<sup>5</sup>  $\mathbf{T}f = -f''$  having the domain  $\mathcal{D}(\mathbf{T}) = \{f \in \mathcal{L}^2(0, 1) \mid f'' \in \mathcal{L}^2(0, 1), f(0) = f(1) = 0\}$ . The reader may

<sup>4</sup>The sum of two bounded operators is bounded.

<sup>5</sup>We shall depart from our convention here and shall not use the Dirac bar-ket notation although the use of abstract operators encourages their use. The reason is that in this example, we are dealing with functions, and it is more convenient to undress the functions from their Dirac clothing.

check that zero is not an eigenvalue of  $T$ . Therefore, we may choose  $R_0(T) = T^{-1}$ . We shall study a systematic way of finding inverses of some specific differential operators in the upcoming chapters on Green's functions. At this point, suffice it to say that  $T^{-1}$  can be written as a Hilbert-Schmidt integral operator with kernel

$$K(x, t) = \begin{cases} x(1-t) & \text{if } 0 \leq x \leq t \leq 1, \\ (1-x)t & \text{if } 0 \leq t \leq x \leq 1. \end{cases}$$

Thus, if  $Tf = g$ , i.e., if  $f'' = -g$ , then  $T^{-1}g = f$ , or  $f = K[g]$ , i.e.,

$$f(x) = K[g](x) = \int_0^1 K(x, t)g(t) dt = \int_0^x (1-x)tg(t) dt + \int_x^1 (1-x)tg(t) dt.$$

It is readily verified that  $K[g](0) = K[g](1) = 0$  and  $f''(x) = K[g]''(x) = -g$ .

We can now use Theorem 18.1.7 with  $\lambda = 0$  to find all the eigenvalues of  $T$ :  $\mu_n$  is an eigenvalue of  $T$  if and only if  $1/\mu_n$  is an eigenvalue of  $T^{-1}$ . These eigenvalues should have finite-dimensional eigenspaces, and we should be able to arrange them in increasing order of magnitude without bound. To verify this, we solve  $f'' = -\mu f$ , whose solutions are  $\mu_n = n^2\pi^2$  and  $f_n(x) = \sin n\pi x$ . Note that there is only one eigenfunction corresponding to each eigenvalue. Therefore, the eigenspaces are finite- (one-) dimensional. ■

The example above is a special case of a large class of DOs occurring in mathematical physics. Recall from Theorem 13.5.4 that all linear second-order differential equations can be made self-adjoint. Moreover, Example 13.4.12 showed that *any* SOLDE can be transformed into a form in which the first-derivative term is absent. By dividing the DE by the coefficient of the second-derivative term if necessary, the study of the most general second-order linear differential operators boils down to that of the so-called **Sturm-Liouville (S-L) operators**

Sturm-Liouville operators

$$L_x \equiv \frac{d^2}{dx^2} - q(x), \tag{18.1}$$

which are assumed to be self-adjoint. Differential operators are necessarily accompanied by boundary conditions that specify their domains. So, to be complete, let us assume that the DO in Equation (18.1) acts on the subset of  $\mathcal{L}^2(a, b)$  consisting of functions  $u$  that satisfy the following so-called **separated boundary conditions**:

separated boundary conditions

$$\begin{aligned} \alpha_1 u(a) + \beta_1 u'(a) &= 0, \\ \alpha_2 u(b) + \beta_2 u'(b) &= 0, \end{aligned} \tag{18.2}$$

where  $\alpha_1, \alpha_2, \beta_1,$  and  $\beta_2$  are real constants with the property that the matrix of coefficients has no zero rows. The collection of the DO and the boundary conditions above is called a **regular Sturm-Liouville system**.

regular Sturm-Liouville systems

We now show that the DO of a regular Sturm-Liouville system has compact resolvent. First observe that by adding  $\alpha u$ —with  $\alpha$  an arbitrary number different from all eigenvalues of the DO—to both sides of the eigenvalue equation  $u'' -$



$qu = \lambda u$ , we can assume<sup>6</sup> that zero is not an eigenvalue of  $L_x$ . Next, suppose that  $u_1(x)$  and  $u_2(x)$  are the two linearly independent solutions of the homogeneous DE satisfying the first and the second boundary conditions of Equation (18.2), respectively. The operator whose kernel is

$$K(x, t) = \begin{cases} -u_1(x)u_2(t)/W(a) & \text{if } a \leq x \leq t \leq b, \\ -u_1(t)u_2(x)/W(a) & \text{if } a \leq t \leq x \leq b, \end{cases}$$

in which  $W$  is the Wronskian of the solutions, is a Hilbert–Schmidt operator and therefore compact. We now show that  $K(x, t)$  is the resolvent  $\mathbf{R}_0(L_x) = L_x^{-1} \equiv \mathbf{K}$  of our DO. To see this, write  $L_x u = v$ , and

$$u(x) = K[v](x) = -\frac{u_2(x)}{W(a)} \int_a^x u_1(t)v(t) dt - \frac{u_1(x)}{W(a)} \int_x^b u_2(t)v(t) dt.$$

Differentiating this once gives

$$u'(x) = -\frac{u_2'(x)}{W(a)} \int_a^x u_1(t)v(t) dt - \frac{u_1'(x)}{W(a)} \int_x^b u_2(t)v(t) dt,$$

and a second differentiation yields

$$u''(x) = -\frac{u_2''(x)}{W(a)} \int_a^x u_1(t)v(t) dt - \frac{u_1''(x)}{W(a)} \int_x^b u_2(t)v(t) dt + v(x).$$

The last equation follows from the fact that the Wronskian  $u_1' u_2 - u_2' u_1$  is constant for a DE of the form  $u'' - qu = 0$ . By substituting  $u_1'' = qu_1$  and  $u_2'' = qu_2$  in the last equation, we verify that  $u = K[v]$  is indeed a solution of the Sturm–Liouville system  $L_x u = v$ .

Next, we show that the eigensolutions of the S-L system are nondegenerate, i.e., the eigenspaces are one-dimensional. Suppose  $f_1$  and  $f_2$  are any two eigenfunctions corresponding to the same eigenvalue. Then both must satisfy the same DE and the same boundary conditions; in particular, we must have

$$\begin{aligned} \alpha_1 f_1(a) + \beta_1 f_1'(a) &= 0, \\ \alpha_1 f_2(a) + \beta_1 f_2'(a) &= 0 \end{aligned} \Rightarrow \begin{pmatrix} f_1(a) & f_1'(a) \\ f_2(a) & f_2'(a) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{18.3}$$

If  $\alpha_1$  and  $\beta_1$  are not both zero, the Wronskian—the determinant of the matrix above—must vanish. Therefore, the two functions must be linearly dependent. Finally, recall that a Hilbert space on which a compact operator  $\mathbf{K}$  is defined can be written as a direct sum of the latter’s eigenspaces. More specifically,  $\mathcal{H} = \sum_{j=0}^N \oplus \mathcal{M}_j$ , where each  $\mathcal{M}_j$  is finite-dimensional for  $j = 1, 2, \dots$ , and

---

<sup>6</sup>Although this will change  $q$ —and the original operator—no information will be lost because the eigenvectors will be the same and all eigenvalues will be changed by  $\alpha$ .

$N$  can be finite or infinite. If  $N$  is finite, then  $\mathcal{M}_0$ , which can be considered as the eigenspace of zero eigenvalue,<sup>7</sup> will be infinite-dimensional. If  $\mathcal{M}_0$  is finite-dimensional (or absent), then  $N$  must be infinite, and the eigenvectors of  $\mathbf{K}$  will span the entire space, i.e., they will form a complete orthogonal system. We now show that this holds for the regular Sturm–Liouville operator.

**Jacques Charles Francois Sturm** (1803–1855) made the first accurate determination of the velocity of sound in water in 1826, working with the Swiss engineer Daniel Colladon. He became a French citizen in 1833 and worked in Paris at the *École Polytechnique* where he became a professor in 1838. In 1840 he succeeded Poisson in the chair of mechanics in the *Faculté des Sciences*, Paris.



The problems of determining the eigenvalues and eigenfunctions of an ordinary differential equation with boundary conditions and of expanding a given function in terms of an infinite series of the eigenfunctions, which date from about 1750, became more prominent as new coordinate systems were introduced and new classes of functions arose as the eigenfunctions of ordinary differential equations. Sturm and his friend Joseph Liouville decided to tackle the general problem for any second-order linear differential equation.

Sturm had been working since 1833 on problems of partial differential equations, primarily on the flow of heat in a bar of variable density, and hence was fully aware of the eigenvalue and eigenfunction problem. The mathematical ideas he applied to this problem are closely related to his investigations of the reality and distribution of the roots of algebraic equations. His ideas on differential equations, he says, came from the study of difference equations and a passage to the limit. Liouville, informed by Sturm of the problems he was working on, took up the same subject. The results of their joint work was published in several papers which are quite detailed.

Suppose that the above Hilbert–Schmidt operator  $\mathbf{K}$  has a zero eigenvalue. Then, there must exist a nonzero function  $v$  such that  $K[v](x) = 0$ , i.e.,

$$-\frac{u_2(x)}{W(a)} \int_a^x u_1(t)v(t) dt - \frac{u_1(x)}{W(a)} \int_x^b u_2(t)v(t) dt = 0 \quad (18.4)$$

for all  $x$ . Differentiate this twice to get

$$-\frac{u_2''(x)}{W(a)} \int_a^x u_1(t)v(t) dt - \frac{u_1''(x)}{W(a)} \int_x^b u_2(t)v(t) dt + v(x) = 0.$$

Now substitute  $u_1'' = qu_1$  and  $u_2'' = qu_2$  in this equation and use Equation (18.4) to conclude that  $v = 0$ . This is impossible because no eigenvector can be zero.

<sup>7</sup>The reader recalls that when  $\mathbf{K}$  acts on  $\mathcal{M}_0$ , it yields zero.

Hence, zero is not an eigenvalue of  $\mathbf{K}$ , i.e.,  $\mathcal{M}_0 = \{0\}$ . Since eigenvectors of  $\mathbf{K} = \mathbf{L}_x^{-1}$  coincide with eigenvectors of  $\mathbf{L}_x$ , and eigenvalues of  $\mathbf{L}_x$  are the reciprocals of the eigenvalues of  $\mathbf{K}$ , we have the following result.

Theorem for regular Sturm-Liouville systems

**18.1.9. Theorem.** *A regular Sturm–Liouville system has a countable number of eigenvalues that can be arranged in an increasing sequence that has infinity as its limit. The eigenvectors of the Sturm–Liouville operator are nondegenerate and constitute a complete orthogonal set. Furthermore, the eigenfunction  $u_n(x)$  corresponding to the eigenvalue  $\lambda_n$  has exactly  $n$  zeros in its interval of definition.*

The last statement is not a result of operator theory, but can be derived using the theory of differential equations. We shall not present the details of its derivation. We need to emphasize that the boundary conditions are an integral part of S-L systems. Changing the boundary conditions so that, for example, they are no longer separated may destroy the regularity of the S-L system.

## 18.2 Sturm–Liouville Systems and SOLDES

We are now ready to combine our discussion of the preceding section with the knowledge gained from our study of differential equations. We saw in Chapter 12 that the separation of PDEs normally results in expressions of the form

$$\mathbf{L}[u] + \lambda u = 0, \quad \text{or} \quad p_2(x) \frac{d^2u}{dx^2} + p_1(x) \frac{du}{dx} + p_0(x)u + \lambda u = 0, \tag{18.5}$$

where  $u$  is a function of a single variable and  $\lambda$  is, a priori, an arbitrary constant. This is an eigenvalue equation for the operator  $\mathbf{L}$ , which is not, in general, self-adjoint. If we use Theorem 13.5.4 and multiply (18.5) by

$$w(x) = \frac{1}{p_2(x)} \exp \left[ \int^x \frac{p_1(t)}{p_2(t)} dt \right],$$

it becomes self-adjoint for real  $\lambda$ , and can be written as

$$\frac{d}{dx} \left[ p(x) \frac{du}{dx} \right] + [\lambda w(x) - q(x)]u = 0 \tag{18.6}$$

with  $p(x) = w(x)p_2(x)$  and  $q(x) = -p_0(x)w(x)$ . Equation (18.6) is the standard form of the S-L equation. However, it is not in the form studied in the previous section. To turn it into that form one changes both the independent and dependent variables via the so-called **Liouville substitution**:

Liouville substitution

$$u(x) = v(t)[p(x)w(x)]^{-1/4}, \quad t = \int_a^x \sqrt{\frac{w(s)}{p(s)}} ds. \tag{18.7}$$

It is then a matter of chain-rule differentiation to show that Equation (18.6) becomes

$$\frac{d^2v}{dt^2} + [\lambda - Q(t)]v = 0, \quad (18.8)$$

where

$$Q(t) = \frac{q(x(t))}{w(x(t))} + [p(x(t))w(x(t))]^{-1/4} \frac{d^2}{dt^2} [(pw)^{1/4}].$$

Therefore, Theorem 18.1.9 still holds.

**Joseph Liouville** (1809–1882) was a highly respected professor at the *Collège de France*, in Paris, and the founder and editor of the *Journal des Mathématiques Pures et Appliquées*, a famous periodical that played an important role in French mathematical life through the latter part of the nineteenth century. His own remarkable achievements as a creative mathematician have only recently received the appreciation they deserve.



He was the first to solve a boundary value problem by solving an equivalent integral equation. His ingenious theory of fractional differentiation answered the long-standing question of what reasonable meaning can be assigned to the symbol  $d^n y/dx^n$  when  $n$  is not a positive integer. He discovered the fundamental result in complex analysis that a bounded entire function is necessarily a constant and used it as the basis for his own theory of elliptic functions. There is also a well-known **Liouville theorem** in Hamiltonian mechanics, which states that volume integrals are time-invariant in phase space. In collaboration with Sturm, he also investigated the eigenvalue problem of second-order differential equations.

The theory of transcendental numbers is another branch of mathematics that originated in Liouville's work. The irrationality of  $\pi$  and  $e$  (the fact that they are not solutions of any linear equations) had been proved in the eighteenth century by Lambert and Euler. In 1844 Liouville showed that  $e$  is not a root of any quadratic equation with integral coefficients as well. This led him to conjecture that  $e$  is transcendental, which means that it does not satisfy any polynomial equation with integral coefficients.

**18.2.1. Example.** The Liouville substitution [Equation (18.7)] transforms the Bessel DE  $(xu')' + (k^2x - v^2/x)u = 0$  into

$$\frac{d^2v}{dt^2} + \left[ k^2 - \frac{v^2 - 1/4}{t^2} \right] v = 0,$$

from which we can obtain an interesting result when  $v = \frac{1}{2}$ . In that case we have  $\ddot{v} + k^2v = 0$ , whose solutions are of the form  $\cos kt$  and  $\sin kt$ . Noting that  $u(x) = J_{1/2}(x)$ , Equation (18.7) gives

$$J_{1/2}(kt) = A \frac{\sin kt}{\sqrt{t}} \quad \text{or} \quad J_{1/2}(kt) = B \frac{\cos kt}{\sqrt{t}},$$

and since  $J_{1/2}(x)$  is analytic at  $x = 0$ , we must have  $J_{1/2}(kt) = A \sin kt/\sqrt{t}$ , which is the result obtained in Chapter 14. ■

The appearance of  $w$  is the result of our desire to render the differential operator self-adjoint. It also appears in another context. Recall the Lagrange identity for a self-adjoint differential operator  $L$ :

$$uL[v] - vL[u] = \frac{d}{dx} \{p(x)[u(x)v'(x) - v(x)u'(x)]\}. \tag{18.9}$$

If we specialize this identity to the S-L equation of (18.6) with  $u = u_1$  corresponding to the eigenvalue  $\lambda_1$  and  $v = u_2$  corresponding to the eigenvalue  $\lambda_2$ , we obtain for the LHS

$$u_1L[u_2] - u_2L[u_1] = u_1(-\lambda_2 wu_2) + u_2(\lambda_1 wu_1) = (\lambda_1 - \lambda_2)wu_1u_2.$$

Integrating both sides of (18.9) then yields

$$(\lambda_1 - \lambda_2) \int_a^b wu_1u_2 dx = \{p(x)[u_1(x)u_2'(x) - u_2(x)u_1'(x)]\}_a^b. \tag{18.10}$$

A desired property of the solutions of a self-adjoint DE is their orthogonality when they belong to different eigenvalues. This property will be satisfied if we assume an inner product integral with weight function  $w(x)$ , and if the RHS of Equation (18.10) vanishes. There are various boundary conditions that fulfill the latter requirement. For example,  $u_1$  and  $u_2$  could satisfy the boundary conditions of Equation (18.2). Another set of appropriate boundary conditions (BC) is the **periodic BC** given by

periodic boundary conditions

$$u(a) = u(b) \quad \text{and} \quad u'(a) = u'(b). \tag{18.11}$$

However, as the following example shows, the latter BCs do not lead to a regular S-L system.

**18.2.2. Example.** (a) The S-L system consisting of the S-L equation  $d^2u/dt^2 + \omega^2u = 0$  in the interval  $[0, T]$  with the separated BCs  $u(0) = 0$  and  $u(T) = 0$  has the eigenfunctions  $u_n(t) = \sin \frac{n\pi}{T}t$  with  $n = 1, 2, \dots$  and the eigenvalues  $\lambda_n = \omega_n^2 = (n\pi/T)^2$  with  $n = 1, 2, \dots$

(b) Let the S-L equation be the same as in part (a) but change the interval to  $[-T, +T]$  and the BCs to a periodic one such as  $u(-T) = u(T)$  and  $u'(-T) = u'(T)$ . The eigenvalues are the same as before, but the eigenfunctions are  $1, \sin(n\pi t/T),$  and  $\cos(n\pi t/T),$  where  $n$  is a positive integer. Note that there is a degeneracy here in the sense that there are two linearly independent eigenfunctions having the same eigenvalue  $(n\pi/T)^2$ . By Theorem 18.1.9, the S-L system is not regular.

(c) The Bessel equation for a given fixed  $\nu^2$  is

$$u'' + \frac{1}{x}u' + \left(k^2 - \frac{\nu^2}{x^2}\right)u = 0, \quad \text{where } a \leq x \leq b,$$

and it can be turned into an S-L system if we multiply it by

$$w(x) = \frac{1}{p_2(x)} \exp \left[ \int^x \frac{p_1(t)}{p_2(t)} dt \right] = \exp \left[ \int^x \frac{dt}{t} \right] = x.$$

Then we can write

$$\frac{d}{dx} \left( x \frac{du}{dx} \right) + \left( k^2 x - \frac{v^2}{x} \right) u = 0,$$

which is in the form of Equation (18.6) with  $p = w = x$ ,  $\lambda = k^2$ , and  $q(x) = v^2/x$ . If  $a > 0$ , we can obtain a regular S-L system by applying appropriate separated BCs. ■

singular S-L systems

A regular S-L system is too restrictive for applications where either  $a$  or  $b$  or both may be infinite or where either  $a$  or  $b$  may be a singular point of the S-L equation. A **singular S-L system** is one for which one or more of the following conditions hold:

1. The interval  $[a, b]$  stretches to infinity in either or both directions.
2. Either  $p$  or  $w$  vanishes at one or both end points  $a$  and  $b$ .
3. The function  $q(x)$  is not continuous in  $[a, b]$ .
4. Any one of the functions  $p(x)$ ,  $q(x)$ , and  $w(x)$  is singular at  $a$  or  $b$ .

Even though the conclusions concerning eigenvalues of a regular S-L system cannot be generalized to the singular S-L system, the orthogonality of eigenfunctions corresponding to different eigenvalues can, as long as the eigenfunctions are square-integrable with weight function  $w(x)$ :

**18.2.3. Box.** *The eigenfunctions of a singular S-L system are orthogonal if the RHS of (18.10) vanishes.*

**18.2.4. Example.** Bessel functions  $J_\nu(x)$  are entire functions. Thus, they are square-integrable in the interval  $[0, b]$  for any finite positive  $b$ . For fixed  $\nu$  the DE

$$r^2 \frac{d^2 u}{dr^2} + r \frac{du}{dr} + (k^2 r^2 - \nu^2) u = 0 \tag{18.12}$$

transforms into the Bessel equation  $x^2 u'' + x u' + (x^2 - \nu^2) u = 0$  if we make the substitution  $kr = x$ . Thus, the solution of the *singular* S-L equation (18.12) that is analytic at  $r = 0$  and corresponds to the eigenvalue  $k^2$  is  $u_k(r) = J_\nu(kr)$ . For two different eigenvalues,  $k_1^2$  and  $k_2^2$ , the eigenfunctions are orthogonal if the boundary term of (18.10) corresponding to Equation (18.12) vanishes, that is, if

$$\{r[J_\nu(k_1 r)J'_\nu(k_2 r) - J_\nu(k_2 r)J'_\nu(k_1 r)]\}_0^b$$

vanishes, which will occur if and only if  $J_\nu(k_1b)J'_\nu(k_2b) - J_\nu(k_2b)J'_\nu(k_1b) = 0$ . A common choice is to take  $J_\nu(k_1b) = 0 = J_\nu(k_2b)$ , that is, to take both  $k_1b$  and  $k_2b$  as (different) roots of the Bessel function of order  $\nu$ . We thus have  $\int_0^b r J_\nu(k_i r) J_\nu(k_j r) dr = 0$  if  $k_i$  and  $k_j$  are different roots of  $J_\nu(kb) = 0$ . The Legendre equation

$$\frac{d}{dx} \left[ (1-x^2) \frac{du}{dx} \right] + \lambda u = 0, \quad \text{where } -1 < x < 1,$$

is already self-adjoint. Thus,  $w(x) = 1$ , and  $p(x) = 1 - x^2$ . The eigenfunctions of this singular S-L system [singular because  $p(1) = p(-1) = 0$ ] are regular at the end points  $x = \pm 1$  and are the Legendre polynomials  $P_n(x)$  corresponding to  $\lambda = n(n+1)$ . The boundary term of (18.10) clearly vanishes at  $a = -1$  and  $b = +1$ . Since  $P_n(x)$  are square-integrable on  $[-1, +1]$ , we obtain the familiar orthogonality relation:  $\int_{-1}^{+1} P_n(x) P_m(x) dx = 0$  if  $m \neq n$ .

The Hermite DE is

$$u'' - 2xu' + \lambda u = 0. \quad (18.13)$$

It is transformed into an S-L system if we multiply it by  $w(x) = e^{-x^2}$ . The resulting S-L equation is

$$\frac{d}{dx} \left[ e^{-x^2} \frac{du}{dx} \right] + \lambda e^{-x^2} u = 0. \quad (18.14)$$

The boundary term corresponding to the two eigenfunctions  $u_1(x)$  and  $u_2(x)$  having the respective eigenvalues  $\lambda_1$  and  $\lambda_2 \neq \lambda_1$  is

$$\{e^{-x^2} [u_1(x)u'_2(x) - u_2(x)u'_1(x)]\}_a^b.$$

This vanishes for arbitrary  $u_1$  and  $u_2$  (because they are Hermite polynomials) if  $a = -\infty$  and  $b = +\infty$ .

The function  $u$  is an eigenfunction of (18.14) corresponding to the eigenvalue  $\lambda$  if and only if it is a solution of (18.13). Solutions of this DE corresponding to  $\lambda = 2n$  are the Hermite polynomials  $H_n(x)$  discussed in Chapter 7. We can therefore write  $\int_{-\infty}^{+\infty} e^{-x^2} H_n(x) H_m(x) dx = 0$  if  $m \neq n$ . This orthogonality relation was also derived in Chapter 7. ■

## 18.3 Other Properties of Sturm–Liouville Systems

The S-L problem is central to the solution of many DEs in mathematical physics. In some cases the S-L equation has a direct bearing on the physics. For example, the eigenvalue  $\lambda$  may correspond to the orbital angular momentum of an electron in an atom (see the treatment of spherical harmonics in Chapter 12) or to the energy levels of a particle in a potential (see Example 14.5.2). In many cases, then, it is worthwhile to gain some knowledge of the behavior of an S-L system in the limit of large  $\lambda$ —high angular momentum or high energy. Similarly, it is useful to understand the behavior of the solutions for large values of their arguments. We therefore devote this section to a discussion of the behavior of solutions of an S-L system in the limit of large eigenvalues and large independent variable.

### 18.3.1 Asymptotic Behavior for Large Eigenvalues

We assume that the S-L operator has the form given in Equation (18.1). This can always be done for an arbitrary second-order linear DE by multiplying it by a proper function (to make it self-adjoint) followed by a Liouville substitution. So, consider an S-L systems of the following form:

$$u'' + [\lambda - q(x)]u \equiv u'' + Q(x)u = 0 \quad \text{where } Q = \lambda - q \quad (18.15)$$

with separated BCs of (18.2). Let us assume that  $Q(x) > 0$  for all  $x \in [a, b]$ , that is,  $\lambda > q(x)$ . This is reasonable, since we are interested in very large  $\lambda$ .

The study of the system of (18.15) and (18.2) is simplified if we make the

Prüfer substitution **Prüfer substitution:**

$$u = RQ^{-1/4} \sin \phi, \quad u' = RQ^{1/4} \cos \phi, \quad (18.16)$$

where  $R(x, \lambda)$  and  $\phi(x, \lambda)$  are  $\lambda$ -dependent functions of  $x$ . This substitution transforms the S-L equation of (18.15) into a pair of equations (see Problem 18.3):

$$\begin{aligned} \frac{d\phi}{dx} &= \sqrt{\lambda - q(x)} - \frac{q'}{4[\lambda - q(x)]} \sin 2\phi, \\ \frac{dR}{dx} &= \frac{Rq'}{4[\lambda - q(x)]} \cos 2\phi. \end{aligned} \quad (18.17)$$

The function  $R(x, \lambda)$  is assumed to be positive because any negativity of  $u$  can be transferred to the phase  $\phi(x, \lambda)$ . Also,  $R$  cannot be zero at any point of  $[a, b]$ , because both  $u$  and  $u'$  would vanish at that point, and, by Lemma 13.3.3,  $u(x) = 0$ . Equation (18.17) is very useful in discussing the asymptotic behavior of solutions of S-L systems both when  $\lambda \rightarrow \infty$  and when  $x \rightarrow \infty$ . Before we discuss such asymptotics, we need to make a digression.

It is often useful to have a notation for the behavior of a function  $f(x, \lambda)$  for large  $\lambda$  and all values of  $x$ . If the function remains bounded for all values of  $x$  as  $\lambda \rightarrow \infty$ , we write  $f(x, \lambda) = O(1)$ . Intuitively, this means that as  $\lambda$  gets larger and larger, the magnitude of the function  $f(x, \lambda)$  remains of order 1. In other words, for no value of  $x$  is  $\lim_{\lambda \rightarrow \infty} f(x, \lambda)$  infinite. If  $\lambda^n f(x, \lambda) = O(1)$ , then we can write  $f(x, \lambda) = O(1)/\lambda^n$ . This means that as  $\lambda$  tends to infinity,  $f(x, \lambda)$  goes to zero as fast as  $1/\lambda^n$  does. Sometimes this is written as  $f(x, \lambda) = O(\lambda^{-n})$ . Some properties of  $O(1)$  are as follows:

1. If  $a$  is a finite real number, then  $O(1) + a = O(1)$ .
2.  $O(1) + O(1) = O(1)$ , and  $O(1)O(1) = O(1)$ .
3. For finite  $a$  and  $b$ ,  $\int_a^b O(1) dx = O(1)$ .
4. If  $r$  and  $s$  are real numbers with  $r \leq s$ , then

$$O(1)\lambda^r + O(1)\lambda^s = O(1)\lambda^s.$$



5. If  $g(x)$  is any bounded function of  $x$ , then a Taylor series expansion yields

$$\begin{aligned} [\lambda + g(x)]^r &= \lambda^r \left[ 1 + \frac{g(x)}{\lambda} \right]^r \\ &= \lambda^r \left\{ 1 + r \frac{g(x)}{\lambda} + \frac{r(r-1)}{2} \left[ \frac{g(x)}{\lambda} \right]^2 + \frac{O(1)}{\lambda^3} \right\} \\ &= \lambda^r + r g(x) \lambda^{r-1} + O(1) \lambda^{r-2} = \lambda^r + O(1) \lambda^{r-1} \\ &= O(1) \lambda^r. \end{aligned}$$

Returning to Equation (18.17) and expanding its RHSs using property 5, we obtain

$$\frac{d\phi}{dx} = \sqrt{\lambda} + \frac{O(1)}{\sqrt{\lambda}} + \frac{O(1)}{\lambda} = \sqrt{\lambda} + \frac{O(1)}{\sqrt{\lambda}} \quad \frac{dR}{dx} = \frac{O(1)}{\lambda}.$$

Taylor series expansion of  $\phi(x, \lambda)$  and  $R(x, \lambda)$  about  $x = a$  then yields

$$\begin{aligned} \phi(x, \lambda) &= \phi(a, \lambda) + (x - a) \sqrt{\lambda} + \frac{O(1)}{\sqrt{\lambda}}, \\ R(x, \lambda) &= R(a, \lambda) + \frac{O(1)}{\lambda} \end{aligned} \tag{18.18}$$

for  $\lambda \rightarrow \infty$ . These results are useful in determining the behavior of  $\lambda_n$  for large  $n$ . As an example, we use (18.2) and (18.16) to write

$$-\frac{\alpha_1}{\beta_1} = \frac{u'(a)}{u(a)} = \frac{R(a, \lambda) Q^{1/4}(a, \lambda) \cos[\phi(a, \lambda)]}{R(a, \lambda) Q^{-1/4}(a, \lambda) \sin[\phi(a, \lambda)]} = Q^{1/2}(a, \lambda) \cot[\phi(a, \lambda)],$$

where we have assumed that  $\beta_1 \neq 0$ . If  $\beta_1 = 0$ , we can take the ratio  $\beta_1/\alpha_1$ , which is finite because at least one of the two constants must be different from zero. Let  $A = -\alpha_1/\beta_1$  and write  $\cot[\phi(a, \lambda)] = A/\sqrt{\lambda - q(a)}$ . Similarly,  $\cot[\phi(b, \lambda)] = B/\sqrt{\lambda - q(b)}$ , where  $B = -\alpha_2/\beta_2$ . Let us concentrate on the  $n$ th eigenvalue and write

$$\phi(a, \lambda_n) = \cot^{-1} \frac{A}{\sqrt{\lambda_n - q(a)}}, \quad \phi(b, \lambda_n) = \cot^{-1} \frac{A}{\sqrt{\lambda_n - q(b)}}.$$

For large  $\lambda_n$  the argument of  $\cot^{-1}$  is small. Therefore, we can expand the RHS in a Taylor series about zero:

$$\cot^{-1} \epsilon = \cot^{-1}(0) - \epsilon + \dots = \frac{\pi}{2} - \epsilon + \dots = \frac{\pi}{2} + \frac{O(1)}{\sqrt{\lambda_n}}$$

for  $\epsilon = O(1)/\sqrt{\lambda_n}$ . It follows that

$$\phi(a, \lambda_n) = \frac{\pi}{2} + \frac{O(1)}{\sqrt{\lambda_n}}, \quad \phi(b, \lambda_n) = \frac{\pi}{2} + n\pi + \frac{O(1)}{\sqrt{\lambda_n}}. \tag{18.19}$$

The term  $n\pi$  appears in (18.19) because, by Theorem 18.1.9, the  $n$ th eigenfunction has  $n$  zeros between  $a$  and  $b$ . Since  $u = RQ^{-1/4} \sin \phi$ , this means that  $\sin \phi$  must go through  $n$  zeros as  $x$  goes from  $a$  to  $b$ . Thus, at  $x = b$  the phase  $\phi$  must be  $n\pi$  larger than at  $x = a$ .

Substituting  $x = b$  in the first equation of (18.18), with  $\lambda \rightarrow \lambda_n$ , and using (18.19), we obtain

$$\frac{\pi}{2} + n\pi + \frac{O(1)}{\sqrt{\lambda_n}} = \frac{\pi}{2} + \frac{O(1)}{\sqrt{\lambda_n}} + (b-a)\sqrt{\lambda_n} + \frac{O(1)}{\sqrt{\lambda_n}},$$

or

$$(b-a)\sqrt{\lambda_n} = n\pi + \frac{O(1)}{\sqrt{\lambda_n}}. \tag{18.20}$$

One consequence of this result is that,  $\lim_{n \rightarrow \infty} n\lambda_n^{-1/2} = (b-a)/\pi$ . Thus,  $\sqrt{\lambda_n} = C_n n$ , where  $\lim_{n \rightarrow \infty} C_n = \pi/(b-a)$ , and Equation (18.20) can be rewritten as

$$\sqrt{\lambda_n} = \frac{n\pi}{b-a} + \frac{O(1)}{C_n n} = \frac{n\pi}{b-a} + \frac{O(1)}{n}. \tag{18.21}$$

This equation describes the asymptotic behavior of eigenvalues. The following theorem, stated without proof, describes the asymptotic behavior of eigenfunctions.

**18.3.1. Theorem.** Let  $\{u_n(x)\}_{n=0}^\infty$  be the normalized eigenfunctions of the regular  $S$ - $L$  system given by Equations (18.15) and (18.2) with  $\beta_1\beta_2 \neq 0$ . Then, for  $n \rightarrow \infty$ ,

asymptotic behavior  
of solutions of large  
order

$$u_n(x) = \sqrt{\frac{2}{b-a}} \cos \frac{n\pi(x-a)}{b-a} + \frac{O(1)}{n}.$$

**18.3.2. Example.** Let us derive an asymptotic formula for the Legendre polynomials  $P_n(x)$ . We first make the Liouville substitution to transform the Legendre DE  $[(1-x^2)P']' + n(n+1)P_n = 0$  into

$$\frac{d^2v}{dt^2} + [\lambda_n - Q(t)]v = 0, \quad \text{where } \lambda_n = n(n+1). \tag{18.22}$$

Here  $p(x) = 1 - x^2$  and  $w(x) = 1$ , so  $t = \int^x ds/\sqrt{1-s^2} = \cos^{-1} x$ , or  $x(t) = \cos t$ , and

$$P_n(x(t)) = v(t)[1-x^2(t)]^{-1/4} = \frac{v(t)}{\sqrt{\sin t}}. \tag{18.23}$$

In Equation (18.22)

$$\begin{aligned} Q(t) &= (1-x^2)^{-1/4} \frac{d^2}{dt^2} [(1-x^2)^{1/4}] \\ &= \frac{1}{\sqrt{\sin t}} \frac{d^2}{dt^2} [\sqrt{\sin t}] = -\frac{1}{4} \left( 1 + \frac{1}{\sin^2 t} \right). \end{aligned}$$

For large  $n$  we can neglect  $Q(t)$ , make the approximation  $\lambda_n \approx (n + \frac{1}{2})^2$ , and write  $\ddot{v} + (n + \frac{1}{2})^2 v = 0$ , whose general solution is

$$v(t) = A \cos[(n + \frac{1}{2})t + \alpha],$$

where  $A$  and  $\alpha$  are arbitrary constants. Substituting this solution in (18.23) yields  $P_n(\cos t) = A \cos[(n + \frac{1}{2})t + \alpha] / \sqrt{\sin t}$ . To determine  $\alpha$  we note that  $P_n(0) = 0$  if  $n$  is odd. Thus, if we let  $t = \pi/2$ , the cosine term vanishes for odd  $n$  if  $\alpha = -\pi/4$ . Thus, the general asymptotic formula for Legendre polynomials is

$$P_n(\cos t) = \frac{A}{\sqrt{\sin t}} \cos \left[ \left( n + \frac{1}{2} \right) t - \frac{\pi}{4} \right]. \quad \blacksquare$$

### 18.3.2 Asymptotic Behavior for Large $x$

Liouville and Prüfer substitutions are useful in investigating the behavior of the solutions of S-L systems for large  $x$  as well. The general procedure is to transform the DE into the form of Equation (18.8) by the Liouville substitution; then make the Prüfer substitution of (18.16) to obtain two DEs in the form of (18.17). Solving Equation (18.17) when  $x \rightarrow \infty$  determines the behavior of  $\phi$  and  $R$  and, subsequently, of  $u$ , the solution. Problem 18.4 illustrates this procedure for the Bessel functions. We simply quote the results:

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}} \cos \left[ x - \left( \nu + \frac{1}{2} \right) \frac{\pi}{2} + \frac{\nu^2 - 1/4}{2x} \right] + \frac{O(1)}{x^{5/2}},$$

$$Y_\nu(x) = \sqrt{\frac{2}{\pi x}} \sin \left[ x - \left( \nu + \frac{1}{2} \right) \frac{\pi}{2} + \frac{\nu^2 - 1/4}{2x} \right] + \frac{O(1)}{x^{5/2}}.$$

These two relations easily yield the asymptotic expressions for the Hankel functions:

$$H_\nu^{(1)}(x) = J_\nu(x) + iY_\nu(x)$$

$$= \sqrt{\frac{2}{\pi x}} \exp \left\{ i \left[ x - \left( \nu + \frac{1}{2} \right) \frac{\pi}{2} + \frac{\nu^2 - 1/4}{2x} \right] \right\} + \frac{O(1)}{x^{5/2}},$$

$$H_\nu^{(2)}(x) = J_\nu(x) - iY_\nu(x)$$

$$= \sqrt{\frac{2}{\pi x}} \exp \left\{ -i \left[ x - \left( \nu + \frac{1}{2} \right) \frac{\pi}{2} + \frac{\nu^2 - 1/4}{2x} \right] \right\} + \frac{O(1)}{x^{5/2}}.$$

If the last term in the exponent—which vanishes as  $x \rightarrow \infty$ —is ignored, the asymptotic expression for  $H_\nu^{(1)}(x)$  matches what was obtained in Chapter 15 using the method of steepest descent.

## 18.4 Problems

**18.1.** Show that the Liouville substitution transforms regular S-L systems into regular S-L systems and separated and periodic BCs into separated and periodic BCs, respectively.

**18.2.** Let  $u_1(x)$  and  $u_2(x)$  be transformed, respectively into  $v_1(t)$  and  $v_2(t)$  by the Liouville substitution. Show that the inner product on  $[a, b]$  with weight function  $w(x)$  is transformed into the inner product on  $[0, c]$  with unit weight, where  $c = \int_a^b \sqrt{w/p} dx$ .

**18.3.** Derive Equation (18.17) from (18.15) using Prüfer substitution.

**18.4.** (a) Show that the Liouville substitution transforms the Bessel DE into

$$\frac{d^2v}{dt^2} + \left( k^2 - \frac{v^2 - 1/4}{t^2} \right) v = 0.$$

(b) Find the equations obtained from the Prüfer substitution, and show that for large  $x$  these equations reduce to

$$\phi' = k \left( 1 - \frac{a}{2k^2x^2} \right) + \frac{O(1)}{x^3} \quad \frac{R'}{R} = \frac{O(1)}{x^3},$$

where  $a = v^2 - \frac{1}{4}$ .

(c) Integrate these equations from  $x$  to  $b > x$  and take the limit as  $b \rightarrow \infty$  to get

$$\phi(x) = \phi_\infty + kx + \frac{a}{2kx} + \frac{O(1)}{x^2}, \quad R(x) = R_\infty + \frac{O(1)}{x^2},$$

where  $\phi_\infty = \lim_{b \rightarrow \infty} (\phi(b) - kb)$  and  $R_\infty = \lim_{b \rightarrow \infty} R(b)$ .

(d) Substitute these and the appropriate expression for  $Q^{-1/4}$  in Equation (18.16) and show that

$$v(x) = \frac{R_\infty}{\sqrt{k}} \cos \left( kx - kx_\infty + \frac{v^2 - 1/4}{2kx} \right) + \frac{O(1)}{x^2},$$

where  $kx_\infty \equiv \pi/2 - \phi_\infty$ .

(e) Choose  $R_\infty = \sqrt{2/\pi}$  for all solutions of the Bessel DE, and let

$$kx_\infty = \left( \nu + \frac{1}{2} \right) \frac{\pi}{2} \quad \text{and} \quad kx_\infty = \left( \nu + \frac{3}{2} \right) \frac{\pi}{2}$$

for the Bessel functions  $J_\nu(x)$  and the Neumann functions  $Y_\nu(x)$ , respectively, and find the asymptotic behavior of these two functions.

## Additional Reading

1. Birkhoff, G. and Rota, G.-C. *Ordinary Differential Equations*, 3rd ed., Wiley, 1978. Has a good discussion of Sturm–Liouville differential equations and their asymptotic behavior.
2. Boccara, N. *Functional Analysis*, Academic Press, 1990. Discusses the Sturm–Liouville operators in the same spirit as presented in this chapter.
3. Hellwig, G. *Differential Operators of Mathematical Physics*, Addison-Wesley, 1967. An oldie, but goodie! It gives a readable account of the Sturm–Liouville systems.

# Sturm–Liouville Systems: Examples

Chapter 12 showed how the solution of many PDEs can be written as the product of the solutions of the separated ODEs. These DEs are usually of Sturm-Liouville type. We saw this in the construction of spherical harmonics. In this chapter, consisting mainly of illustrative examples, we shall consider the use of other coordinate systems and construct solutions to DEs as infinite series expansions in terms of S-L eigenfunctions.

## 19.1 Expansions in Terms of Eigenfunctions

Central to the expansion of solutions in terms of S-L eigenfunctions is the question of their completeness. This completeness was established for a *regular* S-L system in Theorem 18.1.9.

We shall shortly state an analogous theorem (without proof) that establishes the completeness of the eigenfunctions of more general S-L systems. This theorem requires the following generalization of the separated and the periodic BCs:

$$\begin{aligned} \mathbf{R}_1 u &\equiv \alpha_{11}u(a) + \alpha_{12}u'(a) + \alpha_{13}u(b) + \alpha_{14}u'(b) = 0, \\ \mathbf{R}_2 u &\equiv \alpha_{21}u(a) + \alpha_{22}u'(a) + \alpha_{23}u(b) + \alpha_{24}u'(b) = 0, \end{aligned} \quad (19.1)$$

where  $\alpha_{ij}$  are numbers such that the rank of the following matrix is 2:

$$\mathbf{a} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \end{pmatrix}.$$

The separated BCs correspond to the case for which  $\alpha_{11} = \alpha_1, \alpha_{12} = \beta_1, \alpha_{23} = \alpha_2,$  and  $\alpha_{24} = \beta_2,$  with all other  $\alpha_{ij}$  zero. Similarly, the periodic BC is a special case for which  $\alpha_{11} = -\alpha_{13} = \alpha_{22} = -\alpha_{24} = 1,$  with all other  $\alpha_{ij}$  zero. It is easy to

verify that the rank of the matrix  $a$  is 2 for these two special cases. Let

$$\mathcal{U} = \{u \in \mathcal{C}^2[a, b] \mid \mathbf{R}_j u = 0, \quad \text{for } j = 1, 2\} \quad (19.2)$$

be a subspace of  $\mathcal{L}_w^2(a, b)$ , and—to assure the vanishing of the RHS of the Lagrange identity—assume that the following equality holds:

$$p(b) \det \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} = p(a) \det \begin{pmatrix} \alpha_{13} & \alpha_{14} \\ \alpha_{23} & \alpha_{24} \end{pmatrix}. \quad (19.3)$$

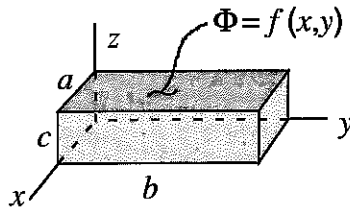
We are now ready to consider the theorem (for a proof, see [Hell 67, Chapter 7]).

**19.1.1. Theorem.** *The eigenfunctions  $\{u_n(x)\}_{n=1}^{\infty}$  of an S-L system consisting of the S-L equation  $(pu')' + (\lambda w - q)u = 0$  and the BCs of (19.1) form a complete basis of the subspace  $\mathcal{U}$  of  $\mathcal{L}_w^2(a, b)$  described in (19.2). The eigenvalues are real and countably infinite and each one has a multiplicity of at most 2. They can be ordered according to size  $\lambda_1 \leq \lambda_2 \leq \dots$ , and their only limit point is  $+\infty$ .*

First note that Equation (19.3) contains both separated and periodic BCs as special cases (Problem 19.1). In the case of periodic BCs, we assume that  $p(a) = p(b)$ . Thus, all the eigenfunctions discussed so far are covered by Theorem 19.1.1. Second, the orthogonality of eigenfunctions corresponding to different eigenvalues and the fact that there are infinitely many distinct eigenvalues assure the existence of infinitely many eigenfunctions. Third, the eigenfunctions form a basis of  $\mathcal{U}$  and not the whole  $\mathcal{L}_w^2(a, b)$ . Only those functions  $u \in \mathcal{L}_w^2(a, b)$  that satisfy the BC in (19.1) are expandable in terms of  $u_n(x)$ . Finally, the last statement of Theorem 19.1.1 is a repetition of part of Theorem 18.1.9 but is included because the conditions under which Theorem 19.1.1 holds are more general than those applying to Theorem 18.1.9.

Part II discussed orthogonal functions in detail and showed how other functions can be expanded in terms of them. However, the procedure used in Part II was ad hoc from a logical standpoint. After all, the orthogonal polynomials were invented by nineteenth-century mathematical physicists who, in their struggle to solve the PDEs of physics using the separation of variables, came across various ODEs of the second order, all of which were recognized later as S-L systems. From a logical standpoint, therefore, this chapter should precede Part II. But the order of the chapters was based on clarity and ease of presentation and the fact that the machinery of differential equations was a prerequisite for such a discussion.

Theorem 19.1.1 is the important link between the algebraic and the analytic machinery of differential equation theory. This theorem puts at our disposal concrete mathematical functions that are calculable to any desired accuracy (on a computer, say) and can serve as basis functions for all the expansions described in Part II. The remainder of this chapter is devoted to solving some PDEs of mathematical physics using the separation of variables and Theorem 19.1.1.



**Figure 19.1** A rectangular conducting box of which one face is held at the potential  $f(x, y)$  and the other faces are grounded.

## 19.2 Separation in Cartesian Coordinates

Problems most suitable for Cartesian coordinates have boundaries with rectangular symmetry such as boxes or planes.

rectangular  
conducting box

### 19.2.1. Example. RECTANGULAR CONDUCTING BOX

Consider a rectangular conducting box with sides  $a$ ,  $b$ , and  $c$  (see Figure 19.1). All faces are held at zero potential except the top face, whose potential is given by a function  $f(x, y)$ . Let us find the potential at all points inside the box.

The relevant PDE for this situation is Laplace's equation,  $\nabla^2 \Phi = 0$ . Writing  $\Phi(x, y, z)$  as a product of three functions,  $\Phi(x, y, z) = X(x)Y(y)Z(z)$ , yields three ODEs (see Problem 19.2):

$$\frac{d^2 X}{dx^2} + \lambda X = 0, \quad \frac{d^2 Y}{dy^2} + \mu Y = 0, \quad \frac{d^2 Z}{dz^2} + \nu Z = 0, \quad (19.4)$$

where  $\lambda + \mu + \nu = 0$ . The vanishing of  $\Phi$  at  $x = 0$  and  $x = a$  means that

$$\Phi(0, y, z) = X(0)Y(y)Z(z) = 0 \quad \forall y, z \Rightarrow X(0) = 0,$$

$$\Phi(a, y, z) = X(a)Y(y)Z(z) = 0 \quad \forall y, z \Rightarrow X(a) = 0.$$

We thus obtain an S-L system,  $X'' + \lambda X = 0$ ,  $X(0) = 0 = X(a)$ , whose BC is neither separated nor periodic, but satisfies (19.1) with  $\alpha_{11} = \alpha_{23} = 1$  and all other  $\alpha_{ij}$  zero. This S-L system has the eigenvalues and eigenfunctions

$$\lambda_n = \left(\frac{n\pi}{a}\right)^2 \quad \text{and} \quad X_n(x) = \sin\left(\frac{n\pi}{a}x\right) \quad \text{for } n = 1, 2, \dots$$

Similarly, the second equation in (19.4) leads to

$$\mu_m = \left(\frac{m\pi}{b}\right)^2 \quad \text{and} \quad Y_m(y) = \sin\left(\frac{m\pi}{b}y\right) \quad \text{for } m = 1, 2, \dots$$

On the other hand, the third equation in (19.4) does not lead to an S-L system because the BC for the top of the box does not fit (19.1). This is as expected because the "eigenvalue"  $\nu$  is already determined by  $\lambda$  and  $\mu$ . Nevertheless, we can find a solution for that equation. The substitution

$$\gamma_{mn}^2 = \left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2$$



changes the  $Z$  equation to  $Z'' - \gamma_{mn}^2 Z = 0$ , whose solution, consistent with  $Z(0) = 0$ , is  $Z(z) = C_{mn} \sinh(\gamma_{mn} z)$ .

We note that  $X(x)$  and  $Y(y)$  are functions satisfying  $\mathbf{R}_1 X = 0 = \mathbf{R}_2 X$ . Thus, by Theorem 19.1.1, they can be written as a linear combination of  $X_n(x)$  and  $Y_m(y)$ :  $X(x) = \sum_{n=1}^{\infty} A_n \sin(n\pi x/a)$  and  $Y(y) = \sum_{m=1}^{\infty} B_m \sin(m\pi b/y)$ . Consequently, the most general solution can be expressed as

$$\Phi(x, y, z) = X(x)Y(y)Z(z) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{mn} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right) \sinh(\gamma_{mn}z),$$

where  $A_{mn} = A_n B_m C_{mn}$ .

To specify  $\Phi$  completely, we must determine the arbitrary constants  $A_{mn}$ . This is done by imposing the remaining BC,  $\Phi(x, y, c) = f(x, y)$ , yielding the identity

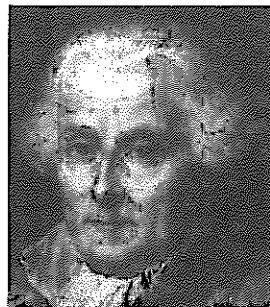
$$\begin{aligned} f(x, y) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{mn} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right) \sinh(\gamma_{mn}c) \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} B_{mn} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right), \end{aligned}$$

where  $B_{mn} \equiv A_{mn} \sinh(\gamma_{mn}c)$ . This is a two-dimensional Fourier series (see Chapter 8) whose coefficients are given by

$$B_{mn} = \frac{4}{ab} \int_0^a dx \int_0^b dy f(x, y) \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right). \quad \blacksquare$$

**Pierre Simon de Laplace** (1749–1827) was a French mathematician and theoretical astronomer who was so famous in his own time that he was known as the Newton of France. His main interests throughout his life were celestial mechanics, the theory of probability, and personal advancement.

At the age of 24 he was already deeply engaged in the detailed application of Newton's law of gravitation to the solar system as a whole, in which the planets and their satellites are not governed by the sun alone, but interact with one another in a bewildering variety of ways. Even Newton had been of the opinion that divine intervention would occasionally be needed to prevent this complex mechanism from degenerating into chaos. Laplace decided to seek reassurance elsewhere, and succeeded in proving that the ideal solar system of mathematics is a stable dynamical system that will endure unchanged for all time. This achievement was only one of the long series of triumphs recorded in his monumental treatise *Mécanique Céleste* (published in five volumes from 1799 to 1825), which summed up the work on gravitation of several generations of illustrious mathematicians. Unfortunately for his later reputation, he omitted all reference to the discoveries of his predecessors and contemporaries, and left it to be inferred that the ideas were entirely his own. Many anecdotes are associated with this work. One of the best known describes the occasion on which Napoleon tried to get a



rise out of Laplace by protesting that he had written a huge book on the system of the world without once mentioning God as the author of the universe. Laplace is supposed to have replied, "Sire, I had no need of that hypothesis." The principal legacy of the *Mécanique Céleste* to later generations lay in Laplace's wholesale development of potential theory, with its far-reaching implications for a dozen different branches of physical science ranging from gravitation and fluid mechanics to electromagnetism and atomic physics. Even though he lifted the idea of the potential from Lagrange without acknowledgment, he exploited it so extensively that ever since his time the fundamental equation of potential theory has been known as Laplace's equation. After the French Revolution, Laplace's political talents and greed for position came to full flower. His compatriots speak ironically of his "suppleness" and "versatility" as a politician. What this really means is that each time there was a change of regime (and there were many), Laplace smoothly adapted himself by changing his principles—back and forth between fervent republicanism and fawning royalism—and each time he emerged with a better job and grander titles. He has been aptly compared with the apocryphal Vicar of Bray in English literature, who was twice a Catholic and twice a Protestant. The Vicar is said to have replied as follows to the charge of being a turncoat: "Not so, neither, for if I changed my religion, I am sure I kept true to my principle, which is to live and die the Vicar of Bray."

To balance his faults, Laplace was always generous in giving assistance and encouragement to younger scientists. From time to time he helped forward in their careers such men as the chemist Gay-Lussac, the traveler and naturalist Humboldt, the physicist Poisson, and—appropriately—the young Cauchy, who was destined to become one of the chief architects of nineteenth century mathematics.

Laplace's equation describes not only electrostatics, but also heat transfer. When the transfer (diffusion) of heat takes place with the temperature being independent of time, the process is known as **steady-state heat transfer**. The diffusion equation,  $\partial T/\partial t = a^2 \nabla^2 T$ , becomes Laplace's equation,  $\nabla^2 T = 0$ , and the technique of the preceding example can be used. It is easy to see that the diffusion equation allows us to perform any linear transformation on  $T$ , such as  $T \rightarrow \alpha T + \beta$ , and still satisfy that equation. This implies that  $T$  can be measured in any scale such as Kelvin, Celsius, and Fahrenheit.

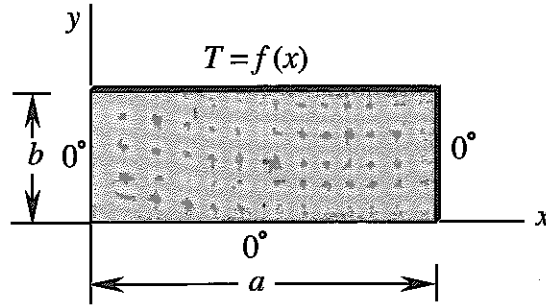
steady-state  
heat-conducting  
plate  
heat-conducting  
plate: steady state

### 19.2.2. Example. STEADY-STATE HEAT-CONDUCTING PLATE

Let us consider a rectangular heat-conducting plate with sides of lengths  $a$  and  $b$ . Three of the sides are held at  $T = 0$ , and the fourth side has a temperature variation  $T = f(x)$  (see Figure 19.2). The flat faces are insulated, so they cannot lose heat to the surroundings. Assuming a steady-state heat transfer, let us calculate the variation of  $T$  over the plate. The problem is two-dimensional. The separation of variables leads to

$$\frac{d^2 X}{dx^2} + \lambda X = 0, \quad \frac{d^2 Y}{dy^2} + \mu Y = 0, \quad \text{where } \lambda + \mu = 0. \quad (19.5)$$

The  $X$  equation and the BCs  $T(0, y) = T(a, y) = 0$  form an S-L system whose eigenvalues and eigenfunctions are  $\lambda_n = (n\pi/a)^2$  and  $X_n(x) = \sin(n\pi x/a)$  for  $n = 1, 2, \dots$ . Thus, according to Theorem 19.1.1, a general  $X(x)$  can be written as  $X(x) = \sum_{n=1}^{\infty} A_n \sin(n\pi x/a)$ .



**Figure 19.2** A heat-conducting rectangular plate.

The  $Y$  equation, on the other hand, does not form an S-L system due to the fact that its “eigenvalue” is predetermined by the third equation in (19.5). Nevertheless, we can solve the equation  $Y'' - (n\pi/a)^2 Y = 0$  to obtain the general solution  $Y(y) = Ae^{n\pi y/a} + Be^{-n\pi y/a}$ . Since  $T(x, 0) = 0 \forall x$ , we must have  $Y(0) = 0$ . This implies that  $A+B = 0$ , which, in turn, reduces the solution to  $Y = A \sinh(n\pi y/a)$ . Thus, the most general solution, consistent with the three BCs  $T(0, y) = T(a, y) = T(x, 0) = 0$ , is

$$T(x, y) = X(x)Y(y) = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi}{a}x\right) \sinh\left(\frac{n\pi}{a}y\right).$$

The fourth BC gives a Fourier series,

$$f(x) = \sum_{n=1}^{\infty} \left[ B_n \sinh\left(\frac{n\pi}{a}b\right) \right] \sin\left(\frac{n\pi}{a}x\right) \equiv \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi}{a}x\right),$$

whose coefficients can be determined from

$$C_n = B_n \sinh\left(\frac{n\pi}{a}b\right) = \frac{2}{a} \int_0^a \sin\left(\frac{n\pi}{a}x\right) f(x) dx.$$

In particular, if the fourth side is held at the constant temperature  $T_0$ , then

$$C_n = \frac{2T_0}{a} \left(\frac{a}{n\pi}\right) [1 - (-1)^n] = \begin{cases} \frac{4T_0}{n\pi} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even,} \end{cases}$$

and we obtain

$$T(x, y) = \frac{4T_0}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \frac{\sin[(2k+1)\pi x/a] \sinh[(2k+1)\pi y/a]}{\sinh[(2k+1)\pi b/a]}. \quad (19.6)$$

If the temperature variation of the fourth side is of the form  $f(x) = T_0 \sin(\pi x/a)$ , then

$$C_n = \frac{2T_0}{a} \int_0^a \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{\pi x}{a}\right) dx = \frac{2T_0}{a} \left(\frac{a}{2}\right) \delta_{n,1} = T_0 \delta_{n,1}$$

and  $B_n = C_n / \sinh(n\pi b/a) = [T_0 / \sinh(n\pi b/a)]\delta_{n,1}$ , and we have

$$T(x, y) = T_0 \frac{\sin(\pi x/a) \sinh(\pi y/a)}{\sinh(\pi b/a)}. \quad (19.7)$$

Only one term of the series survives in this case because the variation on the fourth side happens to be one of the harmonics of the expansion.

Note that the temperature variations given by (19.6) and (19.7) are independent of the material of the plate because we are dealing with a steady state. The conductivity of the material is a factor only in the process of heat transfer, leading to the steady state. Once equilibrium has been reached, the distribution of temperature will be the same for all materials. ■

The preceding two examples concerned themselves with static situations. The remaining examples of this section are drawn from the (time-dependent) diffusion equation, the Schrödinger equation, and the wave equation.

conduction of heat in  
a rectangular plate

### 19.2.3. Example. CONDUCTION OF HEAT IN A RECTANGULAR PLATE

Consider a rectangular heat-conducting plate with sides of length  $a$  and  $b$  all held at  $T = 0$ . Assume that at time  $t = 0$  the temperature has a distribution function  $f(x, y)$ . Let us find the variation of temperature for all points  $(x, y)$  at all times  $t > 0$ .

The diffusion equation for this problem is

$$\frac{\partial T}{\partial t} = k^2 \nabla^2 T = k^2 \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right).$$

A separation of variables,  $T(x, y, t) = X(x)Y(y)g(t)$ , leads to three DEs:

$$\frac{d^2 X}{dx^2} + \lambda X = 0, \quad \frac{d^2 Y}{dy^2} + \mu Y = 0, \quad \frac{dg}{dt} + k^2(\lambda + \mu)g = 0.$$

The BCs  $T(0, y, t) = T(a, y, t) = T(x, 0, t) = T(x, b, t) = 0$ , together with the three ODEs, give rise to two S-L systems. The solutions to both of these are easily found:

$$\begin{aligned} \lambda_n &= \left(\frac{n\pi}{a}\right)^2 & \text{and} & & X_n(x) &= \sin\left(\frac{n\pi}{a}x\right) & \text{for } n = 1, 2, \dots, \\ \mu_m &= \left(\frac{m\pi}{b}\right)^2 & \text{and} & & Y_m(y) &= \sin\left(\frac{m\pi}{b}y\right) & \text{for } m = 1, 2, \dots \end{aligned}$$

These give rise to the general solutions

$$X(x) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{a}x\right), \quad Y(y) = \sum_{m=1}^{\infty} B_m \sin\left(\frac{m\pi}{b}y\right).$$

With  $\gamma_{mn} \equiv k^2 \pi^2 (n^2/a^2 + m^2/b^2)$ , the solution to the  $g$  equation can be expressed as  $g(t) = C_{mn} e^{-\gamma_{mn}t}$ . Putting everything together, we obtain

$$T(x, y, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{mn} e^{-\gamma_{mn}t} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right),$$

where  $A_{mn} = A_n B_m C_{mn}$  is an arbitrary constant. To determine it, we impose the initial condition  $T(x, y, 0) = f(x, y)$ . This yields

$$f(x, y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{mn} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right),$$

which determines the coefficients  $A_{mn}$ :

$$A_{mn} = \frac{4}{ab} \int_0^a dx \int_0^b dy f(x, y) \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right). \quad \blacksquare$$

quantum particle in a box

**19.2.4. Example.** QUANTUM PARTICLE IN A BOX

The behavior of an atomic particle of mass  $\mu$  confined in a rectangular box with sides  $a, b,$  and  $c$  (an infinite three-dimensional potential well) is governed by the Schrödinger equation for a free particle,

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2\mu} \left( \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right),$$

and the BC that  $\psi(x, y, z, t)$  vanishes at all sides of the box for all time.

A separation of variables  $\psi(x, y, z, t) = X(x)Y(y)Z(z)T(t)$  yields the ODEs

$$\begin{aligned} \frac{d^2 X}{dx^2} + \lambda X &= 0, & \frac{d^2 Y}{dy^2} + \sigma Y &= 0, & \frac{d^2 Z}{dz^2} + \nu Z &= 0, \\ \frac{dT}{dt} + i\omega T &= 0, & \text{where } \omega &\equiv \frac{\hbar}{2\mu}(\lambda + \sigma + \nu). \end{aligned}$$

The spatial equations, together with the BCs

$$\begin{aligned} \psi(0, y, z, t) = \psi(a, y, z, t) = 0 &\Rightarrow X(0) = 0 = X(a), \\ \psi(x, 0, z, t) = \psi(x, b, z, t) = 0 &\Rightarrow Y(0) = 0 = Y(b), \\ \psi(x, y, 0, t) = \psi(x, y, c, t) = 0 &\Rightarrow Z(0) = 0 = Z(c), \end{aligned}$$

lead to three S-L systems, whose solutions are easily found:

$$\begin{aligned} X_n(x) &= \sin\left(\frac{n\pi}{a}x\right), & \lambda_n &= \left(\frac{n\pi}{a}\right)^2, & \text{for } n &= 1, 2, \dots, \\ Y_m(y) &= \sin\left(\frac{m\pi}{b}y\right), & \sigma_m &= \left(\frac{m\pi}{b}\right)^2, & \text{for } m &= 1, 2, \dots, \\ Z_l(z) &= \sin\left(\frac{l\pi}{c}z\right), & \nu_l &= \left(\frac{l\pi}{c}\right)^2, & \text{for } l &= 1, 2, \dots \end{aligned}$$

The time equation, on the other hand, has a solution of the form

$$T(t) = C_{lmn} e^{-i\omega_{lmn}t} \quad \text{where } \omega_{lmn} = \frac{\hbar}{2\mu} \left[ \left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2 + \left(\frac{l\pi}{c}\right)^2 \right].$$

The solution of the Schrödinger equation that is consistent with the BCs is therefore

$$\psi(x, y, z, t) = \sum_{l,m,n=1}^{\infty} A_{lmn} e^{-i\omega_{lmn}t} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right) \sin\left(\frac{l\pi}{c}z\right).$$

The constants  $A_{lmn}$  are determined by the initial shape,  $\psi(x, y, z, 0)$  of the wave function. The energy of the particle is

$$E = \hbar\omega_{lmn} = \frac{\hbar^2\pi^2}{2\mu} \left( \frac{n^2}{a^2} + \frac{m^2}{b^2} + \frac{l^2}{c^2} \right).$$

Each set of three positive integers  $(n, m, l)$  represents a state of the particle. For a cube,  $a = b = c \equiv L$ , and the energy of the particle is

$$E = \frac{\hbar^2\pi^2}{2\mu L^2}(n^2 + m^2 + l^2) = \frac{\hbar^2\pi^2}{2\mu V^{2/3}}(n^2 + m^2 + l^2) \quad (19.8)$$

where  $V = L^3$  is the volume of the box. The ground state is  $(1, 1, 1)$ , has energy  $E = 3\hbar^2\pi^2/2\mu V^{2/3}$ , and is nondegenerate (only one state corresponds to this energy). However, the higher-level states are degenerate. For instance, the three distinct states  $(1, 1, 2)$ ,  $(1, 2, 1)$ , and  $(2, 1, 1)$  all correspond to the same energy,  $E = 6\hbar^2\pi^2/2\mu V^{2/3}$ . The degeneracy increases rapidly with larger values of  $n, m$ , and  $l$ .

Equation (19.8) can be written as

$$n^2 + m^2 + l^2 = R^2, \quad \text{where} \quad R^2 = \frac{2\mu E V^{2/3}}{\hbar^2\pi^2}.$$

This looks like the equation of a sphere in the  $nml$ -space. If  $R$  is large, the number of states contained within the sphere of radius  $R$  (the number of states with energy less than or equal to  $E$ ) is simply the volume of the first octant<sup>1</sup> of the sphere. If  $N$  is the number of such states, we have

$$N = \frac{1}{8} \left( \frac{4\pi}{3} \right) R^3 = \frac{\pi}{6} \left( \frac{2\mu E V^{2/3}}{\hbar^2\pi^2} \right)^{3/2} = \frac{\pi}{6} \left( \frac{2\mu E}{\hbar^2\pi^2} \right)^{3/2} V.$$

density of states Thus the **density of states** (the number of states per unit volume) is

$$n = \frac{N}{V} = \frac{\pi}{6} \left( \frac{2\mu}{\hbar^2\pi^2} \right)^{3/2} E^{3/2}. \quad (19.9)$$

This is an important formula in solid-state physics, because the energy  $E$  is (with minor modifications required by spin) the **Fermi energy**. If the Fermi energy is denoted by  $E_f$ , Equation (19.9) gives  $E_f = \alpha n^{2/3}$  where  $\alpha$  is some constant. ■

Fermi energy

In the preceding examples the time variation is given by a first derivative. Thus, as far as time is concerned, we have a FODE. It follows that the initial specification of the physical quantity of interest (temperature  $T$  or Schrödinger wave function  $\psi$ ) is sufficient to determine the solution uniquely.

A second kind of time-dependent PDE occurring in physics is the wave equation, which contains time derivatives of the second order. Thus, there are two arbitrary parameters in the general solution. To determine these, we expect two initial conditions. For example, if the wave is standing, as in a rope clamped at both

<sup>1</sup>This is because  $n, m$ , and  $l$  are all positive.

ends, the boundary conditions are not sufficient to determine the wave function uniquely. One also needs to specify the initial (transverse) velocity of each point of the rope. For traveling waves, specification of the wave shape and velocity shape is not as important as the mode of propagation. For instance, in the theory of wave guides, after the time variation is separated, a particular time variation, such as  $e^{+i\omega t}$ , and a particular direction for the propagation of the wave, say the  $z$ -axis, are chosen. Thus, if  $u$  denotes a component of the electric or the magnetic field, we can write  $u(x, y, z, t) = \psi(x, y)e^{i(\omega t \pm kz)}$ , where  $k$  is the wave number. The wave equation then reduces to

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \left( \frac{\omega^2}{c^2} - k^2 \right) \psi = 0.$$

Introducing  $\gamma^2 = \omega^2/c^2 - k^2$  and the transverse gradient  $\nabla_t = (\partial/\partial x, \partial/\partial y)$  and writing the above equation in terms of the full vectors, we obtain

$$(\nabla_t^2 + \gamma^2) \begin{Bmatrix} \mathbf{E} \\ \mathbf{B} \end{Bmatrix} = 0, \quad \text{where} \quad \begin{Bmatrix} \mathbf{E} \\ \mathbf{B} \end{Bmatrix} = \begin{Bmatrix} \mathbf{E}(x, y) \\ \mathbf{B}(x, y) \end{Bmatrix} e^{i(\omega t \pm kz)}. \quad (19.10)$$

These are the basic equations used in the study of electromagnetic wave guides and resonant cavities.

guided waves

Maxwell's equations in conjunction with Equation (19.10) gives the transverse components (components perpendicular to the propagation direction)  $\mathbf{E}_t$  and  $\mathbf{B}_t$  in terms of the longitudinal components  $E_z$  and  $B_z$  (see [Lorr 88, Chapter 33]):

$$\begin{aligned} \gamma^2 \mathbf{E}_t &= \nabla_t \left( \frac{\partial E_z}{\partial z} \right) - i \frac{\omega}{c} \hat{\mathbf{e}}_z \times (\nabla_t B_z), \\ \gamma^2 \mathbf{B}_t &= \nabla_t \left( \frac{\partial B_z}{\partial z} \right) + i \frac{\omega}{c} \hat{\mathbf{e}}_z \times (\nabla_t E_z). \end{aligned} \quad (19.11)$$

Three types of guided waves are usually studied.

1. Transverse magnetic (TM) waves have  $B_z = 0$  everywhere. The BC on  $\mathbf{E}$  demands that  $E_z$  vanish at the conducting walls of the guide.
2. Transverse electric (TE) waves have  $E_z = 0$  everywhere. The BC on  $\mathbf{B}$  requires that the normal directional derivative

$$\frac{\partial B_z}{\partial n} \equiv \hat{\mathbf{e}}_n \cdot (\nabla B_z)$$

vanish at the walls.

3. Transverse electromagnetic (TEM) waves have  $B_z = 0 = E_z$ . For a nontrivial solution, Equation (19.11) demands that  $\gamma^2 = 0$ . This form resembles a free wave with no boundaries.

We will discuss the TM mode briefly (see any book on electromagnetic theory for further details). The basic equations in this mode are

$$\begin{aligned}(\nabla_t^2 + \gamma^2)E_z &= 0, & B_z &= 0, \\ \gamma^2 \mathbf{E}_t &= \nabla_t \left( \frac{\partial E_z}{\partial z} \right), & \gamma^2 \mathbf{B}_t &= i \frac{\omega}{c} \hat{\mathbf{e}}_z \times (\nabla_t E_z).\end{aligned}\quad (19.12)$$

rectangular wave  
guides

### 19.2.5. Example. RECTANGULAR WAVE GUIDES

For a wave guide with a rectangular cross section of sides  $a$  and  $b$  in the  $x$  and the  $y$  directions, respectively, we have

$$\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + \gamma^2 E_z = 0.$$

A separation of variables,  $E_z(x, y) = X(x)Y(y)$ , leads to two S-L systems,

$$\begin{aligned}\frac{d^2 X}{dx^2} + \lambda X &= 0, & X(0) &= 0 = X(a), \\ \frac{d^2 Y}{dy^2} + \mu Y &= 0, & Y(0) &= 0 = Y(b),\end{aligned}$$

where  $\gamma^2 = \lambda + \mu$ . These equations have the solutions

$$\begin{aligned}X_n(x) &= \sin\left(\frac{n\pi}{a}x\right), & \lambda_n &= \left(\frac{n\pi}{a}\right)^2 & \text{for } n = 1, 2, \dots, \\ Y_m(y) &= \sin\left(\frac{m\pi}{b}y\right), & \mu_m &= \left(\frac{m\pi}{b}\right)^2 & \text{for } m = 1, 2, \dots\end{aligned}$$

The wave number is given by

$$k_{mn} = \sqrt{\frac{\omega^2}{c^2} - \left(\frac{n\pi}{a}\right)^2 - \left(\frac{m\pi}{b}\right)^2},$$

which has to be real if the wave is to propagate (an imaginary  $k$  leads to exponential decay or growth along the  $z$ -axis). Thus, there is a cutoff frequency,

$$\omega_{mn} = c \sqrt{\left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2} \quad \text{for } m, n \geq 1,$$

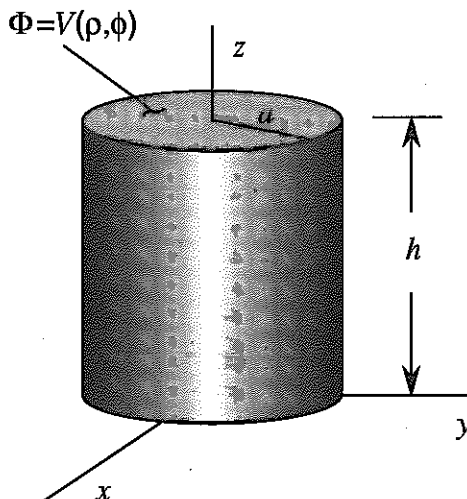
below which the wave cannot propagate through the wave guide. It follows that for a TM wave the lowest frequency that can propagate along a rectangular wave guide is  $\omega_{11} = \pi c \sqrt{a^2 + b^2} / (ab)$ .

The most general solution for  $E_z$  is therefore

$$E_z = \sum_{m,n=1}^{\infty} A_{mn} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right) e^{i(\omega t \pm k_{mn}z)}.$$

The constants  $A_{mn}$  are arbitrary and can be determined from the initial shape of the wave, but that is not commonly done. Once  $E_z$  is found, the other components can be calculated using Equation (19.12). ■





**Figure 19.3** A conducting cylindrical can whose top has a potential given by  $V(\rho, \phi)$ , with the rest of the surface grounded.

## 19.3 Separation in Cylindrical Coordinates

When the geometry of the boundaries is cylindrical, the appropriate coordinate system is the cylindrical one. This usually leads to Bessel functions “of some kind.”

Before working specific examples of cylindrical geometry, let us consider a question that has more general implications. We saw in the previous section that separation of variables leads to ODEs in which certain constants (eigenvalues) appear. Different choices of signs for these constants can lead to different functional forms of the general solution. For example, an equation such as  $d^2x/dt^2 - kx = 0$  can have exponential solutions if  $k > 0$  or trigonometric solutions if  $k < 0$ . One cannot *a priori* assign a specific sign to  $k$ . Thus, the *general* form of the solution is indeterminate. However, once the boundary conditions are imposed, the unique solutions will emerge regardless of the initial functional form of the solutions (see [Hass 99] for a thorough discussion of this point).

conducting  
cylindrical can

### 19.3.1. Example. CONDUCTING CYLINDRICAL CAN

Consider a cylindrical conducting can of radius  $a$  and height  $h$  (see Figure 19.3). The potential varies at the top face as  $V(\rho, \phi)$ , while the lateral surface and the bottom face are grounded. Let us find the electrostatic potential at all points inside the can.

A separation of variables transforms Laplace’s equation into three ODEs:

$$\frac{d}{d\rho} \left( \rho \frac{dR}{d\rho} \right) + \left( k^2 \rho - \frac{m^2}{\rho} \right) R = 0,$$

$$\frac{d^2 S}{d\varphi^2} + m^2 S = 0, \quad \frac{d^2 Z}{dz^2} - k^2 Z = 0,$$

where in anticipation of the correct BCs, we have written the constants as  $k^2$  and  $-m^2$  with  $m$  an integer. The first of these is the Bessel equation, whose general solution can be written as  $R(\rho) = AJ_m(k\rho) + BY_m(k\rho)$ . The second DE, when the extra condition of periodicity is imposed on the potential, has the general solution

$$S(\varphi) = C \cos m\varphi + D \sin m\varphi.$$

Finally the third DE has a general solution of the form

$$Z(z) = Ee^{kz} + Fe^{-kz}.$$

We note that none of the three ODEs lead to an S-L system of Theorem 19.1.1 because the BCs associated with them do not satisfy (19.1). However, we can still solve the problem by imposing the given BCs.

The fact that the potential must be finite everywhere inside the can (including at  $\rho = 0$ ) forces  $B$  to vanish because the Neumann function  $Y_m(k\rho)$  is not defined at  $\rho = 0$ . On the other hand, we want  $\Phi$  to vanish at  $\rho = a$ . This gives  $J_m(ka) = 0$ , which demands that  $ka$  be a root of the Bessel function of order  $m$ . Denoting by  $x_{mn}$  the  $n$ th zero of the Bessel function of order  $m$ , we have  $ka = x_{mn}$ , or  $k = x_{mn}/a$  for  $n = 1, 2, \dots$

Similarly, the vanishing of  $\Phi$  at  $z = 0$  implies that

$$E = -F \quad \text{and} \quad Z(z) = E \sinh\left(\frac{x_{mn}z}{a}\right).$$

We can now multiply  $R$ ,  $S$ , and  $Z$  and sum over all possible values of  $m$  and  $n$ , keeping in mind that negative values of  $m$  give terms that are linearly dependent on the corresponding positive values. The result is the so-called **Fourier-Bessel series**:

Fourier-Bessel series

$$\Phi(\rho, \varphi, z) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} J_m\left(\frac{x_{mn}}{a}\rho\right) \sinh\left(\frac{x_{mn}}{a}z\right) (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi), \quad (19.13)$$

where  $A_{mn}$  and  $B_{mn}$  are constants to be determined by the remaining BC. To find these constants we use the orthogonality of the trigonometric and Bessel functions. For  $z = h$  Equation (19.13) reduces to

$$V(\rho, \varphi) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} J_m\left(\frac{x_{mn}}{a}\rho\right) \sinh\left(\frac{x_{mn}}{a}h\right) (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi),$$

from which we obtain

$$A_{mn} = \frac{2}{\pi a^2 J_{m+1}^2(x_{mn}) \sinh(x_{mn}h/a)} \int_0^{2\pi} d\varphi \int_0^a d\rho \rho V(\rho, \varphi) J_m\left(\frac{x_{mn}}{a}\rho\right) \cos m\varphi,$$

$$B_{mn} = \frac{2}{\pi a^2 J_{m+1}^2(x_{mn}) \sinh(x_{mn}h/a)} \int_0^{2\pi} d\varphi \int_0^a d\rho \rho V(\rho, \varphi) J_m\left(\frac{x_{mn}}{a}\rho\right) \sin m\varphi,$$

where we have used the following result derived in Problem 14.39:

$$\int_0^a \rho J_m^2\left(\frac{x_{mn}}{a}\rho\right) d\rho = \frac{a^2}{2} J_{m+1}^2(x_{mn}). \quad (19.14)$$

For the special but important case of azimuthal symmetry, for which  $V$  is independent of  $\varphi$ , we obtain

$$A_{mn} = \frac{4\delta_{m,0}}{a^2 J_1^2(x_{0n}) \sinh(x_{0n}h/a)} \int_0^a d\rho \rho V(\rho) J_0\left(\frac{x_{0n}}{a} \rho\right),$$

$$B_{mn} = 0.$$

The reason we obtained discrete values for  $k$  was the demand that  $\Phi$  vanish at  $\rho = a$ . If we let  $a \rightarrow \infty$ , then  $k$  will be a continuous variable, and instead of a sum over  $k$ , we will obtain an integral. This is completely analogous to the transition from a Fourier series to a Fourier transform, but we will not pursue it further.

circular  
heat-conducting  
plate

**19.3.2. Example.** CIRCULAR HEAT-CONDUCTING PLATE

Consider a circular heat-conducting plate of radius  $a$  whose temperature at time  $t = 0$  has a distribution function  $f(\rho, \varphi)$ . Let us find the variation of  $T$  for all points  $(\rho, \varphi)$  on the plate for time  $t > 0$  when the edge is kept at  $T = 0$ .

This is a two-dimensional problem involving the heat equation,

$$\frac{\partial T}{\partial t} = k^2 \nabla^2 T = k^2 \left[ \frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial T}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 T}{\partial \varphi^2} \right].$$

A separation of variables,  $T(\rho, \varphi, t) = R(\rho)S(\varphi)g(t)$ , leads to the following ODEs:

$$\frac{dg}{dt} = k^2 \lambda g, \quad \frac{d^2 S}{d\varphi^2} + \mu S = 0,$$

$$\frac{d^2 R}{d\rho^2} + \frac{1}{\rho} \frac{dR}{d\rho} - \left( \frac{\mu}{\rho^2} + \lambda \right) R = 0.$$

To obtain exponential decay rather than growth for the temperature, we demand that  $\lambda \equiv b^2 < 0$ . To ensure periodicity (see the discussion at the beginning of this section), we must have  $\mu = m^2$ , where  $m$  is an integer. To have finite  $T$  at  $\rho = 0$ , no Neumann function is to be present. This leads to the following solutions:

$$g(t) = Ae^{-k^2 b^2 t}, \quad S(\varphi) = B \cos m\varphi + C \sin m\varphi, \quad R(\rho) = DJ_m(b\rho).$$

If the temperature is to be zero at  $\rho = a$ , we must have  $J_m(ba) = 0$ , or  $b = x_{mn}/a$ . It follows that the general solution can be written as

$$T(\rho, \varphi, t) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} e^{-k^2 (x_{mn}/a)^2 t} J_m\left(\frac{x_{mn}}{a} \rho\right) (A_{mn} \cos m\varphi + B_{mn} \sin m\varphi).$$

$A_{mn}$  and  $B_{mn}$  can be determined as in the preceding example.

cylindrical wave  
guide

**19.3.3. Example.** CYLINDRICAL WAVE GUIDE

For a TM wave propagating along the  $z$ -axis in a hollow circular conductor, we have [see Equation (19.12)]

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial E_z}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 E_z}{\partial \varphi^2} + \gamma^2 E_z = 0.$$

The separation  $E_z = R(\rho)S(\varphi)$  yields  $S(\varphi) = A \cos m\varphi + B \sin m\varphi$  and

$$\frac{d^2R}{d\rho^2} + \frac{1}{\rho} \frac{dR}{d\rho} + \left( \gamma^2 - \frac{m^2}{\rho^2} \right) R = 0.$$

The solution to this equation, which is regular at  $\rho = 0$  and vanishes at  $\rho = a$ , is

$$R(\rho) = C J_m\left(\frac{x_{mn}}{a} \rho\right) \quad \text{and} \quad \gamma = \frac{x_{mn}}{a}.$$

Recalling the definition of  $\gamma$ , we obtain

$$\frac{\omega^2}{c^2} - k^2 = \gamma^2 = \frac{x_{mn}^2}{a^2} \Rightarrow k = \sqrt{\frac{\omega^2}{c^2} - \frac{x_{mn}^2}{a^2}}.$$

This gives the cut-off frequency  $\omega_{mn} = cx_{mn}/a$ .

The solution for the azimuthally symmetric case ( $m = 0$ ) is

$$E_z(\rho, \varphi, t) = \sum_{n=1}^{\infty} A_n J_0\left(\frac{x_{0n}}{a} \rho\right) e^{i(\omega t \pm k_n z)} \quad \text{and} \quad B_z = 0,$$

where  $k_n = \sqrt{\omega^2/c^2 - x_{0n}^2/a^2}$ . ■

There are many variations on the theme of Bessel functions. We have encountered three kinds of Bessel functions, as well as modified Bessel functions. Another variation encountered in applications leads to what are known as *Kelvin functions*, introduced in the following example.

current distribution  
in a circular wire

**19.3.4. Example.** CURRENT DISTRIBUTION IN A CIRCULAR WIRE

Consider the flow of charges in an infinitely long wire with a circular cross section of radius  $a$ . We are interested in calculating the variation of the current density in the wire as a function of time and location. The relevant equation can be obtained by starting with Maxwell's equations for negligible charge density ( $\nabla \cdot \mathbf{E} = 0$ ), Ohm's law ( $\mathbf{j} = \sigma \mathbf{E}$ ), the assumption of high electrical conductivity ( $|\sigma \mathbf{E}| \gg |\partial \mathbf{E} / \partial t|$ ), and the usual procedure of obtaining the wave equation from Maxwell's equations. The result is

$$\nabla^2 \mathbf{j} - \frac{4\pi\sigma}{c^2} \frac{\partial \mathbf{j}}{\partial t} = 0.$$

Moreover, we make the simplifying assumptions that the wire is along the  $z$ -axis and that there is no turbulence, so  $\mathbf{j}$  is also along the  $z$  direction. We further assume that  $\mathbf{j}$  is independent of  $\varphi$  and  $z$ , and that its time-dependence is given by  $e^{-i\omega t}$ . Then we get

$$\frac{d^2 j}{d\rho^2} + \frac{1}{\rho} \frac{d j}{d\rho} + \tau^2 j = 0, \tag{19.15}$$

skin depth  
Kelvin equation

where  $\tau^2 = i4\pi\sigma\omega/c^2 \equiv i2/\delta^2$  and  $\delta = c/\sqrt{2\pi\sigma\omega}$  is called the **skin depth**.

The **Kelvin equation** is usually given as

$$\frac{d^2 w}{dx^2} + \frac{1}{x} \frac{d w}{dx} - ik^2 w = 0. \tag{19.16}$$

If we substitute  $x = \sqrt{t} t/k$ , it becomes  $\ddot{w} + \dot{w}/t + w = 0$  which is a Bessel equation of order zero. If the solution is to be regular at  $x = 0$ , then the only choice is  $w(t) = J_0(t) = J_0(e^{-i\pi/4} kx)$ . This is the **Kelvin function** for Equation (19.16). It is usually written as

$$J_0(e^{-i\pi/4} kx) \equiv \text{ber}(kx) + i \text{bei}(kx)$$

where ber and bei stand for “Bessel real” and “Bessel imaginary,” respectively. If we substitute  $z = e^{-i\pi/4} kx$  in the expansion for  $J_0(z)$  and separate the real and the imaginary parts of the expansion, we obtain

$$\begin{aligned} \text{ber}(x) &= 1 - \frac{(x/2)^4}{(2!)^2} + \frac{(x/2)^8}{(4!)^2} - \dots \\ \text{bei}(x) &= \frac{(x/2)^2}{(1!)^2} - \frac{(x/2)^6}{(3!)^2} + \frac{(x/2)^{10}}{(5!)^2} - \dots \end{aligned}$$

Equation (19.15) is the complex conjugate of (19.16) with  $k^2 = 2/\delta^2$ . Thus, its solution is

$$j(\rho) = A J_0(e^{i\pi/4} k\rho) = A \left\{ \text{ber} \left( \frac{\sqrt{2}}{\delta} \rho \right) - i \text{bei} \left( \frac{\sqrt{2}}{\delta} \rho \right) \right\}.$$

We can compare the value of the current density at  $\rho$  with its value at the surface  $\rho = a$ :

$$\left| \frac{j(\rho)}{j(a)} \right| = \left[ \frac{\text{ber}^2 \left( \frac{\sqrt{2}}{\delta} \rho \right) + \text{bei}^2 \left( \frac{\sqrt{2}}{\delta} \rho \right)}{\text{ber}^2 \left( \frac{\sqrt{2}}{\delta} a \right) + \text{bei}^2 \left( \frac{\sqrt{2}}{\delta} a \right)} \right]^{1/2}$$

For low frequencies,  $\delta$  is large, which implies that  $\rho/\delta$  is small; thus,  $\text{ber}(\sqrt{2}\rho/\delta) \approx 1$  and  $\text{bei}(\sqrt{2}\rho/\delta) \approx 0$ , and  $|j(\rho)/j(a)| \approx 1$ ; i.e., the current density is almost uniform. For higher frequencies the ratio of the current densities starts at a value less than 1 at  $\rho = 0$  and increases to 1 at  $\rho = a$ . The starting value depends on the frequency. For very large frequencies the starting value is almost zero (see [Mari 80, pp 150–156]). ■

quantum particle in a cylindrical can

**19.3.5. Example.** QUANTUM PARTICLE IN A CYLINDRICAL CAN

Let us consider a quantum particle in a cylindrical can. For an atomic particle of mass  $\mu$  confined in a cylindrical can of length  $L$  and radius  $a$ , the relevant Schrödinger equation is

$$i \frac{\partial \psi}{\partial t} = -\frac{\hbar}{2\mu} \left[ \frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial \psi}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 \psi}{\partial \varphi^2} + \frac{\partial^2 \psi}{\partial z^2} \right].$$

Let us solve this equation subject to the BC that  $\psi(\rho, \varphi, z, t)$  vanishes at the sides of the can.

A separation of variables,  $\psi(\rho, \varphi, z, t) = R(\rho)S(\varphi)Z(z)T(t)$ , yields

$$\begin{aligned} \frac{dT}{dt} &= -i\omega T & \frac{d^2Z}{dz^2} + \lambda Z &= 0, & \frac{d^2S}{d\varphi^2} + m^2 S &= 0, \\ \frac{d^2R}{d\rho^2} + \frac{1}{\rho} \frac{dR}{d\rho} + \left( \frac{2\mu}{\hbar} \omega - \lambda - \frac{m^2}{\rho^2} \right) R &= 0. \end{aligned} \tag{19.17}$$

The  $Z$  equation, along with its periodic BCs, constitutes an S-L system with solutions

$$Z(z) = \sin\left(\frac{k\pi}{L}z\right) \quad \lambda = \left(\frac{k\pi}{L}\right)^2 \quad \text{for } k = 1, 2, \dots$$

If we let  $2\mu\omega/\hbar - (k\pi/L)^2 \equiv b^2$ , then the last equation in (19.17) becomes

$$\frac{d^2 R}{d\rho^2} + \frac{1}{\rho} \frac{dR}{d\rho} + \left(b^2 - \frac{m^2}{\rho^2}\right) R = 0,$$

and the solution that is well-behaved at  $\rho = 0$  is  $J_m(b\rho)$ . Since  $R(a) = 0$ , we obtain the quantization condition  $b = x_{mn}/a$  for  $n = 1, 2, \dots$ . Thus, the energy eigenvalues are

$$E_{kmn} \equiv \hbar\omega_{kmn} = \frac{\hbar^2}{2\mu} \left[ \left(\frac{k\pi}{L}\right)^2 + \frac{x_{mn}^2}{a^2} \right],$$

and the general solution can be written as

$$\psi(\rho, \varphi, z, t) = \sum_{\substack{k,n=1 \\ m=0}}^{\infty} e^{-i\omega_{kmn}t} J_m\left(\frac{x_{mn}}{a}\rho\right) \sin\left(\frac{k\pi}{L}z\right) (A_{kmn} \cos m\varphi + B_{kmn} \sin m\varphi). \quad \blacksquare$$

## 19.4 Separation in Spherical Coordinates

Recall that most PDEs encountered in physical applications can be separated, in spherical coordinates, into

$$\begin{aligned} \mathbf{L}^2 Y(\theta, \varphi) &= l(l+1)Y(\theta, \varphi), \\ \frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left[ f(r) - \frac{l(l+1)}{r^2} \right] R &= 0. \end{aligned} \quad (19.18)$$

We discussed the first of these two equations in great detail in Chapter 12. In particular, we constructed  $Y_{lm}(\theta, \varphi)$  in such a way that they formed an orthonormal sequence. However, that construction was purely algebraic and did not say anything about the completeness of  $Y_{lm}(\theta, \varphi)$ . With Theorem 19.1.1 at our disposal, we can separate the first equation of (19.18) into two ODEs by writing  $Y_{lm}(\theta, \varphi) = P_{lm}(\theta)S_m(\varphi)$ . We obtain

$$\begin{aligned} \frac{d^2 S_m}{d\varphi^2} + m^2 S_m &= 0, \\ \frac{d}{dx} \left[ (1-x^2) \frac{dP_{lm}}{dx} \right] + \left[ l(l+1) - \frac{m^2}{1-x^2} \right] P_{lm} &= 0, \end{aligned}$$

where  $x = \cos \theta$ . These are both S-L systems satisfying the conditions of Theorem 19.1.1. Thus, the  $S_m$  are orthogonal among themselves and form a complete set for  $\mathcal{L}^2(0, 2\pi)$ . Similarly, for any fixed  $m$ , the  $P_{lm}(x)$  form a complete orthogonal set for  $\mathcal{L}^2(-1, +1)$  (actually for the subset of  $\mathcal{L}^2(-1, +1)$  that satisfies the same

BC as the  $P_{lm}$  do at  $x = \pm 1$ ). Thus, the products  $Y_{lm}(x, \varphi) = P_{lm}(x)S_m(\varphi)$  form a complete orthogonal sequence in the (Cartesian product) set  $[-1, +1] \times [0, 2\pi]$ , which, in terms of spherical angles, is the unit sphere,  $0 \leq \theta \leq \pi$ ,  $0 \leq \varphi \leq 2\pi$ .

Let us consider some specific examples of expansion in the spherical coordinate system starting with the simplest case, Laplace's equation for which  $f(r) = 0$ . The radial equation is therefore

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} - \frac{l(l+1)}{r^2} R = 0.$$

Multiplying by  $r^2$ , substituting  $r = e^t$ , and using the chain rule and the fact that  $dt/dr = 1/r$  leads to the following SOLDE with constant coefficients:

$$\frac{d^2 R}{dt^2} + \frac{dR}{dt} - l(l+1)R = 0.$$

This has a characteristic polynomial  $p(\lambda) = \lambda^2 + \lambda - l(l+1)$  with roots  $\lambda_1 = l$  and  $\lambda_2 = -(l+1)$ . Thus, a general solution is of the form

$$R(t) = Ae^{\lambda_1 t} + Be^{\lambda_2 t} = A(e^t)^l + B(e^t)^{-l-1},$$

or, in terms of  $r$ ,  $R(r) = Ar^l + Br^{-l-1}$ . Thus, the most general solution of Laplace's equation is

$$\Phi(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l (A_{lm}r^l + B_{lm}r^{-l-1})Y_{lm}(\theta, \varphi).$$

For regions containing the origin, the finiteness of  $\Phi$  implies that  $B_{lm} = 0$ . Denoting the potential in such regions by  $\Phi_{\text{in}}$ , we obtain

$$\Phi_{\text{in}}(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{lm}r^l Y_{lm}(\theta, \varphi).$$

Similarly, for regions including  $r = \infty$ , we have

$$\Phi_{\text{out}}(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l B_{lm}r^{-l-1} Y_{lm}(\theta, \varphi).$$

To determine  $A_{lm}$  and  $B_{lm}$ , we need to invoke appropriate BCs. In particular, for inside a sphere of radius  $a$  on which the potential is given by  $V(\theta, \varphi)$ , we have

$$V(\theta, \varphi) = \Phi_{\text{in}}(a, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{lm}a^l Y_{lm}(\theta, \varphi).$$

Multiplying by  $Y_{kj}^*(\theta, \varphi)$  and integrating over  $d\Omega = \sin \theta d\theta d\varphi$ , we obtain

$$A_{kj} = a^{-k} \iint d\Omega V(\theta, \varphi) Y_{kj}^*(\theta, \varphi) \Rightarrow A_{lm} = a^{-l} \iint d\Omega V(\theta, \varphi) Y_{lm}^*(\theta, \varphi).$$

Similarly, for potential outside the sphere,

$$B_{lm} = a^{l+1} \iint d\Omega V(\theta, \varphi) Y_{lm}^*(\theta, \varphi).$$

In particular, if  $V$  is independent of  $\varphi$ , only the components for which  $m = 0$  are nonzero, and we have

$$A_{l0} = \frac{2\pi}{a^l} \int_0^\pi \sin \theta V(\theta) Y_{l0}^*(\theta) d\theta = \frac{2\pi}{a^l} \sqrt{\frac{2l+1}{4\pi}} \int_0^\pi \sin \theta V(\theta) P_l(\cos \theta) d\theta,$$

which yields

$$\Phi_{\text{in}}(r, \theta) = \sum_{l=0}^{\infty} A_l \left(\frac{r}{a}\right)^l P_l(\cos \theta),$$

where

$$A_l = \frac{2}{2l+1} \int_0^\pi \sin \theta V(\theta) P_l(\cos \theta) d\theta.$$

Similarly,

$$\Phi_{\text{out}}(r, \theta) = \sum_{l=0}^{\infty} A_l \left(\frac{a}{r}\right)^{l+1} P_l(\cos \theta).$$

The next simplest case after Laplace's equation is that for which  $f(r)$  is a constant. The diffusion equation, the wave equation, and the Schrödinger equation for a free particle give rise to such a case once time is separated from the rest of the variables.

Helmholtz equation

The **Helmholtz equation** is

$$\nabla^2 \psi + k^2 \psi = 0, \tag{19.19}$$

and its radial part is

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left[ k^2 - \frac{l(l+1)}{r^2} \right] R = 0. \tag{19.20}$$

(This equation was discussed in Problems 14.26 and 14.35.) The solutions are **spherical Bessel functions**, generically denoted by the corresponding lower case letter as  $z_l(x)$  and given by

spherical Bessel functions

$$z_l(x) \equiv \sqrt{\frac{\pi}{2}} \frac{Z_{l+1/2}(x)}{\sqrt{x}}, \tag{19.21}$$



where  $Z_\nu(x)$  is a solution of the Bessel equation of order  $\nu$ .

A general solution of (19.20) can therefore be written as

$$R_l(r) = A j_l(kr) + B y_l(kr).$$

If the origin is included in the region of interest, then we must set  $B = 0$ . For such a case, the solution to the Helmholtz equation is

$$\psi_k(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{lm} j_l(kr) Y_{lm}(\theta, \varphi). \tag{19.22}$$

The subscript  $k$  indicates that  $\psi$  is a solution of the Helmholtz equation with  $k^2$  as its constant.

particle in a hard sphere

**19.4.1. Example.** PARTICLE IN A HARD SPHERE

The time-independent Schrödinger equation for a particle in a sphere of radius  $a$  is  $-\frac{\hbar^2}{2\mu} \nabla^2 \psi = E \psi$  with the BC  $\psi(a, \theta, \varphi) = 0$ . Here  $E$  is the energy of the particle and  $\mu$  is its mass. We rewrite the Schrödinger equation as  $\nabla^2 \psi + 2\mu E/\hbar^2 = 0$ . With  $k^2 = 2\mu E/\hbar^2$ , we can immediately write the radial solution

$$R_l(r) = A j_l(kr) = A j_l(\sqrt{2\mu E} r/\hbar).$$

The vanishing of  $\psi$  at  $a$  implies that  $j_l(\sqrt{2\mu E} a/\hbar) = 0$ , or

$$\frac{\sqrt{2\mu E} a}{\hbar} = X_{ln} \quad \text{for } n = 1, 2, \dots,$$

where  $X_{ln}$  is the  $n$ th zero of  $j_l(x)$ , which is the same as the zero of  $J_{l+1/2}(x)$ . Thus, the energy is quantized as

$$E_{ln} = \frac{\hbar^2 X_{ln}^2}{2\mu a^2} \quad \text{for } l = 0, 1, \dots, n = 1, 2, \dots$$

The general solution to the Schrödinger equation is

$$\psi(r, \theta, \varphi) = \sum_{n=1}^{\infty} \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{nlm} j_l\left(X_{ln} \frac{r}{a}\right) Y_{lm}(\theta, \varphi). \quad \blacksquare$$

A particularly useful consequence of Equation (19.22) is the expansion of a plane wave in terms of spherical Bessel functions. It is easily verified that if  $\mathbf{k}$  is a vector, with  $\mathbf{k} \cdot \mathbf{k} = k^2$ , then  $e^{i\mathbf{k} \cdot \mathbf{r}}$  is a solution of the Helmholtz equation. Thus,  $e^{i\mathbf{k} \cdot \mathbf{r}}$  can be expanded as in Equation (19.22). Assuming that  $\mathbf{k}$  is along the  $z$ -axis, we get  $\mathbf{k} \cdot \mathbf{r} = kr \cos \theta$ , which is independent of  $\varphi$ . Only the terms of Equation (19.22) for which  $m = 0$  will survive in such a case, and we may write  $e^{ikr \cos \theta} = \sum_{l=0}^{\infty} A_l j_l(kr) P_l(\cos \theta)$ . To find  $A_l$ , let  $u = \cos \theta$ , multiply both sides by  $P_n(u)$ , and integrate from  $-1$  to  $1$ :

$$\int_{-1}^1 P_n(u) e^{ikru} du = \sum_{l=0}^{\infty} A_l j_l(kr) \int_{-1}^1 P_n(u) P_l(u) du = A_n j_n(kr) \frac{2}{2n+1}.$$

Thus

$$\begin{aligned} A_n j_n(kr) &= \frac{2n+1}{2} \int_{-1}^1 P_n(u) e^{ikru} du \\ &= \frac{2n+1}{2} \sum_{m=0}^{\infty} \frac{(ikr)^m}{m!} \int_{-1}^1 P_n(u) u^m du. \end{aligned} \quad (19.23)$$

This equality holds for all values of  $kr$ . In particular, both sides should give the same result in the limit of small  $kr$ . From the definition of  $j_n(kr)$  and the expansion of  $J_n(kr)$ , we obtain

$$j_n(kr) \xrightarrow{kr \rightarrow 0} \frac{\sqrt{\pi}}{2} \left(\frac{kr}{2}\right)^n \frac{1}{\Gamma(n+3/2)}.$$

expansion of  $e^{ikr}$  in spherical harmonics

On the other hand, the first nonvanishing term of the RHS of Equation (19.23) occurs when  $m = n$ . Equating these terms on both sides, we get

$$A_n \frac{\sqrt{\pi}}{2} \left(\frac{kr}{2}\right)^n \frac{2^{2n+1} n!}{(2n+1)! \sqrt{\pi}} = \frac{2n+1}{2} \frac{i^n (kr)^n}{n!} \frac{2^{n+1} (n!)^2}{(2n+1)!}, \quad (19.24)$$

where we have used

$$\Gamma\left(n + \frac{3}{2}\right) = \frac{(2n+1)! \sqrt{\pi}}{2^{2n+1} n!} \quad \text{and} \quad \int_{-1}^1 P_n(u) u^n du = \frac{2^{n+1} (n!)^2}{(2n+1)!}.$$

Equation (19.24) yields  $A_n = i^n (2n+1)$ .

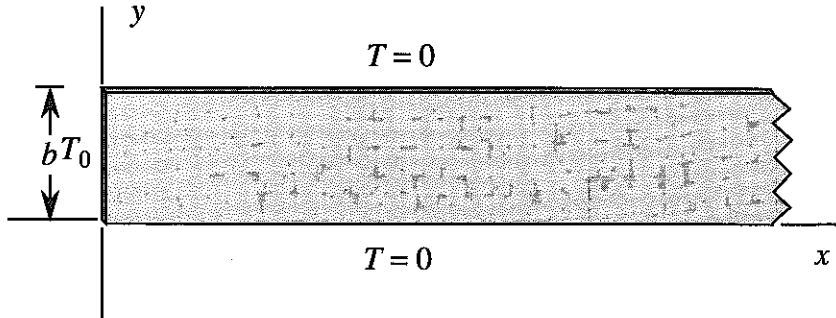
With  $A_n$  thus calculated, we can now write

$$e^{ikr \cos \theta} = \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos \theta). \quad (19.25)$$

For an arbitrary direction of  $\mathbf{k}$ ,  $\mathbf{k} \cdot \mathbf{r} = kr \cos \gamma$ , where  $\gamma$  is the angle between  $\mathbf{k}$  and  $\mathbf{r}$ . Thus, we may write  $e^{i\mathbf{k} \cdot \mathbf{r}} = \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos \gamma)$ , and using the addition theorem for spherical harmonics, we finally obtain

$$e^{i\mathbf{k} \cdot \mathbf{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l i^l j_l(kr) Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi), \quad (19.26)$$

where  $\theta'$  and  $\varphi'$  are the spherical angles of  $\mathbf{k}$  and  $\theta$  and  $\varphi$  are those of  $\mathbf{r}$ . Such a decomposition of plane waves into components with definite orbital angular momenta is extremely useful when working with scattering theory for waves and particles.



**Figure 19.4** A semi-infinite heat-conducting plate.

## 19.5 Problems

**19.1.** Show that separated and periodic BCs are special cases of the equality in Equation (19.3).

**19.2.** Derive Equation (19.4).

**19.3.** A semi-infinite heat-conducting plate of width  $b$  is extended along the positive  $x$ -axis with one corner at  $(0, 0)$  and the other at  $(0, b)$ . The side of width  $b$  is held at temperature  $T_0$ , and the two long sides are held at  $T = 0$  (see Figure 19.4). The two flat faces are insulated. Find the temperature variation of the plate, assuming equilibrium. Repeat the problem with the temperature of the short side held at each of the following:

$$\begin{aligned}
 (a) \quad T &= \begin{cases} 0 & \text{if } 0 < y < b/2, \\ T_0 & \text{if } b/2 < y < b. \end{cases} & (b) \quad \frac{T_0}{b} y, & \quad 0 \leq y \leq b. \\
 (c) \quad T_0 \cos\left(\frac{\pi}{b} y\right), & \quad 0 \leq y \leq b. & (d) \quad T_0 \sin\left(\frac{\pi}{b} y\right), & \quad 0 \leq y \leq b.
 \end{aligned}$$

**19.4.** Find a general solution for the electromagnetic wave propagation in a **resonant cavity**, a rectangular box of sides  $0 \leq x \leq a$ ,  $0 \leq y \leq b$ , and  $0 \leq z \leq d$  with perfectly conducting walls. Discuss the modes the cavity can accommodate.

**19.5.** The lateral faces of a cube are grounded, and its top and bottom faces are held at potentials  $f_1(x, y)$  and  $f_2(x, y)$ , respectively.

(a) Find a general expression for the potential inside the cube.

(b) Find the potential if the top is held at  $V_0$  volts and the bottom at  $-V_0$  volts.

**19.6.** Find the potential inside a semi-infinite cylindrical conductor, closed at the nearby end, whose cross section is a square with sides of length  $a$ . All sides are grounded except the square side, which is held at the constant potential  $V_0$ .

**19.7.** Find the temperature distribution of a rectangular plate (see Figure 19.2) with sides of lengths  $a$  and  $b$  if three sides are held at  $T = 0$  and the fourth side has a temperature variation given by

$$\begin{aligned} (a) \quad & \frac{T_0}{a}x, & 0 \leq x < a. & \quad (b) \quad \frac{T_0}{a^2}x(x-a), & 0 \leq x \leq a. \\ (c) \quad & \frac{T_0}{a} \left| x - \frac{a}{2} \right|, & 0 \leq x < a. & \quad (d) \quad T = 0, & 0 \leq x \leq a. \end{aligned}$$

**19.8.** Consider a thin heat-conducting bar of length  $b$  along the  $x$ -axis with one end at  $x = 0$  held at temperature  $T_0$  and the other end at  $x = b$  held at temperature  $-T_0$ . The lateral surface of the bar is thermally insulated. Find the temperature distribution at all times if initially it is given by

$$\begin{aligned} (a) \quad & T(0, x) = -\frac{2T_0}{b}x + T_0, & \text{where } 0 \leq x \leq b. \\ (b) \quad & T(0, x) = -\frac{2T_0}{b^2}x^2 + T_0, & \text{where } 0 \leq x \leq b. \\ (c) \quad & T(0, x) = \frac{T_0}{b}x + T_0, & \text{where } 0 \leq x < b. \\ (d) \quad & T(0, x) = T_0 \cos\left(\frac{\pi}{b}x\right), & \text{where } 0 \leq x \leq b. \end{aligned}$$

Hint: The solution corresponding to the zero eigenvalue is essential and cannot be excluded.

**19.9.** Determine  $T(x, y, t)$  for the rectangular plate of Example 19.2.3 if initially the lower left quarter is held at  $T_0$  and the rest of the plate is held at  $T = 0$ .

**19.10.** All sides of the plate of Example 19.2.3 are held at  $T = 0$ . Find the temperature distribution for all time if the initial temperature distribution is given by

$$\begin{aligned} (a) \quad & T(x, y, 0) = \begin{cases} T_0 & \text{if } \frac{1}{4}a \leq x \leq \frac{3}{4}a \text{ and } \frac{1}{4}b \leq y \leq \frac{3}{4}b, \\ 0 & \text{otherwise.} \end{cases} \\ (b) \quad & T(x, y, 0) = \frac{T_0}{ab}xy, & \text{where } 0 \leq x < a \text{ and } 0 \leq y < b. \\ (c) \quad & T(x, y, 0) = \frac{T_0}{a}x, & \text{where } 0 \leq x < a \text{ and } 0 < y < b. \end{aligned}$$

**19.11.** Repeat Example 19.2.3 with the temperatures of the sides equal to  $T_1, T_2, T_3,$  and  $T_4$ . Hint: You must include solutions corresponding to the zero eigenvalue.

**19.12.** A string of length  $a$  is fixed at the left end, and the right end moves with displacement  $A \sin \omega t$ . Find  $\psi(x, t)$  and a consistent set of initial conditions for the displacement and the velocity.

**19.13.** Find the equation for a vibrating rectangular membrane with sides of lengths  $a$  and  $b$  rigidly fastened on all sides. For  $a = b$ , show that a given mode frequency may have more than one solution.

**19.14.** Repeat Example 19.3.1 if the can has semi-infinite length, the lateral surface is grounded, and:

(a) the base is held at the potential  $V(\rho, \varphi)$ .

Specialize to the case where the potential of the base is given—in Cartesian coordinates—by

$$(b) V = \frac{V_0}{a}y. \qquad (c) V = \frac{V_0}{a}x. \qquad (d) V = \frac{V_0}{a^2}xy.$$

Hint: Use the integral identity  $\int z^{\nu+1} J_\nu(z) dz = z^{\nu+1} J_{\nu+1}(z)$ .

**19.15.** Find the steady-state temperature distribution  $T(\rho, \varphi, z)$  in a semi-infinite solid cylinder of radius  $a$  if the temperature distribution of the base is  $f(\rho, \varphi)$  and the lateral surface is held at  $T = 0$ .

**19.16.** Find the steady-state temperature distribution of a solid cylinder with a height and radius of 10, assuming that the base and the lateral surface are at  $T = 0$  and the top is at  $T = 100$ .

**19.17.** The circumference of a flat circular plate of radius  $a$ , lying in the  $xy$ -plane, is held at  $T = 0$ . Find the temperature distribution for all time if the temperature distribution at  $t = 0$  is given—in Cartesian coordinates—by

$$(a) \frac{T_0}{a}y. \qquad (b) \frac{T_0}{a}x. \qquad (c) \frac{T_0}{a^2}xy. \qquad (d) T_0.$$

**19.18.** Find the temperature of a circular conducting plate of radius  $a$  at all points of its surface for all time  $t > 0$ , assuming that its edge is held at  $T = 0$  and initially its surface from the center to  $a/2$  is in contact with a heat bath of temperature  $T_0$ .

**19.19.** Find the potential of a cylindrical conducting can of radius  $a$  and height  $h$  whose top is held at a constant potential  $V_0$  while the rest is grounded.

**19.20.** Find the modes and the corresponding fields of a cylindrical resonant cavity of length  $L$  and radius  $a$ . Discuss the lowest TM mode.

**19.21.** Two identical long conducting half-cylindrical shells (cross sections are half-circles) of radius  $a$  are glued together in such a way that they are insulated from one another. One half-cylinder is held at potential  $V_0$  and the other is grounded. Find the potential at any point inside the resulting cylinder. Hint: Separate Laplace's equation in two dimensions.

**19.22.** A linear charge distribution of uniform density  $\lambda$  extends along the  $z$ -axis from  $z = -b$  to  $z = b$ . Show that the electrostatic potential at any point  $r > b$  is given by

$$\Phi(r, \theta, \varphi) = 2\lambda \sum_{k=0}^{\infty} \frac{(b/r)^{2k+1}}{2k+1} P_{2k}(\cos \theta).$$

Hint: Consider a point on the  $z$ -axis at a distance  $r > b$  from the origin. Solve the simple problem by integration and compare the result with the infinite series to obtain the unknown coefficients.

**19.23.** The upper half of a heat-conducting sphere of radius  $a$  has  $T = 100$ ; the lower half is maintained at  $T = -100$ . The whole sphere is inside an infinitely large mass of heat-conducting material. Find the steady-state temperature distribution inside and outside the sphere.

**19.24.** Find the steady-state temperature distribution inside a sphere of radius  $a$  when the surface temperature is given by:

$$\begin{array}{lll} (a) T_0 \cos^2 \theta. & (b) T_0 \cos^4 \theta. & (c) T_0 |\cos \theta|. \\ (d) T_0 (\cos \theta - \cos^3 \theta). & (e) T_0 \sin^2 \theta. & (f) T_0 \sin^4 \theta. \end{array}$$

**19.25.** Find the electrostatic potential both inside and outside a conducting sphere of radius  $a$  when the sphere is maintained at a potential given by

$$\begin{array}{ll} (a) V_0 (\cos \theta - 3 \sin^2 \theta). & (b) V_0 (5 \cos^3 \theta - 3 \sin^2 \theta). \\ (c) \begin{cases} V_0 \cos \theta & \text{for the upper hemisphere,} \\ 0 & \text{for the lower hemisphere.} \end{cases} \end{array}$$

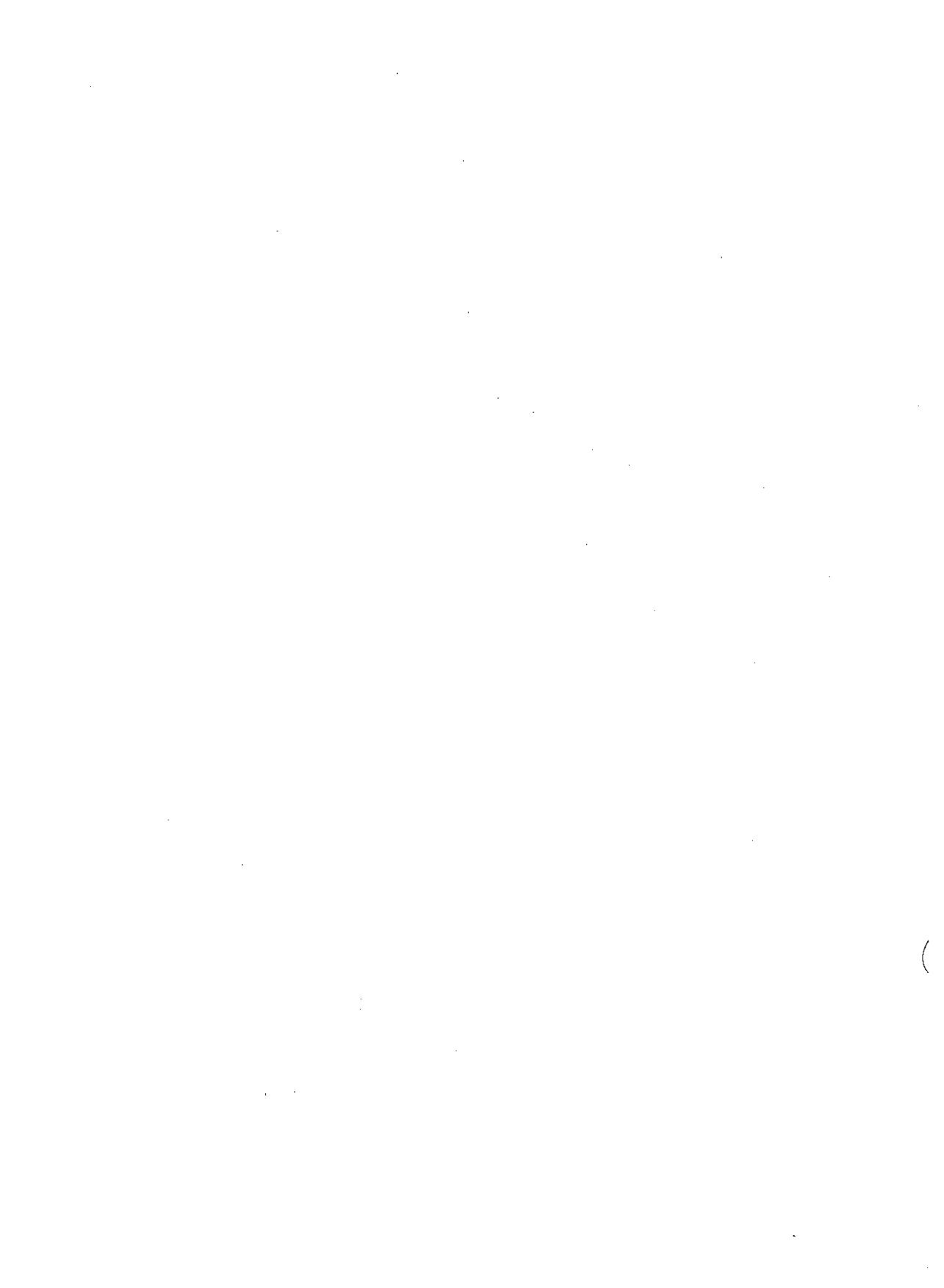
**19.26.** Find the steady-state temperature distribution inside a solid hemisphere of radius  $a$  if the curved surface is held at  $T_0$  and the flat surface at  $T = 0$ . Hint: Imagine completing the sphere and maintaining the lower hemisphere at a temperature such that the overall surface temperature distribution is an *odd* function about  $\theta = \pi/2$ .

**19.27.** Find the steady-state temperature distribution in a spherical shell of inner radius  $R_1$  and outer radius  $R_2$  when the inner surface has a temperature  $T_1$  and the outer surface a temperature  $T_2$ .

## Additional Reading

1. Jackson, J. *Classical Electrodynamics*, 2nd ed., Wiley, 1975. The classic textbook on electromagnetism with many examples and problems on the solutions of Laplace's equation in different coordinate systems.

2. Mathews, J. and Walker, R. *Mathematical Methods of Physics*, 2nd ed., Benjamin, 1970.
3. Morse, P. and Feshbach, M. *Methods of Theoretical Physics*, McGraw-Hill, 1953.





Part VI

---

## Green's Functions



## Green's Functions in One Dimension

Our treatment of differential equations, with the exception of SOLDEs with constant coefficients, did not consider inhomogeneous equations. At this point, however, we can put into use one of the most elegant pieces of machinery in higher mathematics, Green's functions, to solve inhomogeneous differential equations.

This chapter addresses Green's functions in one dimension, that is, Green's functions of ordinary differential equations. Consider the ODE  $\mathbf{L}_x[u] = f(x)$  where  $\mathbf{L}_x$  is a linear differential operator. In the abstract Dirac notation this can be formally written as  $\mathbf{L}|u\rangle = |f\rangle$ . If  $\mathbf{L}$  has an inverse  $\mathbf{L}^{-1} \equiv \mathbf{G}$ , the solution can be formally written as  $|u\rangle = \mathbf{L}^{-1}|f\rangle = \mathbf{G}|f\rangle$ . Multiplying this by  $\langle x|$  and inserting  $\mathbf{1} = \int dy |y\rangle \langle y|$  between  $\mathbf{G}$  and  $|f\rangle$  gives

$$u(x) = \int dy G(x, y) w(y) f(y), \quad (20.1)$$

where the integration is over the range of definition of the functions involved. Once we know  $G(x, y)$ , Equation (20.1) gives the solution  $u(x)$  in an integral form. But how do we find  $G(x, y)$ ?

Sandwiching both sides of  $\mathbf{L}\mathbf{G} = \mathbf{1}$  between  $\langle x|$  and  $|y\rangle$  and using  $\mathbf{1} = \int dx' |x'\rangle \langle x'|$  between  $\mathbf{L}$  and  $\mathbf{G}$  yields  $\int dx' L(x, x') w(x') G(x', y) = \langle x|y\rangle = \delta(x - y)/w(x)$  if we use Equation (6.3). In particular, if  $\mathbf{L}$  is a local differential operator (see Section 16.1), then  $L(x, x') = [\delta(x - x')/w(x)]\mathbf{L}_x$ , and we obtain

$$\mathbf{L}_x G(x, y) = \frac{\delta(x - y)}{w(x)} \quad \text{or} \quad \mathbf{L}_x G(x, y) = \delta(x - y), \quad (20.2)$$

where the second equation makes the frequently used assumption that  $w(x) = 1$ .  $G(x, y)$  is called the **Green's function** (GF) for the differential operator (DO)  $\mathbf{L}_x$ .

differential equation  
for Green's function

Green's function

As discussed in Chapters 16 and 18,  $\mathbf{L}_x$  might not be defined for all functions on  $\mathbb{R}$ . Moreover, a complete specification of  $\mathbf{L}_x$  requires some initial (or boundary) conditions. Therefore, we expect  $G(x, y)$  to depend on such initial conditions as well. We note that when  $\mathbf{L}_x$  is applied to (20.1), we get

$$\mathbf{L}_x u(x) = \int dy [\mathbf{L}_x(G(x, y))] w(y) f(y) = \int dy \frac{\delta(x-y)}{w(x)} w(y) f(y) = f(x),$$

indicating that  $u(x)$  is indeed a solution of the original ODE. Equation (20.2), involving the generalized function  $\delta(x-y)$  (or distribution in the language of Chapter 6), is meaningful only in the same context. Thus, we treat  $G(x, y)$  not as an ordinary function but as a *distribution*. Finally, (20.1) is assumed to hold for an arbitrary (well-behaved) function  $f$ .

## 20.1 Calculation of Some Green's Functions

This section presents some examples of calculating  $G(x, y)$  for very simple DOs. Later we will see how to obtain Green's functions for a general second-order linear differential operator. Although the complete specification of GFs requires boundary conditions, we shall introduce unspecified constants in some of the examples below, and calculate some *indefinite* GFs.

**20.1.1. Example.** Let us find the GF for the simplest DO,  $\mathbf{L}_x = d/dx$ . We need to find a distribution such that its derivative is the Dirac delta function:<sup>1</sup>  $G'(x, y) = \delta(x-y)$ .

In Chapter 6, we encountered such a distribution—the step function  $\theta(x-y)$ . Thus,  $G(x, y) = \theta(x-y) + \alpha(y)$ , where  $\alpha(y)$  is the “constant” of integration. ■

The example above did not include a boundary (or initial) condition. Let us see how boundary conditions affect the resulting GF.

**20.1.2. Example.** Let us solve  $u'(x) = f(x)$  where  $x \in [0, \infty)$  and  $u(0) = 0$ . A general solution of this DE is given by Equation (20.1) and the preceding example:

$$u(x) = \int_0^\infty \theta(x-y) f(y) dy + \int_0^\infty \alpha(y) f(y) dy.$$

The factor  $\theta(x-y)$  in the first term on the RHS chops off the integral at  $x$ :

$$u(x) = \int_0^x f(y) dy + \int_0^\infty \alpha(y) f(y) dy.$$

The BC gives  $0 = u(0) = 0 + \int_0^\infty \alpha(y) f(y) dy$ . The only way that this can be satisfied for arbitrary  $f(y)$  is for  $\alpha(y)$  to be zero. Thus,  $G(x, y) = \theta(x-y)$ , and

$$u(x) = \int_0^\infty \theta(x-y) f(y) dy = \int_0^x f(y) dy.$$

<sup>1</sup>Here and elsewhere in this chapter, a prime over a GF indicates differentiation with respect to its first argument.

This is killing a fly with a sledgehammer! We could have obtained the result by a simple integration. However, the roundabout way outlined here illustrates some important features of GFs that will be discussed later. The BC introduced here is very special. What happens if it is changed to  $u(0) = a$ ? Problem 20.1 answers that. ■

**20.1.3. Example.** A more complicated DO is  $L_x = d^2/dx^2$ . Let us find its indefinite GF. To do so, we integrate  $G''(x, y) = \delta(x - y)$  once with respect to  $x$  to obtain  $\frac{d}{dx}G(x, y) = \theta(x - y) + \alpha(y)$ . A second integration yields  $G(x, y) = \int dx\theta(x - y) + x\alpha(y) + \eta(y)$ , where  $\alpha$  and  $\eta$  are arbitrary functions and the integral is an indefinite integral to be evaluated next.

Let  $\Omega(x, y)$  be the primitive of  $\theta(x - y)$ ; that is,

$$\frac{d\Omega}{dx} = \theta(x - y) = \begin{cases} 1 & \text{if } x > y, \\ 0 & \text{if } x < y. \end{cases} \quad (20.3)$$

The solution to this equation is

$$\Omega(x, y) = \begin{cases} x + a(y) & \text{if } x > y, \\ b(y) & \text{if } x < y. \end{cases}$$

Note that we have not defined  $\Omega(x, y)$  at  $x = y$ . It will become clear below that  $\Omega(x, y)$  is continuous at  $x = y$ . It is convenient to write  $\Omega(x, y)$  as

$$\Omega(x, y) = [x + a(y)]\theta(x - y) + b(y)\theta(y - x). \quad (20.4)$$

To specify  $a(y)$  and  $b(y)$  further, we differentiate (20.4) and compare it with (20.3):

$$\begin{aligned} \frac{d\Omega}{dx} &= \theta(x - y) + [x + a(y)]\delta(x - y) - b(y)\delta(x - y) \\ &= \theta(x - y) + [x - b(y) + a(y)]\delta(x - y), \end{aligned} \quad (20.5)$$

where we have used

$$\frac{d}{dx}\theta(x - y) = -\frac{d}{dx}\theta(y - x) = \delta(x - y).$$

For Equation (20.5) to agree with (20.3), we must have  $[x - b(y) + a(y)]\delta(x - y) = 0$ , which, upon integration over  $x$ , yields  $a(y) - b(y) = -y$ . Substituting this in the expression for  $\Omega(x, y)$  gives

$$\Omega(x, y) = (x - y)\theta(x - y) + b(y)[\theta(x - y) + \theta(y - x)].$$

But  $\theta(x) + \theta(-x) = 1$ ; therefore,  $\Omega(x, y) = (x - y)\theta(x - y) + b(y)$ . It follows, among other things, that  $\Omega(x, y)$  is continuous at  $x = y$ . We can now write

$$G(x, y) = (x - y)\theta(x - y) + x\alpha(y) + \beta(y),$$

where  $\beta(y) = \eta(y) + b(y)$ . ■

The GF in the example above has two arbitrary functions,  $\alpha(y)$  and  $\beta(y)$ , which are the result of underspecification of  $L_x$ : A full specification of  $L_x$  requires BCs, as the following example shows.

**20.1.4. Example.** Let us calculate the GF of  $\mathbf{L}_x[u] = u''(x) = f(x)$  subject to the BC  $u(a) = u(b) = 0$  where  $[a, b]$  is the interval on which  $\mathbf{L}_x$  is defined. Example 20.1.3 gives us the (indefinite) GF for  $\mathbf{L}_x$ . Using that, we can write

$$\begin{aligned} u(x) &= \int_a^b (x-y)\theta(x-y)f(y) dy + x \int_a^b \alpha(y)f(y) dy + \int_a^b \beta(y)f(y) dy \\ &= \int_a^x (x-y)f(y) dy + x \int_a^b \alpha(y)f(y) dy + \int_a^b \beta(y)f(y) dy. \end{aligned}$$

Applying the BCs yields

$$\begin{aligned} 0 = u(a) &= a \int_a^b \alpha(y)f(y) dy + \int_a^b \beta(y)f(y) dy, \\ 0 = u(b) &= \int_a^b (b-y)f(y) dy + b \int_a^b \alpha(y)f(y) dy + \int_a^b \beta(y)f(y) dy. \end{aligned} \quad (20.6)$$

From these two relations it is possible to determine  $\alpha(y)$  and  $\beta(y)$ : Substitute for the last integral on the RHS of the second equation of (20.6) from the first equation and get  $0 = \int_a^b [b-y + b\alpha(y) - a\alpha(y)]f(y) dy$ . Since this must hold for arbitrary  $f(y)$ , we conclude that

$$b-y + (b-a)\alpha(y) = 0 \Rightarrow \alpha(y) = -\frac{b-y}{b-a}.$$

Substituting for  $\alpha(y)$  in the first equation of (20.6) and noting that the result holds for arbitrary  $f$ , we obtain  $\beta(y) = a(b-y)/(b-a)$ . Insertion of  $\alpha(y)$  and  $\beta(y)$  in the expression for  $G(x, y)$  obtained in Example 20.1.3 gives

$$G(x, y) = (x-y)\theta(x-y) + (x-a)\frac{y-b}{b-a} \quad \text{where } a \leq x \text{ and } y \leq b.$$

It is striking that  $G(a, y) = (a-y)\theta(a-y) = 0$  (because  $a-y \leq 0$ ), and

$$G(b, y) = (b-y)\theta(b-y) + (b-a)\frac{y-b}{b-a} = 0$$

because  $\theta(b-y) = 1$  for all  $y \leq b$  [recall that  $x$  and  $y$  lie in the interval  $(a, b)$ ]. These two equations reveal the important fact that as a function of  $x$ ,  $G(x, y)$  satisfies the same (homogeneous) BC as the solution of the DE. This is a general property that will be discussed later. ■

In all the preceding examples, the BCs were very simple. Specifically, the value of the solution and/or its derivative at the boundary points was zero. What if the BCs are not so simple? In particular, how can we handle a case where  $u(a)$  [or  $u'(a)$ ] and  $u(b)$  [or  $u'(b)$ ] are nonzero?

Consider a general (second-order) differential operator  $\mathbf{L}_x$  and the differential equation  $\mathbf{L}_x[u] = f(x)$  subject to the BCs  $u(a) = a_1$  and  $u(b) = b_1$ . We claim that we can reduce this system to the case where  $u(a) = u(b) = 0$ . Recall from Chapter 13 that the most general solution to such a DE is of the form  $u = u_h + u_i$  where  $u_h$ ,

the solution to the homogeneous equation, satisfies  $L_x[u_h] = 0$  and contains the arbitrary parameters inherent in solutions of differential equations. For instance, if the linearly independent solutions are  $v$  and  $w$ , then  $u_h(x) = C_1v(x) + C_2w(x)$  and  $u_i$  is any solution of the inhomogeneous DE.

If we demand that  $u_h(a) = a_1$  and  $u_h(b) = b_1$ , then  $u_i$  satisfies the system

$$L_x[u_i] = f(x), \quad u_i(a) = u_i(b) = 0,$$

which is of the type discussed in the preceding examples. Since  $L_x$  is a SOLDO, we can put all the machinery of Chapter 13 to work to obtain  $v(x)$ ,  $w(x)$ , and therefore  $u_h(x)$ . The problem then reduces to a DE for which the BCs are homogeneous; that is, the value of the solution and/or its derivative is zero at the boundary points.

**20.1.5. Example.** Let us assume that  $L_x = d^2/dx^2$ . Calculation of  $u_h$  is trivial:

$$L_x[u_h] = 0 \Rightarrow \frac{d^2u_h}{dx^2} = 0 \Rightarrow u_h(x) = C_1x + C_2.$$

To evaluate  $C_1$  and  $C_2$ , we impose the BCs  $u_h(a) = a_1$  and  $u_h(b) = b_1$ :

$$C_1a + C_2 = a_1,$$

$$C_1b + C_2 = b_1.$$

This gives  $C_1 = (b_1 - a_1)/(b - a)$  and  $C_2 = (a_1b - ab_1)/(b - a)$ .

The inhomogeneous equation defines a problem identical to that of Example 20.1.4. Thus, we can immediately write  $u_i(x) = \int_a^b G(x, y)f(y) dy$ , where  $G(x, y)$  is as given in that example. Thus, the general solution is

$$u(x) = \frac{b_1 - a_1}{b - a}x + \frac{a_1b - ab_1}{b - a} + \int_a^x (x - y)f(y) dy + \frac{x - a}{b - a} \int_a^b (y - b)f(y) dy \blacksquare$$

Example 20.1.5 shows that an inhomogeneous DE with inhomogeneous BCs can be separated into two DEs, one homogeneous with inhomogeneous BCs and the other inhomogeneous with homogeneous BCs, the latter being appropriate for the GF. Furthermore, all the foregoing examples indicate that solutions of DEs can be succinctly written in terms of GFs that automatically incorporate the BCs as long as the BCs are homogeneous. Can a GF also give the solution to a DE with inhomogeneous BCs?

## 20.2 Formal Considerations

The discussion and examples of the preceding section hint at the power of Green's functions. The elegance of such a function becomes apparent from the realization that it contains all the information about the solutions of a DE for any type of BCs, as we are about to show. Since GFs are inverses of DOs, let us briefly reexamine the inverse of an operator, which is closely tied to its spectrum. The question as

to whether or not an operator  $\mathbf{A}$  in a finite-dimensional vector space is invertible is succinctly answered by the value of its determinant:  $\mathbf{A}$  is invertible if and only if  $\det \mathbf{A} \neq 0$ . In fact, as we saw at the beginning of Chapter 16, one translates the abstract operator equation  $\mathbf{A} |u\rangle = |v\rangle$  into a matrix equation  $Au = v$  and reduces the question to that of the inverse of a matrix. This matrix takes on an especially simple form when  $\mathbf{A}$  is diagonal, that is, when  $A_{ij} = \lambda_i \delta_{ij}$ . For this special situation we have

$$\lambda_i u_i = v_i \quad \text{for } i = 1, 2, \dots, N \quad (\text{no sum over } i). \quad (20.7)$$

This equation has a unique solution (for arbitrary  $v_i$ ) if and only if  $\lambda_i \neq 0$  for all  $i$ . In that case  $u_i = v_i/\lambda_i$  for  $i = 1, 2, \dots, N$ . In particular, if  $v_i = 0$  for all  $i$ , that is, when Equation (20.7) is homogeneous, the unique solution is the trivial solution. On the other hand, when some of the  $\lambda_i$  are zero, there may be no solution to (20.7), but the homogeneous equation has a nontrivial solution ( $u_i$  need not be zero). Recalling (from Chapter 3) that an operator is invertible if and only if none of its eigenvalues is zero, we have the following:

**20.2.1. Proposition.** *The operator  $\mathbf{A} \in \mathcal{L}(\mathcal{V})$  is invertible if and only if the homogeneous equation  $\mathbf{A} |u\rangle = 0$  has no nontrivial solutions.*

In infinite-dimensional (Hilbert) spaces there is no determinant. How can we tell whether or not an operator in a Hilbert space is invertible? The exploitation of the connection between invertibility and eigenvalues has led to Proposition 20.2.1, which can be generalized to an operator acting on any vector space, finite or infinite. Consider the equation  $\mathbf{A} |u\rangle = 0$  in a Hilbert space  $\mathcal{H}$ . In general, neither the domain nor the range of  $\mathbf{A}$  is the whole of  $\mathcal{H}$ . If  $\mathbf{A}$  is invertible, then the only solution to the equation  $\mathbf{A} |u\rangle = 0$  is  $|u\rangle = 0$ . Conversely, assuming that the equation has no nontrivial solution implies that the null space of  $\mathbf{A}$  consists of only the zero vector. Thus,

$$\mathbf{A} |u_1\rangle = \mathbf{A} |u_2\rangle \Rightarrow \mathbf{A} (|u_1\rangle - |u_2\rangle) = 0 \Rightarrow |u_1\rangle - |u_2\rangle = 0.$$

This shows that  $\mathbf{A}$  is injective (one-to-one), i.e.,  $\mathbf{A}$  is a bijective linear mapping from the domain of  $\mathbf{A}$ ,  $\mathcal{D}(\mathbf{A})$ , onto the range of  $\mathbf{A}$ . Therefore,  $\mathbf{A}$  must have an inverse.

The foregoing discussion can be expressed as follows. If  $\mathbf{A} |u\rangle = 0$ , then (by the definition of eigenvectors)  $\lambda = 0$  is an eigenvalue of  $\mathbf{A}$  if and only if  $|u\rangle \neq 0$ . Thus, if  $\mathbf{A} |u\rangle = 0$  has no nontrivial solution, then zero cannot be an eigenvalue of  $\mathbf{A}$ . This can also be stated as follows:

**20.2.2. Theorem.** *An operator  $\mathbf{A}$  on a Hilbert space has an inverse if and only if  $\lambda = 0$  is not an eigenvalue of  $\mathbf{A}$ .*

Green's functions are inverses of differential operators. Therefore, it is important to have a clear understanding of the DOs. An  $n$ th-order linear differential operator (NOLDO) satisfies the following theorem (for a proof, see [Birk 78, Chapter 6]).



**20.2.3. Theorem.** *Let*

$$L_x = p_n(x) \frac{d^n}{dx^n} + p_{n-1}(x) \frac{d^{n-1}}{dx^{n-1}} + \dots + p_1(x) \frac{d}{dx} + p_0(x) \tag{20.8}$$

where  $p_n(x) \neq 0$  in  $[a, b]$ . Let  $x_0 \in [a, b]$  and let  $\{\gamma_k\}_{k=1}^n$  be given numbers and  $f(x)$  a given piecewise continuous function on  $[a, b]$ . Then the initial value problem (IVP)

$$\begin{aligned} L_x[u] &= f && \text{for } x \in [a, b], \\ u(x_0) &= \gamma_1, u'(x_0) = \gamma_2, \dots, u^{(n-1)}(x_0) = \gamma_n \end{aligned} \tag{20.9}$$

has one and only one solution.

initial value problem

This is simply the existence and uniqueness theorem for a NOLDE. Equation (20.9) is referred to as the **IVP with data**  $\{f(x); \gamma_1, \dots, \gamma_n\}$ . This theorem is used to define  $L_x$ . Part of that definition are the BCs that the solutions to  $L_x$  must satisfy.

A particularly important BC is the homogeneous one in which  $\gamma_1 = \gamma_2 = \dots = \gamma_n = 0$ . In such a case it can be shown (see Problem 20.3) that the only nontrivial solution of the homogeneous DE  $L_x[u] = 0$  is  $u \equiv 0$ . Theorem 20.2.2 then tells us that  $L_x$  is invertible; that is, there is a unique operator  $G$  such that  $LG = 1$ . The “components” version of this last relation is part of the content of the next theorem.

**20.2.4. Theorem.** *The DO  $L_x$  of Equation (20.8) associated with the IVP with data  $\{f(x); 0, 0, \dots, 0\}$  is invertible; that is, there exists a function  $G(x, y)$  such that*

$$L_x G(x, y) = \frac{\delta(x - y)}{w(x)}.$$

The importance of homogeneous BCs can now be appreciated. Theorem 20.2.4 is the reason why we had to impose homogeneous BCs to obtain the GF in all the examples of the previous section.

The BCs in (20.9) clearly are not the only ones that can be used. The most general linear BCs encountered in differential operator theory are

$$\begin{aligned} R_1[u] &\equiv \alpha_{11}u(a) + \dots + \alpha_{1n}u^{(n-1)}(a) + \beta_{11}u(b) + \dots + \beta_{1n}u^{(n-1)}(b) = \gamma_1, \\ R_2[u] &\equiv \alpha_{21}u(a) + \dots + \alpha_{2n}u^{(n-1)}(a) + \beta_{21}u(b) + \dots + \beta_{2n}u^{(n-1)}(b) = \gamma_2, \\ &\vdots \\ R_n[u] &\equiv \alpha_{n1}u(a) + \dots + \alpha_{nn}u^{(n-1)}(a) + \beta_{n1}u(b) + \dots + \beta_{nn}u^{(n-1)}(b) = \gamma_n. \end{aligned} \tag{20.10}$$

The  $n$  row vectors  $\{(\alpha_{i1}, \dots, \alpha_{in}, \beta_{i1}, \dots, \beta_{in})\}_{i=1}^n$  are assumed to be independent (in particular, no row is identical to zero). We refer to  $R_i$  as **boundary functionals** because for each (sufficiently smooth) function  $u$ , they give a number  $\gamma_i$ . The

boundary functionals  
and boundary value  
problem

DO of (20.8) and the BCs of (20.10) together form a **boundary value problem (BVP)**. The DE  $\mathbf{L}_x[u] = f$  subject to the BCs of (20.10) is a BVP with data  $\{f(x); \gamma_1, \dots, \gamma_n\}$ .

We note that the  $\mathbf{R}_i$  are linear; that is,

$$\mathbf{R}_i[u_1 + u_2] = \mathbf{R}_i[u_1] + \mathbf{R}_i[u_2] \quad \text{and} \quad \mathbf{R}_i[\alpha u] = \alpha \mathbf{R}_i[u].$$

Since  $\mathbf{L}_x$  is also linear, we conclude that the superposition principle applies to the system consisting of  $\mathbf{L}_x[u] = f$  and the BCs of (20.10), which is sometimes denoted by  $(\mathbf{L}; \mathbf{R}_1, \dots, \mathbf{R}_n)$ . If  $u$  satisfies the BVP with data  $\{f; \gamma_1, \dots, \gamma_n\}$  and  $v$  satisfies the BVP with data  $\{g; \mu_1, \dots, \mu_n\}$ , then  $\alpha u + \beta v$  satisfies the BVP with data  $\{\alpha f + \beta g; \alpha \gamma_1 + \beta \mu_1, \dots, \alpha \gamma_n + \beta \mu_n\}$ . It follows that if  $u$  and  $v$  both satisfy the BVP with data  $\{f; \gamma_1, \dots, \gamma_n\}$ , then  $u - v$  satisfies the BVP with data  $\{0; 0, 0, \dots, 0\}$ , which is called the **completely homogeneous problem**.

completely  
homogeneous  
problem

Unlike the IVP, the BVP with data  $\{0; 0, 0, \dots, 0\}$  may have a nontrivial solution. If the completely homogeneous problem has no nontrivial solution, then the BVP with data  $\{f; \gamma_1, \dots, \gamma_n\}$  has at most one solution (a solution exists for any set of data). On the other hand, if the completely homogeneous problem has nontrivial solutions, then the BVP with data  $\{f; \gamma_1, \dots, \gamma_n\}$  either has no solutions or has more than one solution (see [Stak 79, pp. 203–204]).

Recall that when a differential (unbounded) operator  $\mathbf{L}_x$  acts in a Hilbert space, such as  $\mathcal{L}_w^2(a, b)$ , it acts only on its domain. In the context of the present discussion, this means that not all functions in  $\mathcal{L}_w^2(a, b)$  satisfy the BCs necessary for defining  $\mathbf{L}_x$ . Thus, the functions for which the operator is defined (those that satisfy the BCs) form a subset of  $\mathcal{L}_w^2(a, b)$ , which we called the domain of  $\mathbf{L}_x$  and denoted by  $\mathcal{D}(\mathbf{L}_x)$ . From a formal standpoint it is important to distinguish among maps that have different domains. For instance, the Hilbert–Schmidt integral operators, which are defined on a finite interval, are compact, while those defined on the entire real line are not.

**20.2.5. Definition.** Let  $\mathbf{L}_x$  be the DO of Equation (20.8). Suppose there exists a DO  $\mathbf{L}_x^\dagger$ , with the property that

adjoint of a  
differential operator

$$w \{v^*(\mathbf{L}_x[u]) - u(\mathbf{L}_x^\dagger[v])^*\} = \frac{d}{dx} Q[u, v^*] \quad \text{for } u, v \in \mathcal{D}(\mathbf{L}_x) \cap \mathcal{D}(\mathbf{L}_x^\dagger),$$

conjunct

where  $Q[u, v^*]$ , called the **conjunct** of the functions  $u$  and  $v$ , depends on  $u, v$ , and their derivatives of order up to  $n - 1$ . The DO  $\mathbf{L}_x^\dagger$  is then called the **formal adjoint** of  $\mathbf{L}_x$ . If  $\mathbf{L}_x^\dagger = \mathbf{L}_x$  (without regard to the BCs imposed on their solutions), then  $\mathbf{L}_x$  is said to be **formally self-adjoint**. If  $\mathcal{D}(\mathbf{L}_x^\dagger) \supset \mathcal{D}(\mathbf{L}_x)$  and  $\mathbf{L}_x^\dagger = \mathbf{L}_x$  on  $\mathcal{D}(\mathbf{L}_x)$ , then  $\mathbf{L}_x$  is said to be **hermitian**. If  $\mathcal{D}(\mathbf{L}_x^\dagger) = \mathcal{D}(\mathbf{L}_x)$  and  $\mathbf{L}_x^\dagger = \mathbf{L}_x$ , then  $\mathbf{L}_x$  is said to be **self-adjoint**.

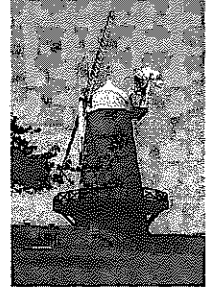
generalized Green's  
identity

The relation given in the definition above involving the conjunct is a generalization of the Lagrange identity and can also be written in integral form:

$$\int_a^b dxw\{v^*(L_x[u])\} - \int_a^b dxw\{u(L_x^\dagger[v])^*\} = Q[u, v^*]_a^b \quad (20.11)$$

This form is sometimes called the **generalized Green's identity**.

**George Green** (1793?–1841) was not appreciated in his lifetime. His date of birth is unknown (however, it is known that he was baptized on 14 July 1793), and no portrait of him survives. He left school, after only one year's attendance, to work in his father's bakery. When the father opened a windmill in Nottingham, the boy used an upper room as a study in which he taught himself physics and mathematics from library books. In 1828, when he was thirty-five years old, he published his most important work, *An Essay on the Application of Mathematical Analysis to the Theory of Electricity and Magnetism* at his own expense. In it Green apologized for any shortcomings in the paper due to his minimal formal education or the limited resources available to him, the latter being apparent in the few previous works he cited. The introduction explained the importance Green placed on the "potential" function. The body of the paper generalizes this idea to electricity and magnetism.



In addition to the physics of electricity and magnetism, Green's first paper also contained the monumental mathematical contributions for which he is now famous: The relationship between surface and volume integrals we now call *Green's theorem*, and the *Green's function*, a ubiquitous solution to partial differential equations in almost every area of physics. With little appreciation for the future impact of this work, one of Green's contemporaries declared the publication "a complete failure." The "Essay", which received little notice because of poor circulation, was saved by Lord Kelvin, who tracked it down in a German journal.

When his father died in 1829, some of George's friends urged him to seek a college education. After four years of self-study, during which he closed the gaps in his elementary education, Green was admitted to Caius College of Cambridge University at the age of 40, from which he graduated four years later after a disappointing performance on his final examinations. Later, however, he was appointed Perce Fellow of Caius College. Two years after his appointment he died, and his famous 1828 paper was republished, this time reaching a much wider audience. This paper has been described as "the beginning of mathematical physics in England."

He published only ten mathematical works. In 1833 he wrote three further papers. Two on electricity were published by the Cambridge Philosophical Society. One on hydrodynamics was published by the Royal Society of Edinburgh (of which he was a Fellow) in 1836. He also had two papers on hydrodynamics (in particular wave motion in canals), two papers on reflection and refraction of light, and two papers on reflection and refraction of sound published in Cambridge.

In 1923 the Green windmill was partially restored by a local businessman as a gesture of tribute to Green. Einstein came to pay homage. Then a fire in 1947 destroyed the renovations. Thirty years later the idea of a memorial was once again mooted, and sufficient money was raised to purchase the mill and present it to the sympathetic Nottingham City Council. In 1980 the George Green Memorial Appeal was launched to secure \$20,000 to get the sails

turning again and the machinery working once more. Today, Green's restored mill stands as a mathematics museum in Nottingham.

### 20.2.1 Second-Order Linear DOs

Since second-order linear differential operators (SOLDOs) are sufficiently general for most physical applications, we will concentrate on them. Because homogeneous BCs are important in constructing Green's functions, let us first consider BCs of the form

$$\begin{aligned} \mathbf{R}_1[u] &\equiv \alpha_{11}u(a) + \alpha_{12}u'(a) + \beta_{11}u(b) + \beta_{12}u'(b) = 0, \\ \mathbf{R}_2[u] &\equiv \alpha_{21}u(a) + \alpha_{22}u'(a) + \beta_{21}u(b) + \beta_{22}u'(b) = 0, \end{aligned} \quad (20.12)$$

where it is assumed, as usual, that  $(\alpha_{11}, \alpha_{12}, \beta_{11}, \beta_{12})$  and  $(\alpha_{21}, \alpha_{22}, \beta_{21}, \beta_{22})$  are linearly independent.

If we define the inner product as an integral with weight  $w$ , Equation (20.11) can be formally written as

$$\langle v | \mathbf{L} | u \rangle = \langle u | \mathbf{L}^\dagger | v \rangle^* + Q[u, v^*]_a^b.$$

This would coincide with the usual definition of the adjoint if the surface term vanishes, that is, if

$$Q[u, v^*]_{x=b} = Q[u, v^*]_{x=a}. \quad (20.13)$$

For this to happen, we need to impose BCs on  $v$ . To find these BCs, let us rewrite Equation (20.12) in a more compact form. Linear independence of the two row vectors of coefficients implies that the  $2 \times 4$  matrix of coefficients has rank two. This means that the  $2 \times 4$  matrix has an invertible  $2 \times 2$  submatrix. By rearranging the terms in Equation (20.12) if necessary, we can assume that the second of the two  $2 \times 2$  submatrices is invertible. The homogeneous BCs can then be conveniently written as

$$\mathbf{R}[u] = \begin{pmatrix} \mathbf{R}_1[u] \\ \mathbf{R}_2[u] \end{pmatrix} = (\mathbf{A} \quad \mathbf{B}) \begin{pmatrix} u_a \\ u_b \end{pmatrix} = \mathbf{A}u_a + \mathbf{B}u_b = 0, \quad (20.14)$$

where

$$\mathbf{A} \equiv \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}, \quad \mathbf{B} \equiv \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}, \quad u_a \equiv \begin{pmatrix} u(a) \\ u'(a) \end{pmatrix}, \quad u_b \equiv \begin{pmatrix} u(b) \\ u'(b) \end{pmatrix},$$

and  $\mathbf{B}$  is invertible.

The most general form of the conjunct for a SOLDO is

$$Q[u, v^*](x) = q_{11}(x)u(x)v^*(x) + q_{12}(x)u(x)v'^*(x) + q_{21}(x)u'(x)v^*(x) + q_{22}(x)u'(x)v'^*(x),$$

which can be written in matrix form as

$$Q[u, v^*](x) = u_x^t Q_x v_x^* \quad \text{where} \quad Q_x = \begin{pmatrix} q_{11}(x) & q_{12}(x) \\ q_{21}(x) & q_{22}(x) \end{pmatrix}, \quad (20.15)$$

and  $u_x$  and  $v_x^*$  have similar definitions as  $u_a$  and  $u_b$  above. The vanishing of the surface term becomes

$$u_b^t Q_b v_b^* = u_a^t Q_a v_a^*. \quad (20.16)$$

We need to translate this equation into a condition on  $v^*$  alone.<sup>2</sup> This is accomplished by solving for two of the four quantities  $u(a)$ ,  $u'(a)$ ,  $u(b)$ , and  $u'(b)$  in terms of the other two, substituting the result in Equation (20.16), and setting the coefficients of the other two equal to zero. Let us assume, as before, that the submatrix  $B$  is invertible, i.e.,  $u(b)$  and  $u'(b)$  are expressible in terms of  $u(a)$  and  $u'(a)$ . Then  $u_b = -B^{-1}Au_a$ , or  $u_b^t = -u_a^t A^t (B^t)^{-1}$ , and we obtain

$$-u_a^t A^t (B^t)^{-1} Q_b v_b^* = u_a^t Q_a v_a^* \Rightarrow u_a^t [A^t (B^t)^{-1} Q_b v_b^* + Q_a v_a^*] = 0,$$

and the condition on  $v^*$  becomes

$$A^t (B^t)^{-1} Q_b v_b^* + Q_a v_a^* = 0. \quad (20.17)$$

We see that all factors of  $u$  have disappeared, as they should. The expanded version of the BCs on  $v^*$  are written as

$$\begin{aligned} B_1[v^*] &\equiv \sigma_{11}v^*(a) + \sigma_{12}v'^*(a) + \eta_{11}v^*(b) + \eta_{12}v'^*(b) = 0, \\ B_2[v^*] &\equiv \sigma_{21}v^*(a) + \sigma_{22}v'^*(a) + \eta_{21}v^*(b) + \eta_{22}v'^*(b) = 0. \end{aligned} \quad (20.18)$$

adjoint boundary conditions

These homogeneous BCs are said to be **adjoint** to those of (20.12). Because of the difference between BCs and their adjoints, the domain of a differential operator need not be the same as that of its adjoint.

**20.2.6. Example.** Let  $L_x = d^2/dx^2$  with the homogeneous BCs

$$R_1[u] = au(a) - u'(a) = 0 \quad \text{and} \quad R_2[u] = \beta u(b) - u'(b) = 0. \quad (20.19)$$

We want to calculate  $Q[u, v^*]$  and the adjoint BCs for  $v$ . By repeated integration by parts [or by using Equation (13.23)], we obtain  $Q[u, v^*] = u'v^* - uv'^*$ . For the surface term to vanish, we must have

$$u'(a)v^*(a) - u(a)v'^*(a) = u'(b)v^*(b) - u(b)v'^*(b).$$

<sup>2</sup>The boundary conditions on  $v^*$  should not depend on the choice of  $u$ .

Substituting from (20.19) in this equation, we get

$$u(a)[\alpha v^*(a) - v'^*(a)] = u(b)[\beta v^*(b) - v'^*(b)],$$

which holds for arbitrary  $u$  if and only if

$$\mathbf{B}_1[v^*] = \alpha v^*(a) - v'^*(a) = 0 \quad \text{and} \quad \mathbf{B}_2[v^*] = \beta v^*(b) - v'^*(b) = 0. \quad (20.20)$$

This is a special case, in which the adjoint BCs are the same as the original BCs (substitute  $u$  for  $v^*$  to see this).

To see that the original BCs and their adjoints need not be the same, we consider

$$\mathbf{R}_1[u] = u'(a) - \alpha u(b) = 0 \quad \text{and} \quad \mathbf{R}_2[u] = \beta u(a) - u'(b) = 0, \quad (20.21)$$

from which we obtain  $u(a)[\beta v^*(b) + v'^*(a)] = u(b)[\alpha v^*(a) + v'^*(b)]$ . Thus,

$$\mathbf{B}_1[v^*] = \alpha v^*(a) + v'^*(b) = 0 \quad \text{and} \quad \mathbf{B}_2[v^*] = \beta v^*(b) + v'^*(a) = 0, \quad (20.22)$$

mixed and unmixed  
BCs

which is not the same as (20.21). Boundary conditions such as those in (20.19) and (20.20), in which each equation contains the function and its derivative evaluated at the same point, are called **unmixed BCs**. On the other hand, (20.21) and (20.22) are mixed BCs. ■

## 20.2.2 Self-Adjoint SOLDOs

In Chapter 13, we showed that a SOLDO satisfies the generalized Green's identity with  $w(x) = 1$ . In fact, since  $u$  and  $v$  are real, Equation (13.24) is identical to (20.11) if we set  $w = 1$  and

$$Q[u, v] = p_2 v u' - (p_2 v)' u + p_1 u v. \quad (20.23)$$

Also, we have seen that any SOLDO can be made (formally) self-adjoint. Thus, let us consider the formally self-adjoint SOLDO

$$\mathbf{L}_x = \mathbf{L}_x^\dagger = \frac{d}{dx} \left( p \frac{d}{dx} \right) + q$$

where both  $p(x)$  and  $q(x)$  are real functions and the inner product is defined with weight  $w = 1$ . If we are interested in formally self-adjoint operators with respect to a general weight  $w > 0$ , we can construct them as follows. We first note that if  $\mathbf{L}_x$  is formally self-adjoint with respect to a weight of unity, then  $(1/w)\mathbf{L}_x$  is self-adjoint with respect to weight  $w$ . Next, we note that  $\mathbf{L}_x$  is formally self-adjoint for all functions  $q$ , in particular, for  $wq$ . Now we define

$$\mathbf{L}_x^{(w)} = \frac{d}{dx} \left( p \frac{d}{dx} \right) + qw$$

and note that  $\mathbf{L}_x^{(w)}$  is formally self-adjoint with respect to a weight of unity, and therefore

$$\mathbf{L}_x \equiv \frac{1}{w} \mathbf{L}_x^{(w)} = \frac{1}{w} \frac{d}{dx} \left( p \frac{d}{dx} \right) + q \quad (20.24)$$

is formally self-adjoint with respect to weight  $w(x) > 0$ .

For SOLDOS that are formally self-adjoint with respect to weight  $w$ , the conjunct given in (20.23) reduces to

$$Q[u, v] = p(x)w(x)(vu' - uv'). \tag{20.25}$$

Thus, the surface term in the generalized Green's identity vanishes if and only if

$$p(b)w(b)[v(b)u'(b) - u(b)v'(b)] = p(a)w(a)[v(a)u'(a) - u(a)v'(a)]. \tag{20.26}$$

common types of boundary conditions for a SOLDE

The DO becomes self-adjoint if  $u$  and  $v$  satisfy Equation (20.26) as well as the same BCs. It can easily be shown that the following four types of BCs on  $u(x)$  assure the validity of Equation (20.26) and therefore define a self-adjoint operator  $L_x$  given by (20.24):

1. The **Dirichlet** BCs:  $u(a) = u(b) = 0$
2. The **Neumann** BCs:  $u'(a) = u'(b) = 0$
3. **General unmixed** BCs:  $\alpha u(a) - u'(a) = \beta u(b) - u'(b) = 0$
4. **Periodic** BCs:  $u(a) = u(b)$  and  $u'(a) = u'(b)$

### 20.3 Green's Functions for SOLDOS

We are now in a position to find the Green's function for a SOLDOS. First, note that a complete specification of  $L_x$  requires not only knowledge of  $p_0(x)$ ,  $p_1(x)$ , and  $p_2(x)$ —its coefficient functions—but also knowledge of the BCs imposed on the solutions. The most general BCs for a SOLDOS are of the type given in Equation (20.10) with  $n = 2$ . Thus, to specify  $L_x$  uniquely, we consider the system ( $L$ ;  $R_1, R_2$ ) with data ( $f$ ;  $\gamma_1, \gamma_2$ ). This system defines a unique BVP:

$$L_x[u] = p_2(x)\frac{d^2u}{dx^2} + p_1(x)\frac{du}{dx} + p_0(x)u = f(x), \quad R_i[u] = \gamma_i, \quad i = 1, 2. \tag{20.27}$$

A necessary condition for  $L_x$  to be invertible is that the homogeneous DE  $L_x[u] = 0$  have only the trivial solution  $u = 0$ . For  $u = 0$  to be the *only* solution, it must be a solution. This means that it must meet all the conditions in Equation (20.27). In particular, since  $R_i$  are linear functionals of  $u$ , we must have  $R_i[0] = 0$ . This can be stated as follows:

**20.3.1. Lemma.** *A necessary condition for a second-order linear DO to be invertible is for its associated BCs to be homogeneous.*<sup>3</sup>

<sup>3</sup>The lemma applies to all linear DOs, not just second order ones.

Thus, to study Green's functions we must restrict ourselves to problems with homogeneous BCs. This at first may seem restrictive, since not all problems have homogeneous BCs. Can we solve the others by the Green's function method? The answer is yes, as will be shown later in this chapter.

The above discussion clearly indicates that the Green's function of  $L_x$ , being its "inverse," is defined only if we consider the system  $(L; R_1, R_2)$  with data  $(f; 0, 0)$ . If the Green's function exists, it must satisfy the DE of Theorem 20.2.4, in which  $L_x$  acts on  $G(x, y)$ . But part of the definition of  $L_x$  are the BCs imposed on the solutions. Thus, if the LHS of the DE is to make any sense,  $G(x, y)$  must also satisfy those same BCs. We therefore make the following definition:

formal definition of  
Green's function

**20.3.2. Definition.** *The Green's function of a DO  $L_x$  is a function  $G(x, y)$  that satisfies both the DE*

$$L_x G(x, y) = \frac{\delta(x - y)}{w(x)}$$

and, as a function of  $x$ , the homogeneous BCs  $R_i[G] = 0$  for  $i = 1, 2$  where the  $R_i$  are defined as in Equation (20.12).

It is convenient to study the Green's function for the adjoint of  $L_x$  simultaneously. Denoting this by  $g(x, y)$ , we have

$$L_x^\dagger g(x, y) = \frac{\delta(x - y)}{w(x)}, \quad B_i[g] = 0, \quad \text{for } i = 1, 2, \quad (20.28)$$

adjoint Green's  
function

where  $B_i$  are the boundary functionals adjoint to  $R_i$  and given in Equation (20.18). The function  $g(x, y)$  is known as the **adjoint Green's function** associated with the DE of (20.27).

We can now use (20.27) and (20.28) to find the solutions to

$$\begin{aligned} L_x[u] &= f(x), & R_i[u] &= 0 & \text{for } i = 1, 2, \\ L_x^\dagger[v] &= h(x), & B_i[v^*] &= 0 & \text{for } i = 1, 2. \end{aligned} \quad (20.29)$$

With  $v(x) = g(x, y)$  in Equation (20.11)—whose RHS is assumed to be zero—we get  $\int_a^b w g^*(x, y) L_x[u] dx = \int_a^b w u(x) (L_x^\dagger[g])^* dx$ . Using (20.28) on the RHS and (20.29) on the LHS, we obtain

$$u(y) = \int_a^b g^*(x, y) w(x) f(x) dx.$$

Similarly, with  $u(x) = G(x, y)$ , Equation (20.11) gives

$$v^*(y) = \int_a^b G(x, y) w(x) h^*(x) dx,$$



or, since  $w(x)$  is a (positive) real function,

$$v(y) = \int_a^b G^*(x, y)w(x)h(x) dx.$$

These equations for  $u(y)$  and  $v(y)$  are not what we expect [see, for instance, Equation (20.1)]. However, if we take into account certain properties of Green's functions that we will discuss next, these equations become plausible.

### 20.3.1 Properties of Green's functions

Let us rewrite the generalized Green's identity [Equation (20.11)], with the RHS equal to zero, as

$$\int_a^b dt w(t) \{v^*(t)(L_t[u])\} = \int_a^b dt w(t) \{u(t)(L_t^\dagger[v])^*\}. \quad (20.30)$$

Green's identity This is sometimes called **Green's identity**. Substituting  $G(t, y)$  for  $u(t)$  and  $g(t, x)$  for  $v(t)$  gives

$$\int_a^b dt w(t) g^*(t, x) \frac{\delta(t-y)}{w(t)} = \int_a^b dt w(t) G(t, y) \frac{\delta(t-x)}{w(t)},$$

or  $g^*(y, x) = G(x, y)$ . A consequence of this identity is that

**20.3.3. Box.**  $G(x, y)$  must satisfy the adjoint BCs with respect to its second argument.

If for the time being we assume that the Green's function associated with a system  $(L; R_1, R_2)$  is unique, then, since for a self-adjoint differential operator,  $L_x$  and  $L_x^\dagger$  are identical and  $u$  and  $v$  both satisfy the same BCs, we must have  $G(x, y) = g(x, y)$  or, using  $g^*(y, x) = G(x, y)$ , we get  $G(x, y) = G^*(y, x)$ . In particular, if the coefficient functions of  $L_x$  are all real,  $G(x, y)$  will be real, and we have  $G(x, y) = G(y, x)$ . This means that  $G$  is a symmetric function of its two arguments.

The last property is related to the continuity of  $G(x, y)$  and its derivative at  $x = y$ . For a SOLDO, we have

$$L_x G(x, y) = p_2(x) \frac{\partial^2 G}{\partial x^2} + p_1(x) \frac{\partial G}{\partial x} + p_0(x) G = \frac{\delta(x-y)}{w(x)},$$

where  $p_0$ ,  $p_1$ , and  $p_2$  are assumed to be real and continuous in the interval  $[a, b]$ , and  $w(x)$  and  $p_2(x)$  are assumed to be positive for all  $x \in [a, b]$ . We multiply both sides of the DE by

$$h(x) = \frac{\mu(x)}{p_2(x)}, \quad \text{where} \quad \mu(x) \equiv \exp \left[ \int_a^x \frac{p_1(t)}{p_2(t)} dt \right],$$

noting that  $d\mu/dx = (p_1/p_2)\mu$ . This transforms the DE into

$$\frac{\partial}{\partial x} \left[ \mu(x) \frac{\partial}{\partial x} G(x, y) \right] + \frac{p_0(x)\mu(x)}{p_2(x)} G(x, y) = \frac{\mu(y)}{p_2(y)w(y)} \delta(x - y).$$

Integrating this equation gives

$$\mu(x) \frac{\partial}{\partial x} G(x, y) + \int_a^x \frac{p_0(t)\mu(t)}{p_2(t)} G(t, y) dt = \frac{\mu(y)}{p_2(y)w(y)} \theta(x - y) + \alpha(y) \tag{20.31}$$

because the primitive of  $\delta(x - y)$  is  $\theta(x - y)$ . Here  $\alpha(y)$  is the "constant" of integration. First consider the case where  $p_0 = 0$ , for which the Green's function will be denoted by  $G_0(x, y)$ . Then Equation (20.31) becomes

$$\mu(x) \frac{\partial}{\partial x} G_0(x, y) = \frac{\mu(y)}{p_2(y)w(y)} \theta(x - y) + \alpha_1(y),$$

which (since  $\mu, p_2$ , and  $w$  are continuous on  $[a, b]$ , and  $\theta(x - y)$  has a discontinuity only at  $x = y$ ) indicates that  $\partial G_0/\partial x$  is continuous everywhere on  $[a, b]$  except at  $x = y$ . Now divide the last equation by  $\mu$  and integrate the result to get

$$G_0(x, y) = \frac{\mu(y)}{p_2(y)w(y)} \int_a^x \frac{\theta(t - y)}{\mu(t)} dt + \alpha_1(y) \int_a^x \frac{dt}{\mu(t)} + \alpha_2(y).$$

Every term on the RHS is continuous except possibly the integral involving the  $\theta$ -function. However, that integral can be written as

$$\int_a^x \frac{\theta(t - y)}{\mu(t)} dt = \theta(x - y) \int_y^x \frac{dt}{\mu(t)}. \tag{20.32}$$

The  $\theta$ -function in front of the integral is needed to ensure that  $a \leq y \leq x$  as demanded by the LHS of Equation (20.32). The RHS of Equation (20.32) is continuous at  $x = y$  with limit being zero as  $x \rightarrow y$ .

Next, we write  $G(x, y) = G_0(x, y) + H(x, y)$ , and apply  $\mathbf{L}_x$  to both sides. This gives

$$\begin{aligned} \frac{\delta(x - y)}{w(x)} &= \left( p_2 \frac{d^2}{dx^2} + p_1 \frac{d}{dx} \right) G_0 + p_0 G_0 + \mathbf{L}_x H(x, y) \\ &= \frac{\delta(x - y)}{w(x)} + p_0 G_0 + \mathbf{L}_x H(x, y), \end{aligned}$$

or  $p_2 H'' + p_1 H' + p_0 H = -p_0 G_0$ . The continuity of  $G_0, p_0, p_1$ , and  $p_2$  on  $[a, b]$  implies the continuity of  $H$ , because a discontinuity in  $H$  would entail a delta function discontinuity in  $dH/dx$ , which is impossible because there are no delta functions in the equation for  $H$ . Since both  $G_0$  and  $H$  are continuous,  $G$  must also be continuous on  $[a, b]$ .

We can now calculate the jump in  $\partial G/\partial x$  at  $x = y$ . We denote the jump as  $\Delta G'(y)$  and define it as follows:

$$\Delta G'(y) \equiv \lim_{\epsilon \rightarrow 0} \left[ \frac{\partial G}{\partial x}(x, y) \Big|_{x=y+\epsilon} - \frac{\partial G}{\partial x}(x, y) \Big|_{x=y-\epsilon} \right].$$

Dividing (20.31) by  $\mu(x)$  and taking the above limit for all terms, we obtain

$$\begin{aligned} \Delta G'(y) + \lim_{\epsilon \rightarrow 0} & \left[ \frac{1}{\mu(y + \epsilon)} \int_a^{y+\epsilon} \frac{p_0(t)\mu(t)}{p_2(t)} G(t, y) dt \right. \\ & \left. - \frac{1}{\mu(y - \epsilon)} \int_a^{y-\epsilon} \frac{p_0(t)\mu(t)}{p_2(t)} G(t, y) dt \right] \\ & = \frac{\mu(y)}{p_2(y)w(y)} \lim_{\epsilon \rightarrow 0} \left[ \overbrace{\frac{\theta(+\epsilon)}{\mu(y + \epsilon)}}{=1} - \overbrace{\frac{\theta(-\epsilon)}{\mu(y - \epsilon)}}{=0} \right]. \end{aligned}$$

The second term on the LHS is zero because all functions are continuous at  $y$ . The limit on the RHS is simply  $1/\mu(y)$ . We therefore obtain

$$\Delta G'(y) = \frac{1}{p_2(y)w(y)}. \tag{20.33}$$

### 20.3.2 Construction and Uniqueness of Green's Functions

We are now in a position to calculate the Green's function for a general SOLDOS and show that it is unique.

existence and uniqueness of GF for a second order linear differential operator

**20.3.4. Theorem.** Consider the system  $(\mathbf{L}; \mathbf{R}_1, \mathbf{R}_2)$  with data  $(f; 0, 0)$ , in which  $\mathbf{L}_x$  is a SOLDOS. If the homogeneous DE  $\mathbf{L}_x[u] = 0$  has no nontrivial solution, then the GF associated with the given system exists and is unique. The solution of the system is

$$u(x) = \int_a^b dy w(y) G(x, y) f(y)$$

and is also unique.

*Proof.* The GF satisfies the DE  $\mathbf{L}_x G(x, y) = 0$  for all  $x \in [a, b]$  except  $x = y$ . We thus divide  $[a, b]$  into two intervals,  $I_1 = [a, y)$  and  $I_2 = (y, b]$ , and note that a general solution to the above homogeneous DE can be written as a linear combination of a basis of solutions,  $u_1$  and  $u_2$ . Thus, we can write the solution of the DE as

$$\begin{aligned} G_l(x, y) &= c_1 u_1(x) + c_2 u_2(x) & \text{for } x \in I_1 \\ G_r(x, y) &= d_1 u_1(x) + d_2 u_2(x) & \text{for } x \in I_2 \end{aligned}$$

and define the GF as

$$G(x, y) = \begin{cases} G_l(x, y) & \text{if } x \in I_1, \\ G_r(x, y) & \text{if } x \in I_2, \end{cases} \quad (20.34)$$

where  $c_1, c_2, d_1$ , and  $d_2$  are, in general, functions of  $y$ . To determine  $G(x, y)$  we must determine four unknowns. We also have four relations: the continuity of  $G$ , the jump in  $\partial G/\partial x$  at  $x = y$ , and the two BCs  $\mathbf{R}_1[G] = \mathbf{R}_2[G] = 0$ . The continuity of  $G$  gives

$$c_1(y)u_1(y) + c_2(y)u_2(y) = d_1(y)u_1(y) + d_2(y)u_2(y).$$

The jump of  $\partial G/\partial x$  at  $x = y$  yields

$$c_1(y)u_1'(y) + c_2(y)u_2'(y) - d_1(y)u_1'(y) - d_2(y)u_2'(y) = -\frac{1}{p_2(y)w(y)}.$$

Introducing  $b_1 = c_1 - d_1$  and  $b_2 = c_2 - d_2$  changes the two preceding equations to

$$\begin{aligned} b_1u_1 + b_2u_2 &= 0, \\ b_1u_1' + b_2u_2' &= -\frac{1}{p_2w}. \end{aligned}$$

These equations have a unique solution iff

$$\det \begin{pmatrix} u_1 & u_2 \\ u_1' & u_2' \end{pmatrix} \neq 0.$$

But the determinant is simply the Wronskian of the two independent solutions and therefore cannot be zero. Thus,  $b_1(y)$  and  $b_2(y)$  are determined in terms of  $u_1, u_1', u_2, u_2', p_2$ , and  $w$ .

We now define

$$h(x, y) \equiv \begin{cases} b_1(y)u_1(x) + b_2(y)u_2(x) & \text{if } x \in I_1, \\ 0 & \text{if } x \in I_2. \end{cases}$$

so that  $G(x, y) = h(x, y) + d_1(y)u_1(x) + d_2(y)u_2(x)$ . We have reduced the number of unknowns to two,  $d_1$  and  $d_2$ . Imposing the BCs gives two more relations:

$$\begin{aligned} \mathbf{R}_1[G] &= \mathbf{R}_1[h] + d_1\mathbf{R}_1[u_1] + d_2\mathbf{R}_1[u_2] = 0, \\ \mathbf{R}_2[G] &= \mathbf{R}_2[h] + d_1\mathbf{R}_2[u_1] + d_2\mathbf{R}_2[u_2] = 0. \end{aligned}$$

Can we solve these equations and determine  $d_1$  and  $d_2$  uniquely? We can, if

$$\det \begin{pmatrix} \mathbf{R}_1[u_1] & \mathbf{R}_1[u_2] \\ \mathbf{R}_2[u_1] & \mathbf{R}_2[u_2] \end{pmatrix} \neq 0.$$

It can be shown that this determinant is nonzero (see Problem 20.5).

Having found the unique  $\{b_i, d_i\}_{i=1}^2$ , we can calculate  $c_i$  uniquely, substitute all of them in Equation (20.34), and obtain the unique  $G(x, y)$ . That  $u(x)$  is also unique can be shown similarly.  $\square$

**20.3.5. Example.** Let us calculate the GF for  $\mathbf{L}_x = d^2/dx^2$  with BCs  $u(a) = u(b) = 0$ . We note that  $\mathbf{L}_x[u] = 0$  with the given BCs has no nontrivial solution (verify this). Thus, the GF exists. The DE for  $G(x, y)$  is  $G'' = 0$  for  $x \neq y$ , whose solutions are

$$G(x, y) = \begin{cases} c_1x + c_2 & \text{if } a \leq x < y, \\ d_1x + d_2 & \text{if } y < x \leq b. \end{cases} \quad (20.35)$$

Continuity at  $x = y$  gives  $c_1y + c_2 = d_1y + d_2$  or  $b_1y + b_2 = 0$  with  $b_i = c_i - d_i$ . The discontinuity of  $dG/dx$  at  $x = y$  gives

$$d_1 - c_1 = \frac{1}{p_2w} = 1 \Rightarrow b_1 = -1$$

assuming that  $w = 1$ . From the equations above we also get  $b_2 = y$ .  $G(x, y)$  must also satisfy the given BCs. Thus,  $G(a, y) = 0 = G(b, y)$ . Since  $a \leq y$  and  $b \geq y$ , we obtain  $c_1a + c_2 = 0$  and  $d_1b + d_2 = 0$ , or, after substituting  $c_i = b_i + d_i$ ,

$$ad_1 + d_2 = a - y, \quad bd_1 + d_2 = 0.$$

The solution to these equations is  $d_1 = (y - a)/(b - a)$  and  $d_2 = -b(y - a)/(b - a)$ . With  $b_1, b_2, d_1$ , and  $d_2$  as given above, we find

$$c_1 = b_1 + d_1 = -\frac{b - y}{b - a} \quad \text{and} \quad c_2 = b_2 + d_2 = a\frac{b - y}{b - a}.$$

Writing Equation (20.35) as

$$G(x, y) = (c_1x + c_2)\theta(y - x) + (d_1x + d_2)\theta(x - y)$$

and using the identity  $\theta(y - x) = 1 - \theta(x - y)$ , we get

$$G(x, y) = c_1x + c_2 - (b_1x + b_2)\theta(x - y).$$

Using the values found for the  $b$ 's and  $c$ 's, we obtain

$$G(x, y) = (a - x)\left(\frac{b - y}{b - a}\right) + (x - y)\theta(x - y),$$

which is the same as the GF obtained in Example 20.1.4.  $\blacksquare$

**20.3.6. Example.** Let us find the GF for  $\mathbf{L}_x = d^2/dx^2 + 1$  with the BCs  $u(0) = u(\pi/2) = 0$ . The general solution of  $\mathbf{L}_x[u] = 0$  is

$$u(x) = A \sin x + B \cos x.$$

If the BCs are imposed, we get  $u = 0$ . Thus,  $G(x, y)$  exists. The general form of  $G(x, y)$  is

$$G(x, y) = \begin{cases} c_1 \sin x + c_2 \cos x & \text{if } 0 \leq x < y, \\ d_1 \sin x + d_2 \cos x & \text{if } y < x \leq \pi/2. \end{cases} \quad (20.36)$$

Continuity of  $G$  at  $x = y$  gives  $b_1 \sin y + b_2 \cos y = 0$  with  $b_i = c_i - d_i$ . The discontinuity of the derivative of  $G$  at  $x = y$  gives  $b_1 \cos y - b_2 \sin y = -1$ , where we have set  $w(x) = 1$ . Solving these equations yields  $b_1 = -\cos y$  and  $b_2 = \sin y$ . The BCs give

$$G(0, y) = 0 \Rightarrow c_2 = 0 \Rightarrow d_2 = -b_2 = -\sin y,$$

$$G(\pi/2, y) = 0 \Rightarrow d_1 = 0 \Rightarrow c_1 = -b_1 = -\cos y.$$

Substituting in Equation (20.36) gives

$$G(x, y) = \begin{cases} -\cos y \sin x & \text{if } x < y, \\ -\sin y \cos x & \text{if } y < x, \end{cases}$$

or, using the theta function,

$$G(x, y) = -\theta(y - x) \cos y \sin x - \theta(y - x) \sin y \cos x$$

$$= -[1 - \theta(x - y)] \cos y \sin x - \theta(x - y) \sin y \cos x$$

$$= -\cos y \sin x + \theta(x - y) \sin(x - y).$$

It is instructive to verify directly that  $G(x, y)$  satisfies  $L_x[G] = \delta(x - y)$ :

$$L_x[G] = -\cos y \underbrace{\left(\frac{d^2}{dx^2} + 1\right) \sin x}_{=0} + \left(\frac{d^2}{dx^2} + 1\right) [\theta(x - y) \sin(x - y)]$$

$$= \frac{d^2}{dx^2} [\theta(x - y) \sin(x - y)] + \theta(x - y) \sin(x - y)$$

$$= \frac{d}{dx} \underbrace{[\delta(x - y) \sin(x - y)]}_{=0} + \theta(x - y) \cos(x - y) + \theta(x - y) \sin(x - y).$$

The first term vanishes because the sine vanishes at the only point where the delta function is nonzero. Thus, we have

$$L_x[G] = [\delta(x - y) \cos(x - y) - \theta(x - y) \sin(x - y)] + \theta(x - y) \sin(x - y)$$

$$= \delta(x - y)$$

because the delta function demands that  $x = y$ , for which  $\cos(x - y) = 1$ . ■

The existence and uniqueness of the Green's function  $G(x, y)$  in conjunction with its properties and its adjoint, imply the existence and uniqueness of the adjoint Green's function  $g(x, y)$ . Using this fact, we can show that the condition for the absence of a nontrivial solution for  $L_x[u] = 0$  is also a necessary condition for the existence of  $G(x, y)$ . That is, if  $G(x, y)$  exists, then  $L_x[u] = 0$  implies that  $u = 0$ . Suppose  $G(x, y)$  exists; then  $g(x, y)$  also exists. In Green's identity let  $v = g(x, y)$ . This gives an identity:

$$\int_a^b w(x) g^*(x, y) (L_x[u]) dx = \int_a^b w(x) u(x) (L_x^\dagger[g])^* dx$$

$$= \int_a^b w(x) u(x) \frac{\delta(x - y)}{w(x)} dx = u(y).$$

In particular, if  $\mathbf{L}_x[u] = 0$ , then  $u(y) = 0$  for all  $y$ . We have proved the following result.

**20.3.7. Proposition.** *The DE  $\mathbf{L}_x[u] = 0$  implies that  $u \equiv 0$  if and only if the GF corresponding to  $\mathbf{L}_x$  and the homogeneous BCs exist.*

It is sometimes stated that the Green's function of a SOLDO with constant coefficients depends on the difference  $x - y$ . This statement is motivated by the observation that if  $u(x)$  is a solution of

$$\mathbf{L}_x[u] = a_2 \frac{d^2u}{dx^2} + a_1 \frac{du}{dx} + a_0u = f(x),$$

then  $u(x - y)$  is the solution of  $a_2u'' + a_1u' + a_0u = f(x - y)$  if  $a_0, a_1$ , and  $a_2$  are constant. Thus, if  $G(x)$  is a solution of  $\mathbf{L}_x[G] = \delta(x)$  [again assuming that  $w(x) = 1$ ], then it seems that the solution of  $\mathbf{L}_x[G] = \delta(x - y)$  is simply  $G(x - y)$ . This is clearly wrong, as Examples 20.3.5 and 20.3.6 showed. The reason is, of course, the BCs. The fact that  $G(x - y)$  satisfies the right DE does not guarantee that it also satisfies the right BCs. The following example, however, shows that the conjecture is true for a homogeneous *initial value problem*.

**20.3.8. Example.** The most general form for the GF is

$$G(x, y) = \begin{cases} c_1u_1(x) + c_2u_2(x) & \text{if } a \leq x < y, \\ d_1u_1(x) + d_2u_2(x) & \text{if } y < x \leq b. \end{cases}$$

The IVP condition  $G(a, y) = 0 = G'(a, y)$  implies

$$c_1u_1(a) + c_2u_2(a) = 0 \quad \text{and} \quad c_1u_1'(a) + c_2u_2'(a) = 0.$$

Linear independence of  $u_1$  and  $u_2$  implies

$$\det \begin{pmatrix} u_1(a) & u_2(a) \\ u_1'(a) & u_2'(a) \end{pmatrix} = W(a; u_1, u_2) \neq 0.$$

Hence,  $c_1 = c_2 = 0$  is the only solution. This gives

$$G(x, y) = \begin{cases} 0 & \text{if } a \leq x < y, \\ d_1u_1(x) + d_2u_2(x) & \text{if } y < x \leq b. \end{cases} \tag{20.37}$$

Continuity of  $G$  at  $x = y$  yields  $d_1u_1(y) + d_2u_2(y) = 0$ , while the discontinuity jump condition in the derivative gives  $d_1u_1'(y) + d_2u_2'(y) = 1$ . Solving these two equations, we get

$$d_1 = \frac{u_2(y)}{u_1'(y)u_2(y) - u_2'(y)u_1(y)}, \quad d_2 = -\frac{u_1(y)}{u_1'(y)u_2(y) - u_2'(y)u_1(y)}.$$

Substituting this in (20.37) gives

$$G(x, y) = \left[ \frac{u_2(y)u_1(x) - u_1(y)u_2(x)}{u_1'(y)u_2(y) - u_2'(y)u_1(y)} \right] \theta(x - y). \tag{20.38}$$

Equation (20.38) holds for any SOLDO with the given BCs. We now use the fact that the SOLDO has *constant coefficients*. In that case, we know the exact form of  $u_1$  and  $u_2$ . There are two cases to consider:

1. If the characteristic polynomial of  $L_x$  has two distinct roots  $\lambda_1$  and  $\lambda_2$ , then  $u_1(x) = e^{\lambda_1 x}$  and  $u_2(x) = e^{\lambda_2 x}$ . Writing  $\lambda_1 = a + b$  and  $\lambda_2 = a - b$  and substituting the exponential functions and their derivatives in Equation (20.38) yields

$$G(x, y) = \left[ \frac{e^{(a-b)y} e^{(a+b)x} - e^{(a+b)y} e^{(a-b)x}}{2be^{2ay}} \right] \theta(x - y) \\ = \frac{1}{2b} [e^{(a+b)(x-y)} - e^{(a-b)(x-y)}] \theta(x - y),$$

which is a function of  $x - y$  alone.

2. If  $\lambda_1 = \lambda_2 = \lambda$ , then  $u_1(x) = e^{\lambda x}$ ,  $u_2(x) = xe^{\lambda x}$ , and substitution of these functions in Equation (20.38) gives

$$G(x, y) = (x - y)e^{\lambda(x-y)}\theta(x - y). \quad \blacksquare$$

### 20.3.3 Inhomogeneous BCs

So far we have concentrated on problems with homogeneous BCs,  $R_i[u] = 0$ , for  $i = 1, 2$ . What if the BCs are inhomogeneous? It turns out that the Green's function method, even though it was derived for homogeneous BCs, solves this kind of problem as well! The secret of this success is the generalized Green's identity.

GF solves inhomogeneous BCs as well

Suppose we are interested in solving the DE

$$L_x[u] = f(x) \quad \text{with} \quad R_i[u] = \gamma_i \quad \text{for} \quad i = 1, 2,$$

and we have the GF for  $L_x$  (with homogeneous BCs, of course). We can substitute  $v = g(x, y) = G^*(y, x)$  in the generalized Green's identity and use the DE to obtain

$$\int_a^b w(x)G(y, x)f(x) dx - \int_a^b w(x)u(x)(L_x^\dagger[g])^* dx = Q[u, g^*(x, y)]|_{x=a}^{x=b},$$

or, using  $L_x^\dagger[g(x, y)] = \delta(x - y)/w(y)$ ,

$$u(y) = \int_a^b w(x)G(y, x)f(x) dx - Q[u, g^*(x, y)]|_{x=a}^{x=b}.$$

To evaluate the surface term, let us write the BCs in matrix form [see Equation (20.17)]:

$$Au_a + Bu_b = \gamma \Rightarrow u_b = B^{-1}\gamma - B^{-1}Au_a,$$

$$AG_a + BG_b = 0 \Rightarrow A'(B')^{-1}Q_b g_b^* + Q_a g_a^* = 0,$$



where  $\gamma$  is a column vector composed of  $\gamma_1$  and  $\gamma_2$ , and we have assumed that  $G(x, y)$  and  $g^*(x, y)$  satisfy, respectively, the homogeneous BCs (with  $\gamma = 0$ ) and their adjoints. We have also assumed that the  $2 \times 4$  matrix of coefficients has rank 2, and without loss of generality, let  $B$  be the invertible  $2 \times 2$  submatrix. Then, assuming the general form of the surface term as in Equation (20.15), we obtain

$$\begin{aligned} Q[u, g^*(x, y)]|_{x=a}^{x=b} &= u_b^t Q_b g_b^* - u_a^t Q_a g_a^* \\ &= (B^{-1}\gamma - B^{-1}Au_a)^t Q_b g_b^* - u_a^t Q_a g_a^* \\ &= \gamma^t (B^t)^{-1} Q_b g_b^* - u_a^t \underbrace{[A^t (B^t)^{-1} Q_b g_b^* + Q_a g_a^*]}_{= 0 \text{ because } g^* \text{ satisfies homogeneous adjoint BC}} \\ &= \gamma^t (B^t)^{-1} Q_b g_b^*, \end{aligned} \tag{20.39}$$

where

$$g_b^* \equiv \left( \begin{array}{c} g^*(b, y) \\ \frac{\partial}{\partial x} g^*(x, y)|_{x=b} \end{array} \right) = \left( \begin{array}{c} G(y, b) \\ \frac{\partial}{\partial x} G(y, x)|_{x=b} \end{array} \right).$$

It follows that  $Q[u, g^*(x, y)]|_{x=a}^{x=b}$  is given entirely in terms of  $G$ , its derivative, the coefficient functions of the DE (hidden in the matrix  $Q$ ), the homogeneous BCs (hidden in  $B$ ), and the constants  $\gamma_1$  and  $\gamma_2$ . The fact that  $g^*$  and  $\partial g^*/\partial x$  appear to be evaluated at  $x = b$  is due to the simplifying (but harmless) assumption that  $B$  is invertible, i.e., that  $u(b)$  and  $u'(b)$  can be written in terms of  $u(a)$  and  $u'(a)$ . Of course, this may not be possible; then we have to find another pair of the four quantities in terms of the other two, in which case the matrices and the vectors will change but the argument, as well as the conclusion, will remain valid. We can now write

$$u(y) = \int_a^b w(x)G(y, x)f(x) dx - \gamma^t M g^*, \tag{20.40}$$

where a general matrix  $M$  has been introduced, and the subscript  $b$  has been removed to encompass cases where submatrices other than  $B$  are invertible. Equation (20.40) shows that  $u$  can be determined completely once we know  $G(x, y)$ , even though the BCs are inhomogeneous. In practice, there is no need to calculate  $M$ . We can use the expression for  $Q[u, g^*]$  obtained from the Lagrange identity of Chapter 13 and evaluate it at  $b$  and  $a$ . This, in general, involves evaluating  $u$  and  $G$  and their derivatives at  $a$  and  $b$ . We know how to handle the evaluation of  $G$  because we can actually construct it (if it exists). We next find two of the four quantities corresponding to  $u$  in terms of the other two and insert the result in the expression for  $Q[u, g^*]$ . Equation (20.39) then guarantees that the coefficients of the other two terms will be zero. Thus, we can simply drop all the terms in  $Q[u, g^*]$  containing a factor of the other two terms.

Specifically, we use the conjunct for a formally self-adjoint SOLDO [see Equation (20.26)] and  $g^*(x, y) = G(y, x)$  to obtain

$$u(y) = \int_a^b w(x)G(y, x)f(x) dx - \left\{ p(x)w(x) \left[ G(y, x) \frac{du}{dx} - u(x) \frac{\partial G}{\partial x}(y, x) \right] \right\}_{x=a}^{x=b}$$

Interchanging  $x$  and  $y$  gives

$$u(x) = \int_a^b w(y)G(x, y)f(y) dy + \left\{ p(y)w(y) \left[ u(y) \frac{\partial G}{\partial y}(x, y) - G(x, y) \frac{du}{dy} \right] \right\}_{y=a}^{y=b} \tag{20.41}$$

This equation is valid only for a self-adjoint SOLDO. That is, using it requires casting the SOLDO into a self-adjoint form (a process that is always possible, in light of Theorem 13.5.4).

By setting  $f(x) = 0$ , we can also obtain the solution to a homogeneous DE  $L_x[u] = 0$  that satisfies the inhomogeneous BCs.

**20.3.9. Example.** Let us find the solution of the simple DE  $d^2u/dx^2 = f(x)$  subject to the simple inhomogeneous BCs  $u(a) = \gamma_1$  and  $u(b) = \gamma_2$ . The GF for this problem has been calculated in Examples 20.1.5 and 20.3.5. Let us begin by calculating the surface term in Equation (20.41). We have  $p(y) = 1$ , and we set  $w(y) = 1$ , then

$$\begin{aligned} \text{surface term} &= u(b) \left. \frac{\partial G}{\partial y} \right|_{y=b} - G(x, b)u'(b) - u(a) \left. \frac{\partial G}{\partial y} \right|_{y=a} + G(x, a)u'(a) \\ &= \gamma_2 \left. \frac{\partial G}{\partial y} \right|_{y=b} - \gamma_1 \left. \frac{\partial G}{\partial y} \right|_{y=a} + G(x, a)u'(a) - G(x, b)u'(b). \end{aligned}$$

That the unwanted (and unspecified) terms are zero can be seen by observing that  $G(x, a) = g^*(a, x) = (g(a, x))^*$ , and that  $g(x, y)$  satisfies the BCs adjoint to the homogeneous BCs (obtained when  $\gamma_i = 0$ ). In this particular and simple case, the BCs happen to be self-adjoint (Dirichlet BCs). Thus,  $u(a) = u(b) = 0$  implies that  $g(a, x) = g(b, x) = 0$  for all  $x \in [a, b]$ . (In a more general case the coefficient of  $u'(a)$  would be more complicated, but still zero.) Thus, we finally have

$$\text{surface term} = \gamma_2 \left. \frac{\partial G}{\partial y} \right|_{y=b} - \gamma_1 \left. \frac{\partial G}{\partial y} \right|_{y=a}$$

Now, using the expression for  $G(x, y)$  obtained in Examples 20.1.5 and 20.3.5, we get

$$\frac{\partial G}{\partial y} = -\frac{a-x}{b-a} - \theta(x-y) - \underbrace{(x-y)\delta(x-y)}_{=0} = \frac{x-a}{b-a} - \theta(x-y).$$

Thus,

$$\left. \frac{\partial G}{\partial y} \right|_{y=b} = \frac{x-a}{b-a}, \quad \left. \frac{\partial G}{\partial y} \right|_{y=a} = \frac{x-a}{b-a} - 1 = \frac{x-b}{b-a}.$$

Substituting in Equation (20.41), we get

$$u(x) = \int_a^b G(x, y) f(y) dy + \frac{\gamma_2 - \gamma_1}{b - a} x + \frac{b\gamma_1 - a\gamma_2}{b - a}.$$

(Compare this with the result obtained in Example 20.1.5.) ■

Green's functions have a very simple and enlightening physical interpretation. An inhomogeneous DE such as  $\mathbf{L}_x[u] = f(x)$  can be interpreted as a black box ( $\mathbf{L}_x$ ) that determines a physical quantity ( $u$ ) when there is a source ( $f$ ) of that physical quantity. For instance, electrostatic potential is a physical quantity whose source is charge; a magnetic field has an electric current as its source; displacements and velocities have forces as their sources; and so forth. Applying this interpretation and assuming that  $w(x) = 1$ , we have  $G(x, y)$  as the physical quantity, evaluated at  $x$  when its source  $\delta(x - y)$  is located at  $y$ . To be more precise, let us say that the strength of the source is  $S_1$  and it is located at  $y_1$ ; then the source becomes  $S_1\delta(x - y_1)$ . The physical quantity, the Green's function, is then  $S_1G(x, y_1)$ , because of the linearity of  $\mathbf{L}_x$ : If  $G(x, y)$  is a solution of  $\mathbf{L}_x[u] = \delta(x - y)$ , then  $S_1G(x, y_1)$  is a solution of  $\mathbf{L}_x[u] = S_1\delta(x - y)$ . If there are many sources located at  $\{y_i\}_{i=1}^N$  with corresponding strengths  $\{S_i\}_{i=1}^N$ , then the overall source  $f$  as a function of  $x$  becomes  $f(x) = \sum_{i=1}^N S_i\delta(x - y_i)$ , and the corresponding physical quantity  $u(x)$  becomes  $u(x) = \sum_{i=1}^N S_iG(x, y_i)$ .

Since the source  $S_i$  is located at  $y_i$ , it is more natural to define a function  $S(x)$  and write  $S_i = S(y_i)$ . When the number of point sources goes to infinity and  $y_i$  becomes a smooth continuous variable, the sums become integrals, and we have

$$f(x) = \int_a^b S(y)\delta(x - y) dy, \quad u(x) = \int_a^b S(y)G(x, y) dy.$$

The first integral shows that  $S(x) = f(x)$ . Thus, the second integral becomes  $u(x) = \int_a^b f(y)G(x, y) dy$  which is precisely what we obtained formally.

## 20.4 Eigenfunction Expansion of Green's Functions

Green's functions are inverses of differential operators. Inverses of operators in a Hilbert space are best studied in terms of resolvents. This is because if an operator  $\mathbf{A}$  has an inverse, zero is in its resolvent set, and

$$\mathbf{R}_0(\mathbf{A}) = \mathbf{R}_\lambda(\mathbf{A})|_{\lambda=0} = (\mathbf{A} - \lambda\mathbf{1})^{-1}|_{\lambda=0} = \mathbf{A}^{-1}.$$

Thus, it is instructive to discuss Green's functions in the context of the resolvent of a differential operator. We will consider only the case where the eigenvalues are discrete, for example, when  $\mathbf{L}_x$  is a Sturm–Liouville operator.

Formally, we have  $(\mathbf{L} - \lambda\mathbf{1})\mathbf{R}_\lambda(\mathbf{L}) = \mathbf{1}$ , which leads to the DE

$$(\mathbf{L}_x - \lambda)R_\lambda(x, y) = \frac{\delta(x - y)}{w(x)},$$

where  $R_\lambda(x, y) = \langle x | \mathbf{R}_\lambda(\mathbf{L}) | y \rangle$ . The DE simply says that  $R_\lambda(x, y)$  is the Green's function for the operator  $\mathbf{L}_x - \lambda$ . So we can rewrite the equation as  $(\mathbf{L}_x - \lambda)G_\lambda(x, y) = \delta(x - y)/w(x)$  where  $\mathbf{L}_x - \lambda$  is a DO having some homogeneous BCs. The GF  $G_\lambda(x, y)$  exists if and only if  $(\mathbf{L}_x - \lambda)[u] = 0$  has no nontrivial solution, which is true only if  $\lambda$  is not an eigenvalue of  $\mathbf{L}_x$ . We choose the BCs in such a way that  $\mathbf{L}_x$  becomes self-adjoint.

Let  $\{\lambda_n\}_{n=1}^\infty$  be the eigenvalues of the system  $\mathbf{L}_x[u] = \lambda u$ ,  $\{\mathbf{R}_i[u] = 0\}_{i=1}^2$ , and let the  $u_n^{(k)}(x)$  be the corresponding eigenfunctions. The index  $k$  distinguishes among the linearly independent vectors corresponding to the same eigenvalue  $\lambda_n$ . Assuming that  $\mathbf{L}$  has compact resolvent (e.g., a Sturm–Liouville operator), these eigenfunctions form a complete set for the subspace of the Hilbert space that consists of those functions that satisfy the same BCs as the  $u_n^{(k)}(x)$ . In particular,  $G_\lambda(x, y)$  can be expanded in terms of  $u_n^{(k)}(x)$ . The expansion coefficients are, of course, functions of  $y$ . Thus, we can write

$$G_\lambda(x, y) = \sum_k \sum_{n=1}^\infty a_n^{(k)}(y) u_n^{(k)}(x)$$

where  $a_n^{(k)}(y) = \int_a^b w(x) u_n^{*(k)}(x) G_\lambda(x, y) dx$ . Using Green's identity, Equation (20.30), and the fact that  $\lambda_n$  is real, we have

$$\begin{aligned} \lambda_n a_n^{(k)}(y) &= \int_a^b w(x) [\lambda_n u_n^{(k)}(x)]^* G_\lambda(x, y) dx \\ &= \int_a^b w(x) G_\lambda(x, y) \{\mathbf{L}_x [u_n^{(k)}(x)]\}^* dx \\ &= \int_a^b w(x) [u_n^{(k)}(x)]^* \mathbf{L}_x [G_\lambda(x, y)] dx \\ &= \int_a^b w(x) u_n^{*(k)}(x) \left[ \frac{\delta(x - y)}{w(x)} + \lambda G_\lambda(x, y) \right] dx \\ &= u_n^{*(k)}(y) + \lambda \int_a^b w(x) u_n^{*(k)}(x) G_\lambda(x, y) dx \\ &= u_n^{*(k)}(y) + \lambda a_n^{(k)}(y). \end{aligned}$$

Thus,  $a_n^{(k)}(y) = u_n^{*(k)}(y)/(\lambda_n - \lambda)$ , and the expansion for the Green's function is

$$G_\lambda(x, y) = \sum_k \sum_{n=1}^\infty \frac{u_n^{*(k)}(y) u_n^{(k)}(x)}{\lambda_n - \lambda}. \tag{20.42}$$

This expansion is valid as long as  $\lambda_n \neq \lambda$  for any  $n = 0, 1, 2, \dots$ . But this is precisely the condition that ensures the existence of an inverse for  $\mathbf{L} - \lambda \mathbf{1}$ .

An interesting result is obtained from Equation (20.42) if  $\lambda$  is considered a complex variable. In that case,  $G_\lambda(x, y)$  has (infinitely many) simple poles at

$\{\lambda_n\}_{n=1}^\infty$ . The residue at the pole  $\lambda_n$  is  $-\sum_k u_n^{*(k)}(y)u_n^{(k)}(x)$ . If  $C_m$  is a contour having the poles  $\{\lambda_n\}_{n=1}^m$  in its interior, then, by the residue theorem, we have

$$\frac{1}{2\pi i} \oint_{C_m} G_\lambda(x, y) d\lambda = -\sum_k \sum_{n=1}^m u_n^{*(k)}(y)u_n^{(k)}(x).$$

In particular, if we let  $m \rightarrow \infty$ , we obtain

$$\frac{1}{2\pi i} \oint_{C_\infty} G_\lambda(x, y) d\lambda = -\sum_k \sum_{n=1}^\infty u_n^{*(k)}(y)u_n^{(k)}(x) = -\frac{\delta(x-y)}{w(x)}, \tag{20.43}$$

where  $C_\infty$  is any contour that encircles all the eigenvalues, and in the last step we used the completeness of the eigenfunctions. Equation (20.43) is the infinite-dimensional analogue of Equation (16.12) with  $f(\mathbf{A}) = 1$  when the latter equation is sandwiched between  $\langle x|$  and  $|y\rangle$ .

**20.4.1. Example.** Consider the DO  $L_x = d^2/dx^2$  with BCs  $u(0) = u(a) = 0$ . This is an S-L operator with eigenvalues and normalized eigenfunctions

$$\lambda_n = \left(\frac{n\pi}{a}\right)^2 \quad \text{and} \quad u_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) \quad \text{for } n = 1, 2, \dots$$

Equation (20.42) becomes

$$G_\lambda(x, y) = -\frac{2}{a} \sum_{n=1}^\infty \frac{\sin(n\pi x/a) \sin(n\pi y/a)}{\lambda - (n\pi/a)^2},$$

which leads to

$$\begin{aligned} \frac{1}{2\pi i} \oint_{C_\infty} G_\lambda(x, y) d\lambda &= \frac{1}{2\pi i} \oint_{C_\infty} \left[ -\frac{2}{a} \sum_{n=1}^\infty \frac{\sin(n\pi x/a) \sin(n\pi y/a)}{\lambda - (n\pi/a)^2} \right] d\lambda \\ &= -\frac{1}{2\pi i} \left(\frac{2}{a}\right) \sum_{n=1}^\infty \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{n\pi}{a}y\right) \oint_{C_\infty} \frac{d\lambda}{\lambda - (n\pi/a)^2} \\ &= -\left(\frac{2}{a}\right) \sum_{n=1}^\infty \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{n\pi}{a}y\right) \text{Res} \left[ \frac{1}{\lambda - (n\pi/a)^2} \right]_{\lambda=\lambda_n} \\ &= -\left(\frac{2}{a}\right) \sum_{n=1}^\infty \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{n\pi}{a}y\right). \end{aligned}$$

The RHS is recognized as  $-\delta(x-y)$ . ■

eigenfunction  
expansion of GF

If zero is not an eigenvalue of  $L_x$ , Equation (20.42) yields

$$G(x, y) \equiv G_0(x, y) = \sum_k \sum_{n=1}^\infty \frac{u_n^{*(k)}(y)u_n^{(k)}(x)}{\lambda_n}, \tag{20.44}$$

which is an expression for the Green's function of  $L_x$  in terms of its eigenvalues and eigenfunctions.

## 20.5 Problems

**20.1.** Using the GF method, solve the DE  $\mathbf{L}_x u(x) = du/dx = f(x)$  subject to the BC  $u(0) = a$ . Hint: Consider the function  $v(x) = u(x) - a$ .

**20.2.** Solve the problem of Example 20.1.4 subject to the BCs  $u(a) = u'(a) = 0$ . Show that the corresponding GF also satisfies these BCs.

**20.3.** Show that the IVP with data  $\{0; 0, 0, \dots, 0\}$  has only  $u \equiv 0$  as a solution. Hint: Assume otherwise, add  $u$  to the solution of the inhomogeneous equation, and invoke uniqueness.

**20.4.** In this problem, we generalize the concepts of exactness and integrating factor to a NOLDE. The DO  $\mathbf{L}_x^{(n)} \equiv \sum_{k=0}^n p_k(x) d^k/dx^k$  is said to be **exact** if there exists a DO  $\mathbf{M}_x^{(n-1)} \equiv \sum_{k=0}^{n-1} a_k(x) d^k/dx^k$  such that

$$\mathbf{L}_x^{(n)}[u] = \frac{d}{dx}(\mathbf{M}_x^{(n-1)}[u]) \quad \forall u \in \mathcal{C}^n[a, b].$$

(a) Show that  $\mathbf{L}_x^{(n)}$  is exact iff  $\sum_{m=0}^n (-1)^m d^m p_m/dx^m = 0$ .

(b) Show that there exists an integrating factor for  $\mathbf{L}_x^{(n)}$ —that is, a function  $\mu(x)$  such that  $\mu(x)\mathbf{L}_x^{(n)}$  is exact—if and only if  $\mu(x)$  satisfies the DE

$$\mathbf{N}_x^{(n)}[\mu] \equiv \sum_{m=0}^n (-1)^m \frac{d^m}{dx^m} (\mu p_m) = 0.$$

The DO  $\mathbf{N}_x^{(n)}$  is the formal adjoint of  $\mathbf{L}_x^{(n)}$ .

**20.5.** Assuming that  $\mathbf{L}_x[u] = 0$  has no nontrivial solution, show that the matrix

$$\mathbf{R} \equiv \begin{pmatrix} \mathbf{R}_1[u_1] & \mathbf{R}_1[u_2] \\ \mathbf{R}_2[u_1] & \mathbf{R}_2[u_2] \end{pmatrix},$$

where  $u_1$  and  $u_2$  are independent solutions of  $\mathbf{L}_x[u] = 0$  and  $\mathbf{R}_i$  are the boundary functionals, has a nonzero determinant. Hint: Assume otherwise and show that the system of homogeneous linear equations  $\alpha \mathbf{R}_1[u_1] + \beta \mathbf{R}_1[u_2] = 0$  and  $\alpha \mathbf{R}_2[u_1] + \beta \mathbf{R}_2[u_2] = 0$  has a nontrivial solution for  $(\alpha, \beta)$ . Reach a contradiction by considering  $u = \alpha u_1 + \beta u_2$  as a solution of  $\mathbf{L}_x[u] = 0$ .

**20.6.** Determine the formal adjoint of each of the operators in (a) through (d) below (i) as a differential operator, and (ii) as an operator, that is, including the BCs. Which operators are formally self-adjoint? Which operators are self-adjoint?

(a)  $\mathbf{L}_x = d^2/dx^2 + 1$  in  $[0, 1]$  with BCs  $u(0) = u(1) = 0$ .

(b)  $\mathbf{L}_x = d^2/dx^2$  in  $[0, 1]$  with BCs  $u(0) = u'(0) = 0$ .

(c)  $\mathbf{L}_x = d/dx$  in  $[0, \infty]$  with BCs  $u(0) = 0$ .

(d)  $\mathbf{L}_x = d^3/dx^3 - \sin x d/dx + 3$  in  $[0, \pi]$  with BCs  $u(0) = u'(0) = 0, u''(0) - 4u(\pi) = 0$ .

**20.7.** Show that the Dirichlet, Neumann, general unmixed, and periodic BCs make the following formally self-adjoint SOLDO self-adjoint:

$$\mathbf{L}_x = \frac{1}{w} \frac{d}{dx} \left( p \frac{d}{dx} \right) + q.$$

**20.8.** Using a procedure similar to that described in the text for SOLDOs, show that for the FOLDO  $\mathbf{L}_x = p_1 d/dx + p_0$

(a) the indefinite GF is

$$G(x, y) \equiv \frac{\mu(y)}{p_1(y)w(y)} \left[ \frac{\theta(x-y)}{\mu(x)} \right] + C(y),$$

where  $\mu(x) = \exp \left[ \int^x \frac{p_0(t)}{p_1(t)} dt \right],$

(b) and the GF itself is discontinuous at  $x = y$  with

$$\lim_{\epsilon \rightarrow 0} [G(y + \epsilon, y) - G(y - \epsilon, y)] = \frac{1}{p_1(y)w(y)}.$$

(c) For the homogeneous BC

$$\mathbf{R}[u] \equiv \alpha_1 u(a) + \alpha_2 u'(a) + \beta_1 u(b) + \beta_2 u'(b) = 0$$

construct  $G(x, y)$  and show that

$$G(x, y) = \frac{1}{p_1(y)w(y)v(y)} v(x)\theta(x-y) + C(y)v(x),$$

where  $v(x)$  is any solution to the homogeneous DE  $\mathbf{L}_x[u] = 0$  and

$$C(y) = \frac{\beta_1 v(b) + \beta_2 v'(b)}{\mathbf{R}[v]p_1(y)w(y)v(y)}, \quad \text{with } \mathbf{R}[v] \neq 0$$

(d) Show directly that  $\mathbf{L}_x[G] = \delta(x-y)/w(x)$ .

**20.9.** Let  $\mathbf{L}_x$  be a NOLDO with constant coefficients. Show that if  $u(x)$  satisfies  $\mathbf{L}_x[u] = f(x)$ , then  $u(x-y)$  satisfies  $\mathbf{L}_x[u] = f(x-y)$ . (Note that no BCs are specified.)

**20.10.** Find the GF for  $\mathbf{L}_x = d^2/dx^2 + 1$  with BCs  $u(0) = u'(0) = 0$ . Show that it can be written as a function of  $x-y$  only.

**20.11.** Find the GF for  $\mathbf{L}_x = d^2/dx^2 + k^2$  with BCs  $u(0) = u(a) = 0$ .

**20.12.** Find the GF for  $\mathbf{L}_x = d^2/dx^2 - k^2$  with BCs  $u(\infty) = u(-\infty) = 0$ .

**20.13.** Find the GF for  $\mathbf{L}_x = (d/dx)(xd/dx)$  given the condition that  $G(x, y)$  is finite at  $x = 0$  and vanishes at  $x = 1$ .

**20.14.** Evaluate the GF and the solutions for each of the following DEs in the interval  $[0, 1]$ .

(a)  $u'' - k^2u = f; u(0) - u'(0) = a, u(1) = b.$

(b)  $u'' = f; u(0) = u'(0) = 0.$

(c)  $u'' + 6u' + 9u = 0; u(0) = 0, u'(0) = 1.$

(d)  $u'' + \omega^2u = f(x), \text{ for } x > 0; u(0) = a, u'(0) = b.$

(e)  $u^{(4)} = f; u(0) = 0, u'(0) = 2u'(1), u(1) = a, u''(0) = 0.$

**20.15.** Use eigenfunction expansion of the GF to solve the BVP  $u'' = x, u(0) = 0, u(1) - 2u'(1) = 0.$

## Additional Reading

1. Dennerly, P. and Krzywicki, A. *Mathematics for Physicists*, Harper and Row, 1967. Contains an exposition of Green's functions for second-order linear differential equations.
2. Roach, G. *Green's Functions*, Van Nostrand, 1970. A readable introduction to Green's functions, especially those we have called "indefinite" (no boundary conditions specified).



# Multidimensional Green's Functions: Formalism

The extensive study of Green's functions in one dimension in the last chapter has no doubt exhibited the power and elegance of their use in solving inhomogeneous differential equations. If the differential equation has a (unique) solution, the GF exists and contains all the information necessary to build it up. The solution results from operating on the inhomogeneous term with an integral operator whose kernel is the appropriate Green's function.

The Green's function's very existence depends on the type of BCs imposed. We encountered two types of problems in solving ODEs. The first, called initial value problems (IVPs), involves fixing (for an  $n$ th-order DE) the value of the solution and its first  $n - 1$  derivatives at a fixed point. Then the ODE, if it is sufficiently well-behaved, will determine the values of the solution in the neighborhood of the fixed point in a unique way. Because of this uniqueness, Green's functions always exist for IVPs.

The second type of problems, called boundary value problems (BVPs), consists—when the DE is second order—of determining a relation between the solution and its derivative evaluated at the boundaries of some interval  $[a, b]$ . These boundary values are relations that we denoted by  $\mathbf{R}_i[u] = \gamma_i$ , where  $i = 1, 2$ . In this case, the existence and uniqueness of the Green's function are not guaranteed.

There is a fundamental (topological) difference between a boundary in one dimension and a boundary in two and more dimensions. In one dimension a boundary consists of *only two points*; in 2 and higher dimensions a boundary has *infinitely many points*. The boundary of a region in  $\mathbb{R}^2$  is a closed curve, in  $\mathbb{R}^3$  it is a closed surface, and in  $\mathbb{R}^m$  it is called a **hypersurface**. This fundamental difference makes the study of Green's functions in higher dimensions more complicated, but also richer and more interesting.

## 21.1 Properties of Partial Differential Equations

This section presents certain facts and properties of PDEs, in particular, how BCs affect their solutions. We shall discover the important difference between ODEs and PDEs:

**21.1.1. Box.** *The existence of a solution to a PDE satisfying a given BC depends on the type of the PDE.*

We shall be concerned exclusively with a linear PDE. A linear PDE of order  $M$  in  $m$  variables is of the form

$$\mathbf{L}_x[u] = f(\mathbf{x}) \quad \text{where} \quad \mathbf{L}_x = \sum_{|J|=1}^M \sum_J a_J(\mathbf{x}) \frac{\partial^{|J|}}{\partial \mathbf{x}^J}, \quad (21.1)$$

where the following notation has been used:

$$\begin{aligned} \mathbf{x} &= (x_1, \dots, x_m), & J &= (j_1, \dots, j_m), \\ |J| &= j_1 + j_2 + \dots + j_m, & \frac{\partial^{|J|}}{\partial \mathbf{x}^J} &= \frac{\partial^{|J|}}{\partial x_1^{j_1} \partial x_2^{j_2} \dots \partial x_m^{j_m}}; \end{aligned}$$

the  $j_k$  are nonnegative integers;  $M$ , the order of the highest derivative, is called the **order** of the PDE. The outer sum in Equation (21.1) is over  $|J|$ ; once  $|J|$  is fixed, the inner summation goes over individual  $j_k$ 's with the restriction that their sum has to equal the given  $|J|$ .

principal part of a  
PDE

The **principal part** of  $\mathbf{L}_x$  is

$$\mathbf{L}_p = \sum_{|J|=M} a_J(x_1, \dots, x_m) \frac{\partial^M}{\partial \mathbf{x}^J}. \quad (21.2)$$

The coefficients  $a_J$  and the inhomogeneous (or source) term  $f$  are assumed to be continuous functions of their arguments.

Cauchy data and  
Cauchy problem

We consider Equation (21.1) as an IVP with appropriate initial data. The most direct generalization of the IVP of ordinary differential equation theory is to specify the values of  $u$  and all its *normal derivatives* of order less than or equal to  $M - 1$  on a hypersurface  $\Gamma$  of dimension  $m - 1$ . This type of initial data is called **Cauchy data**, and the resulting IVP is known as the **Cauchy problem** for  $\mathbf{L}_x$ . The reason that the tangential derivatives do not come into play here is that once we know the values of  $u$  on  $\Gamma$ , we can evaluate  $u$  on two neighboring points on  $\Gamma$ , take the limit as the points get closer and closer, and evaluate the tangential derivatives.

### 21.1.1 Characteristic Hypersurfaces

In contrast to the IVP in one dimension, the Cauchy problem for arbitrary Cauchy data may not have a solution, or if it does, the solution may not be unique.

**21.1.2. Box.** *The existence and uniqueness of the solution of the Cauchy problem depend crucially on the hypersurface  $\Gamma$  and on the type of PDE.*

We assume that  $\Gamma$  can be parametrized by a set of  $m$  functions of  $m - 1$  parameters. These parameters can be thought of as generalized coordinates of points of  $\Gamma$ .

tangential  
coordinates

Consider a point  $P$  on  $\Gamma$ . Introduce  $m - 1$  coordinates  $\xi_2, \dots, \xi_m$ , called **tangential coordinates**, to label points on  $\Gamma$ . Choose, by translation if necessary, coordinates in such a way that  $P$  is the origin, with coordinates  $(0, 0, \dots, 0)$ . Now let  $\nu = \xi_1$  stand for the remaining coordinate normal to  $\Gamma$ . Usually  $\xi_i$  is taken to be the  $i$ th coordinate of the projection of the point on  $\Gamma$  onto the hyperplane tangent to  $\Gamma$  at  $P$ .

As long as we do not move too far away from  $P$ , the Cauchy data on  $\Gamma$  can be written as

$$u(0, \xi_2, \dots, \xi_m), \quad \frac{\partial u}{\partial \nu}(0, \xi_2, \dots, \xi_m), \dots, \quad \frac{\partial^{M-1} u}{\partial \nu^{M-1}}(0, \xi_2, \dots, \xi_m).$$

Using the chain rule,  $\partial u / \partial x_i = \sum_{j=1}^m (\partial u / \partial \xi_j)(\partial \xi_j / \partial x_i)$ , where  $\xi_1 = \nu$ , we can also determine the first  $M - 1$  derivatives of  $u$  with respect to  $x_i$ . The fundamental question is whether we can determine  $u$  uniquely using the above Cauchy data and the DE. To motivate the answer, let's look at the analogous problem in one dimension.

Consider the  $M$ th-order linear ODE

$$\mathbb{L}_x[u] = a_M(x) \frac{d^M u}{dx^M} + \dots + a_1(x) \frac{du}{dx} + a_0(x)u = f(x) \quad (21.3)$$

with the following initial data at  $x_0$ :  $\{u(x_0), u'(x_0), \dots, u^{(M-1)}(x_0)\}$ . If the coefficients  $\{a_k(x)\}_{k=0}^M$  and the inhomogeneous term  $f(x)$  are continuous and if  $a_M(x_0) \neq 0$ , then Theorem 20.2.3 implies that there exists a unique solution to the IVP in a neighborhood of  $x_0$ .

For  $a_M(x_0) \neq 0$ , Equation (21.3), the initial data, and a knowledge of  $f(x_0)$  give  $u^{(M)}(x_0)$  uniquely. Having found  $u^{(M)}(x_0)$ , we can calculate, with arbitrary accuracy (by choosing  $\Delta x$  small enough), the following set of *new* initial data at  $x_1 = x_0 + \Delta x$ :

$$u(x_1) = u(x_0) + u'(x_0)\Delta x, \dots, u^{(M-1)}(x_1) = u^{(M-1)}(x_0) + u^{(M)}(x_0)\Delta x.$$

Using these new initial data and Theorem 20.2.3, we are assured of a unique solution at  $x_1$ . Since  $a_M(x)$  is assumed to be continuous for  $x_1$ , for sufficiently small

$\Delta x$ ,  $a_M(x_0)$  is nonzero, and it is possible to find newer initial data at  $x_2 = x_1 + \Delta x$ . The process can continue until we reach a singularity of the DE, a point where  $a_M(x)$  vanishes. We can thus construct the unique solution of the IVP in an interval  $(x_0, b)$  as long as  $a_M(x)$  does not vanish anywhere in  $[x_0, b]$ . This procedure is analogous to the one used in the analytic continuation of a complex function.

For  $a_M(x_0) = 0$ , however, we cannot calculate  $u^{(M)}(x_0)$  unambiguously. In such a case the LHS of (21.3) is completely determined from the initial data. If the LHS happens to be equal to  $f(x_0)$ , then the equation is satisfied for *any*  $u^{(M)}(x_0)$ , i.e., there exist infinitely many solutions for  $u^{(M)}(x_0)$ ; if the LHS is not equal to  $f(x_0)$ , there are no solutions. The difficulty can be stated in another way, which is useful for generalization to the  $m$ -dimensional case:

**21.1.3. Box.** *If  $a_M(x_0) = 0$  in (21.3), then the initial data determine the function  $\mathbf{L}_x[u]$ .*

Let us now return to the question of constructing  $u$  and investigate conditions under which the Cauchy problem may have a solution. We follow the same steps as for the IVP for ODEs. To construct the solution numerically for points near  $P$  but away from  $\Gamma$  (since the function is completely determined on  $\Gamma$ , not only its  $M$ th derivative but derivatives of all orders are known on  $\Gamma$ ), we must be able to calculate  $\partial^M u / \partial v^M$  at  $P$ . This is not possible if the coefficient of  $\partial^M u / \partial v^M$  in  $\mathbf{L}_x[u]$  is zero when  $x_1, \dots, x_m$  is written in terms of  $v, \xi_2, \dots, \xi_m$ . When this happens,  $\mathbf{L}_x[u]$  itself will be determined by the Cauchy data. This motivates the following definition.

characteristic  
hypersurface

**21.1.4. Definition.** *If  $\mathbf{L}_x[u]$  can be evaluated at a point  $P$  on  $\Gamma$  from the Cauchy data alone, then  $\Gamma$  is said to be **characteristic** for  $\mathbf{L}_x$  at  $P$ . If  $\Gamma$  is characteristic for all its points, then it is called a **characteristic hypersurface** for  $\mathbf{L}_x$ . The Cauchy problem does not have a solution at a point on the characteristic hypersurface.*

The following theorem characterizes  $\Gamma$ :

**21.1.5. Theorem.** *Let  $\Gamma$  be a smooth  $(m - 1)$ -dimensional hypersurface. Let  $\mathbf{L}_x[u] = f$  be an  $M$ th-order linear PDE in  $m$  variables. Then  $\Gamma$  is characteristic at  $P \in \Gamma$  if and only if the coefficient of  $\partial^M u / \partial v^M$  vanishes when  $\mathbf{L}_x$  is expressed in terms of the normal-tangential coordinate system  $(v, \xi_2, \dots, \xi_m)$ .*

One can rephrase the foregoing theorem as follows:

**21.1.6. Box.** *The hypersurface  $\Gamma$  is not characteristic at  $P$  if and only if all  $M$ th-order partial derivatives of  $u$  with respect to  $\{x_i\}_{i=1}^m$  are unambiguously determined at  $P$  by the DE and the Cauchy data on  $\Gamma$ .*

Characteristic  
"hypersurfaces" of  
ODEs are points!

In the one-dimensional case the difficulty arose when  $a_M(x_0) = 0$ . In the language being used here, we could call  $x_0$  a "characteristic point." This makes sense because in this special case ( $m = 1$ ), the hypersurfaces can only be of dimension 0. Thus, we can say that in the neighborhood of a characteristic point, the IVP has no well-defined solution.<sup>1</sup> For the general case ( $m > 1$ ), we can similarly say that the Cauchy problem has no well-defined solution in the neighborhood of  $P$  if  $P$  happens to lie on a characteristic hypersurface of the differential operator. Thus, it is important to determine the characteristic hypersurfaces of PDEs.

**21.1.7. Example.** Let us consider the first-order PDE in two variables

$$\mathbf{L}_x[u] = a(x, y) \frac{\partial u}{\partial x} + b(x, y) \frac{\partial u}{\partial y} + F(x, y, u) = 0 \quad (21.4)$$

where  $F(x, y, u) \equiv c(x, y)u + d(x, y)$ . For this discussion the form of  $F$  is irrelevant.

We wish to find the characteristic hypersurfaces (in this case, curves) of  $\mathbf{L}$ . The Cauchy data consist of a simple determination of  $u$  on  $\Gamma$ . By Theorem 21.1.5, we need to derive relations that ensure that  $\partial u/\partial x$  and  $\partial u/\partial y$  cannot be unambiguously determined at  $P = (x, y)$ . Using an obvious notation, the PDE of Equation (21.4) gives

$$-F(P, u(P)) = a(P) \frac{\partial u}{\partial x}(P) + b(P) \frac{\partial u}{\partial y}(P).$$

On the other hand, if  $Q \equiv (x + dx, y + dy)$  lies on the curve  $\Gamma$ , then

$$u(Q) - u(P) = dx \frac{\partial u}{\partial x}(P) + dy \frac{\partial u}{\partial y}(P).$$

The Cauchy data determine the LHS of both of the preceding equations. Treating these equations as a system of two linear equations in two unknowns,  $\partial u/\partial x(P)$  and  $\partial u/\partial y(P)$ , we conclude that the system has a unique solution if and only if the matrix of coefficients is invertible. Thus, by Box 21.1.6,  $\Gamma$  is a characteristic curve if and only if

$$\det \begin{pmatrix} dx & dy \\ a(P) & b(P) \end{pmatrix} = b(P) dx - a(P) dy = 0,$$

or  $dy/dx = b(x, y)/a(x, y)$ , assuming that  $a(x, y) \neq 0$ . Solving this FODE yields  $y$  as a function of  $x$ , thus determining the characteristic curve. Note that a general solution of this FODE involves an arbitrary constant, resulting in a family of characteristic curves. ■

---

**Sofia Vasilyevna Kovalevskaya** (1850–1891) is considered the greatest woman mathematician prior to the twentieth century. She grew up in a well-educated family of the Russian nobility, her father being an artillery general and reputed to be a descendant of a Hungarian king, Mathias Korvin. Sonja was educated by a British governess and enjoyed life at the large country estate of her father's family, although the rather progressive thinking of the Kovalevsky sisters did not always meet with approval from their father.

<sup>1</sup>Here lies the crucial difference between ODEs and PDEs: All ODEs have a universal characteristic hypersurface, i.e., a point. PDEs, on the other hand, can have a variety of hypersurfaces.

Sonja has written of two factors that attracted her to the study of mathematics. The first was her Uncle Pyotr, who had studied the subject on his own and would speak of squaring the circle and of the asymptote, as well as of many other things that excited her imagination. The second was a curious "wallpaper" that was used to cover one of the children's rooms at Polibino, which turned out to be lecture notes on differential and integral calculus that had been purchased by her father in student days. These sheets fascinated her and she would spend hours trying to decipher separate phrases and to find the proper ordering of the pages.

In the autumn of 1867 Sonja went to St. Petersburg, where she studied calculus with Alexander Strannolyubsky, a teacher of mathematics at the naval school. While there, she consulted the prominent Russian mathematician Chebyshev about her mathematical studies, but since Russian universities were closed to women, there seemed to be no way that she could pursue advanced studies in her native land.

In order to escape the oppression of women common in Russia at the time, young ladies of ambition and ability would often arrange a marriage of convenience in order to allow study at a foreign university. At the age of 18, Sonya arranged such a marriage with Vladimir Kovalevsky, a paleontologist, and in 1869 the couple moved to Heidelberg, where Sonja took courses from Kirchhoff, Helmholtz, and others. Two years later she went to Berlin, where she worked with Weierstrass, who tutored her privately, since she, as a woman, was not allowed to attend lectures.



The three papers she published in the next three years earned her a doctorate in absentia from the University of Göttingen. Unfortunately, even that distinction was not sufficient to gain her a university position anywhere in Europe, despite strong recommendation from the renowned Weierstrass. Her rejections resulted in a six-year period during which time she neither undertook research nor replied to Weierstrass's letters. She was bitter to discover that the best job she was offered was teaching arithmetic to elementary classes of schoolgirls, and remarked, "I was unfortunately weak in the multiplication table."

The existence and uniqueness of solutions to partial differential equations occupied the attention of many notable mathematicians of the last century, including Cauchy, who transformed the problem into his *method of majorant functions*. This method was later extended and refined by Kovalevskaya to include more general cases. The result was the now-famous *Cauchy-Kovalevskaya theorem*. She also contributed to the advancement of the study of Abelian integrals and functions and applied her knowledge of these topics to problems in physics, including her paper "On the Rotation of a Solid Body About a Fixed Point," for which she won a 5000-franc prize. She also performed some investigations into the dynamics of Saturn's rings, inspiring a sonnet in which she is named "Muse of the Heavens." In 1878, Kovalevskaya gave birth to a daughter, but from 1880 increasingly returned to her study of mathematics. In 1882 she began work on the refraction of light, and wrote three articles on the topic. In the spring of 1883, Vladimir, from whom Sonja had been separated for two years, committed suicide. After the initial shock, Kovalevskaya immersed herself in mathematical work in an attempt to rid herself of feelings of guilt. Mittag-Leffler managed to overcome opposition to Kovalevskaya in Stockholm, and obtained for her a position as privat docent. She began to lecture there in early 1884, was appointed to a five-year extraordinary professorship in June of that year, and in June 1889 became the third woman ever to hold a chair at a European university.

During Kovalevskaya's years at Stockholm she carried out important research, taught courses on the latest topics in analysis, and became an editor of the new journal *Acta Mathematica*. She was the liaison with the mathematicians of Paris and Berlin, and took part in the organization of international conferences. Interestingly, Kovalevskaya also nurtured a parallel career in literature, penning several novels and a drama, "The Struggle for Happiness" that was favorably received at the Korsh Theater in Moscow. She died at the pinnacle of her scientific career from a combination of influenza and pneumonia less than two years after her election to both the Swedish and the Russian Academies of Sciences. The latter membership being initiated by Chebyshev, in spite of the Tsarist government's repeated refusal to grant her a university position in her own country.

### 21.1.2 Second-Order PDEs in $m$ Dimensions

Because of their importance in mathematical physics, the rest of this chapter and the next will be devoted to SOPDEs. This subsection classifies SOPDEs and the BCs associated with them.

The most general linear SOPDE in  $m$  variables can be written as

$$\sum_{j,k=1}^m A_{jk}(\mathbf{x}) \frac{\partial^2 u}{\partial x_j \partial x_k} + \sum_{j=1}^m B_j(\mathbf{x}) \frac{\partial u}{\partial x_j} + C(\mathbf{x})u = 0,$$

where  $A_{jk}$  can be assumed to be symmetric in  $j$  and  $k$ . We restrict ourselves to the simpler case in which the matrix  $(A_{jk})$  is diagonal. We therefore consider the PDE

$$\sum_{j=1}^m a_j(\mathbf{x}) \frac{\partial^2 u}{\partial x_j^2} + F\left(\mathbf{x}, u, \frac{\partial u}{\partial \mathbf{x}}\right), \quad (21.5)$$

where the last term collects all the other terms except the second derivatives. We classify SOPDEs as follows:

- |  |  |
|--|--|
| second-order PDE of elliptic type                        | 1. Equation (21.5) is said to be of <b>elliptic type</b> at $\mathbf{x}_0$ if all the coefficients $a_j(\mathbf{x}_0)$ are nonzero and have the same sign.   |
| second-order PDEs of hyperbolic and ultrahyperbolic type | 2. Equation (21.5) is said to be of <b>ultrahyperbolic type</b> at $\mathbf{x}_0$ if all $a_j(\mathbf{x}_0)$ are nonzero but do not have the same sign. If only one of the coefficients has a sign different from the rest, the equation is said to be of <b>hyperbolic type</b> . |
| second-order PDE of parabolic type                       | 3. Equation (21.5) is said to be of <b>parabolic type</b> at $\mathbf{x}_0$ if at least one of the coefficients $a_j(\mathbf{x}_0)$ is zero.   |

If a SOPDE is of a given type at every point of its domain, it is said to be of that given type. In particular, if the coefficients  $a_j$  are constants, the type of the PDE does not change from point to point.

**21.1.8. Example.** In this example, we study the SOPDE in two dimensions. The most general linear SOPDE is

$$L[u] = a \frac{\partial^2 u}{\partial x^2} + 2b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} + F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0, \tag{21.6}$$

where  $a, b,$  and  $c$  are functions of  $x$  and  $y$ .

To determine the characteristic curves of  $L$ , we seek conditions under which all second-order partial derivatives of  $u$  can be determined from the DE and the Cauchy data, which are values of  $u$  and all its first derivatives on  $\Gamma$ . Consider a point  $Q \equiv (x + dx, y + dy)$  close to  $P \equiv (x, y)$ . We can write

$$\begin{aligned} \frac{\partial u}{\partial x}(Q) - \frac{\partial u}{\partial x}(P) &= dx \frac{\partial^2 u}{\partial x^2}(P) + dy \frac{\partial^2 u}{\partial x \partial y}(P), \\ \frac{\partial u}{\partial y}(Q) - \frac{\partial u}{\partial y}(P) &= dx \frac{\partial^2 u}{\partial x \partial y}(P) + dy \frac{\partial^2 u}{\partial y^2}(P), \\ -F\left(P, u(P), \frac{\partial u}{\partial x}(P), \frac{\partial u}{\partial y}(P)\right) &= a(P) \frac{\partial^2 u}{\partial x^2}(P) + 2b(P) \frac{\partial^2 u}{\partial x \partial y}(P) + c(P) \frac{\partial^2 u}{\partial y^2}(P). \end{aligned}$$

This system of three linear equations in the three unknowns—the three second derivatives evaluated at  $P$ —has a unique solution if and only if the determinant of the coefficients is nonzero. Thus, by Box 21.1.6,  $\Gamma$  is a characteristic curve if and only if

$$\det \begin{pmatrix} dx & dy & 0 \\ 0 & dx & dy \\ a(P) & 2b(P) & c(P) \end{pmatrix} = 0,$$

or  $a(x, y)(dy)^2 - 2b(x, y)dx dy + c(x, y)(dx)^2 = 0$ . It then follows, assuming that  $a(x, y) \neq 0$ , that

$$\frac{dy}{dx} = \frac{b \pm \sqrt{b^2 - ac}}{a}. \tag{21.7}$$

There are three cases to consider:

1. If  $b^2 - ac < 0$ , Equation (21.7) has no solution, which implies that no characteristic curves exist at  $P$ . Problem 21.1 shows that the SOPDE is of elliptic type. Thus, the Laplace equation in two dimensions is elliptic because  $b^2 - ac = -1$ . In fact, it is elliptic in the whole plane, or, stated differently, it has no characteristic curve in the entire  $xy$ -plane. This may lead us to believe that the Cauchy problem for the Laplace equation in two dimensions has a unique solution. However, even though the absence of a characteristic hypersurface at  $P$  is a necessary condition for the existence of a solution to the Cauchy problem, it is not sufficient. Problem 21.4 presents a Cauchy problem that is **ill-posed**, meaning that the solution at any fixed point is not a continuous function of the initial data. Satisfying this condition is required of a well-posed problem on both mathematical and physical grounds.
2. If  $b^2 - ac > 0$ , Equation (21.7) has two solutions; that is, there are two characteristic curves passing through  $P$ . Problem 21.1 shows that the SOPDE is of hyperbolic type. The wave equation is such an equation in the entire  $\mathbb{R}^2$ .

ill-posed Cauchy problem



3. If  $b^2 - ac = 0$ , Equation (21.7) has only one solution. In this case there is only one characteristic curve at  $P$ . The SOPDE is parabolic in this case. The one-dimensional diffusion equation is an example of an SOPDE that is parabolic in the entire  $\mathbb{R}^2$ . ■

appropriate BCs are determined by the type of PDE

The question of what type of BCs to use to obtain a unique solution for a PDE is a very intricate mathematical problem. As Problem 21.4 shows, even though it has no characteristic curves in the entire  $\mathbb{R}^2$ , the two-dimensional Laplace equation does not lead to a well-posed Cauchy problem. On the other hand, examples in Chapter 19 that dealt with electrostatic potentials and temperatures led us to believe that a specification of the solution  $u$  on a *closed* curve in 2D, and a closed surface in 3D, gives a unique solution. This has a sound physical basis. After all, specifying the temperature (or electrostatic potential) on a closed surface should be enough to give us information about the temperature (or electrostatic potential) in the region close to the curve. A boundary condition in which the value of the solution is given on a closed hypersurface is called a **Dirichlet boundary condition**, and the associated problem, a **Dirichlet BVP**.

Dirichlet boundary condition and boundary value problem

There is another type of BC, which on physical grounds is appropriate for the Laplace equation. This condition is based on the fact that if the surface charge on a conductor is specified, then the electrostatic potential in the vicinity of the conductor can be determined uniquely. The surface charge on a conductor is proportional to the value of the electric field on the conductor. The electric field, on the other hand, is the normal derivative of the potential. A boundary condition in which the value of the normal derivative of the solution is specified on a closed curve is called a **Neumann BC**, and the associated problem, a **Neumann boundary value problem**. Thus, at least on physical grounds, either a Dirichlet BVP or a Neumann BVP is a well-posed problem for the Laplace equation.

Neumann boundary condition and boundary value problem

boundary conditions appropriate for elliptic, hyperbolic, and parabolic PDEs

For the heat (or diffusion) equation we are given an initial temperature distribution  $f(x)$  on a bar along, say the  $x$ -axis, with end points held at constant temperatures. For a bar with end points at  $x = a$  and  $x = b$ , this is equivalent to the data  $u(0, x) = f(x)$ ,  $u(t, a) = T_1$ , and  $u(t, b) = T_2$ . These are not Cauchy data, so we need not worry about characteristic curves. The boundary curve consists of three parts: (1)  $t = 0$  for  $a \leq x \leq b$ , (2)  $t > 0$  for  $x = a$ , and (3)  $t > 0$ , for  $x = b$ . In the  $xt$ -plane, these form an open rectangle consisting of  $\overline{ab}$  as one side and vertical lines at  $a$  and  $b$  as the other two. The problem is to determine  $u$  on the side that closes the rectangle, that is, on the side  $a \leq x \leq b$  at  $t > 0$ .

The wave equation requires specification of both  $u$  and  $\partial u / \partial t$  at  $t = 0$ . The displacement of the boundaries of the waving medium—a taut rope for example—must also be specified. Again the curve is open, as for the diffusion case, but the initial data are Cauchy. Thus, for the wave equation we do have a Cauchy problem with Cauchy data specified on an open curve. Since the curve, the open rectangle, is not a characteristic curve of the wave equation, the Cauchy problem is well-posed. We can generalize these BCs to  $m$  dimensions and make the following correspondence between a SOPDE with  $m$  variables and the appropriate BCs:

1. Elliptic SOPDE  $\leftrightarrow$  Dirichlet or Neumann BCs on a closed hypersurface.

- 2. Hyperbolic SOPDE ↔ Cauchy data on an open hypersurface.
- 3. Parabolic SOPDE ↔ Dirichlet or Neumann BCs on an open hypersurface.

## 21.2 Multidimensional GFs and Delta Functions

This section will discuss some of the characteristics of Green’s functions in higher dimensions. These characteristics are related to the formal partial differential operator associated with the Green’s function and also to the delta functions.

Using the formal idea of several continuous indices, we can turn the operator equation  $\mathbf{L}G = \mathbf{1}$  into the PDE

$$\mathbf{L}_x G(\mathbf{x}, \mathbf{y}) = \frac{\delta(\mathbf{x} - \mathbf{y})}{w(\mathbf{x})}, \tag{21.8}$$

where  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ ,  $w(\mathbf{x})$  is a weight function that is usually set equal to one, and, only in Cartesian coordinates,

$$\delta(\mathbf{x} - \mathbf{y}) = \delta(x_1 - y_1)\delta(x_2 - y_2) \cdots \delta(x_m - y_m) = \prod_{i=1}^m \delta(x_i - y_i). \tag{21.9}$$

In most applications Cartesian coordinates are not the most convenient to use. Therefore, it is helpful to express Equations (21.8) and (21.9) in other coordinate systems. In particular, it is helpful to know how the delta function transforms under a general coordinate transformation.

Let  $x_i = f_i(\xi_1, \dots, \xi_m), i = 1, 2, \dots, m$ , be a coordinate transformation. Let  $P$  be a point whose coordinates are  $\mathbf{a} = (a_1, \dots, a_m)$  and  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$  in the  $x$  and  $\xi$  coordinate systems, respectively. Let  $J$  be the Jacobian of the transformation, that is, the absolute value of the determinant of a matrix whose elements are  $\partial x_i / \partial \xi_j$ . For a function  $F(\mathbf{x})$  the definition of the delta function gives  $\int d^m x F(\mathbf{x})\delta(\mathbf{x} - \mathbf{a}) = F(\mathbf{a})$ . Expressing this equation in terms of the  $\xi$  coordinate system, recalling that  $d^m x = J d^m \xi$  and  $a_i = f_i(\boldsymbol{\alpha})$ , and introducing the notation  $H(\boldsymbol{\xi}) \equiv F(f_1(\boldsymbol{\xi}), \dots, f_m(\boldsymbol{\xi}))$ , we obtain

$$\int d^m \xi J H(\boldsymbol{\xi}) \prod_{i=1}^m \delta(f_i(\boldsymbol{\xi}) - f_i(\boldsymbol{\alpha})) = H(\boldsymbol{\alpha}). \tag{21.10}$$

This suggests that  $J \prod_{i=1}^m \delta(f_i(\boldsymbol{\xi}) - f_i(\boldsymbol{\alpha})) = \prod_{i=1}^m \delta(\xi_i - \alpha_i)$ , or, in more compact notation,  $J\delta(\mathbf{x} - \mathbf{a}) = \delta(\boldsymbol{\xi} - \boldsymbol{\alpha})$ . It is, of course, understood that  $J \neq 0$  at  $P$ . What happens when  $J = 0$  at  $P$ ? A point at which the Jacobian vanishes is called a **singular point** of the transformation. Thus, all points on the  $z$ -axis, including the origin, are singular points of Cartesian–spherical transformation. Since  $J$  is a determinant, its vanishing at a point signals lack of invertibility at that point.

singular point of a transformation

ignorable  
coordinates

Thus, in the transformation from Cartesian to spherical coordinates, all spherical coordinates  $(\rho, \theta, \varphi)$ , with arbitrary  $\varphi$ , are mapped to the Cartesian coordinates  $(0, 0, -\rho)$ . Similarly, the point  $(0, 0, 0)$  in the Cartesian coordinate system goes to  $(0, \theta, \varphi)$  in the spherical system, with  $\theta$  and  $\varphi$  arbitrary. A coordinate whose value is not determined at a singular point is called an **ignorable coordinate** at that point. Thus, at the origin both  $\theta$  and  $\varphi$  are ignorable.

Among the  $\xi$  coordinates, let  $\{\xi_i\}_{i=k+1}^m$  be ignorable at  $P$  with Cartesian coordinates  $\mathbf{a}$ . This means that any function, when expressed in terms of  $\xi$ 's, will be independent of the ignorable coordinates. A reexamination of Equation (21.10) reveals that (see Problem 21.8)

$$\delta(\mathbf{x} - \mathbf{a}) = \frac{1}{|J_k|} \prod_{i=1}^k \delta(\xi_i - \alpha_i), \quad \text{where} \quad J_k = \int J d\xi_{k+1} \cdots d\xi_m. \tag{21.11}$$

In particular, if the transformation is invertible,  $k = m$  and  $J_m = J$ , and we recover  $J\delta(\mathbf{x} - \mathbf{a}) = \delta(\xi - \alpha)$ .

**21.2.1. Example.** In two dimensions the transformation between Cartesian and polar coordinates is given by  $x_1 \equiv x = r \cos \theta \equiv \xi_1 \cos \xi_2$ ,  $x_2 \equiv y = r \sin \theta \equiv \xi_1 \sin \xi_2$  with the Jacobian

$$J = \det \begin{pmatrix} \partial x_1 / \partial \xi_1 & \partial x_1 / \partial \xi_2 \\ \partial x_2 / \partial \xi_1 & \partial x_2 / \partial \xi_2 \end{pmatrix} = \det \begin{pmatrix} \cos \xi_2 & -\xi_1 \sin \xi_2 \\ \sin \xi_2 & \xi_1 \cos \xi_2 \end{pmatrix} = \xi_1 = r,$$

which vanishes at the origin. The angle  $\theta$  is the only ignorable coordinate at the origin. Thus,  $k = 2 - 1 = 1$ , and

$$J_1 = \int_0^{2\pi} J d\theta = \int_0^{2\pi} r d\theta = 2\pi r \Rightarrow \delta(\mathbf{x}) \equiv \delta(x)\delta(y) = \frac{\delta(r)}{2\pi r}.$$

In three dimensions, the transformation between Cartesian and spherical coordinates yields Jacobian  $J = r^2 \sin \theta$ . This vanishes at the origin regardless of the values of  $\theta$  and  $\varphi$ . We thus have two ignorable coordinates at the origin (therefore,  $k = 3 - 2 = 1$ ), over which we integrate to obtain

$$J_1 = \int_0^{2\pi} d\varphi \int_0^\pi d\theta r^2 \sin \theta = 4\pi r^2 \Rightarrow \delta(\mathbf{x}) = \frac{\delta(r)}{4\pi r^2}. \quad \blacksquare$$

### 21.2.1 Spherical Coordinates in $m$ Dimensions

In discussing Green's functions in  $m$  dimensions, a particular curvilinear coordinate system will prove useful. This system is the generalization of spherical coordinates in three dimensions. The  $m$ -dimensional spherical coordinate system

is defined as

$$\begin{aligned}
 x_1 &= r \sin \theta_1 \cdots \sin \theta_{m-1}, \\
 x_2 &= r \sin \theta_1 \cdots \sin \theta_{m-2} \cos \theta_{m-1}, \\
 &\vdots \\
 x_k &= r \sin \theta_1 \cdots \sin \theta_{m-k} \cos \theta_{m-k+1}, \quad 2 \leq k \leq m-1, \\
 &\vdots \\
 x_m &= r \cos \theta_1.
 \end{aligned} \tag{21.12}$$

(Note that for  $m = 3$ , the first two Cartesian coordinates are switched compared to their usual definitions.)

It is not hard to show (see Example 21.2.2) that the Jacobian of the transformation (21.12) is

$$J = r^{m-1} (\sin \theta_1)^{m-2} (\sin \theta_2)^{m-3} \cdots (\sin \theta_k)^{m-k-1} \cdots \sin \theta_{m-2} \tag{21.13}$$

and that the volume element in terms of these coordinates is

$$d^m x = J dr d\theta_1 \cdots d\theta_{m-1} = r^{m-1} dr d\Omega_m, \tag{21.14}$$

element of the  
 $m$ -dimensional solid  
angle

where

$$d\Omega_m = (\sin \theta_1)^{m-2} (\sin \theta_2)^{m-3} \cdots \sin \theta_{m-2} d\theta_1 d\theta_2 \cdots d\theta_{m-1} \tag{21.15}$$

is the element of the  $m$ -dimensional solid angle.

**21.2.2. Example.** For  $m = 4$  we have

$$\begin{aligned}
 x_1 &= r \sin \theta_1 \sin \theta_2 \sin \theta_3, & x_2 &= r \sin \theta_1 \sin \theta_2 \cos \theta_3, \\
 x_3 &= r \sin \theta_1 \cos \theta_2, & x_4 &= r \cos \theta_1,
 \end{aligned}$$

and the Jacobian is given by

$$J = \det \begin{pmatrix} \partial x_1 / \partial r & \partial x_1 / \partial \theta_1 & \partial x_1 / \partial \theta_2 & \partial x_1 / \partial \theta_3 \\ \partial x_2 / \partial r & \partial x_2 / \partial \theta_1 & \partial x_2 / \partial \theta_2 & \partial x_2 / \partial \theta_3 \\ \partial x_3 / \partial r & \partial x_3 / \partial \theta_1 & \partial x_3 / \partial \theta_2 & \partial x_3 / \partial \theta_3 \\ \partial x_4 / \partial r & \partial x_4 / \partial \theta_1 & \partial x_4 / \partial \theta_2 & \partial x_4 / \partial \theta_3 \end{pmatrix} = r^3 \sin^2 \theta_1 \sin \theta_2.$$

It is readily seen (one can use mathematical induction to prove it rigorously) that the Jacobians for  $m = 2$  ( $J = r$ ),  $m = 3$  ( $J = r^2 \sin \theta_1$ ), and  $m = 4$  ( $J = r^3 \sin^2 \theta_1 \sin \theta_2$ ) generalize to Equation (21.13). ■

Using the integral  $\int_0^\pi \sin^n \theta d\theta = \sqrt{\pi} \Gamma[(n+1)/2] / \Gamma[(n+2)/2]$ , the total solid angle in  $m$  dimensions can be found to be

$$\Omega_m = \frac{2\pi^{m/2}}{\Gamma(m/2)}. \tag{21.16}$$

An interesting result that is readily obtained is an expression of the delta function in terms of spherical coordinates at the origin. Since  $r = 0$ , Equation (21.12) shows that all the angles are ignorable. Thus, we have

$$J_1 = \int J d\theta_1 \cdots d\theta_{m-1} = r^{m-1} \int d\Omega_m = r^{m-1} \Omega_m,$$

which yields

$$\delta(\mathbf{x}) = \delta(x_1) \cdots \delta(x_m) = \frac{\delta(r)}{\Omega_m r^{m-1}} = \frac{\Gamma(m/2)\delta(r)}{2\pi^{m/2}r^{m-1}}. \quad (21.17)$$

### 21.2.2 Green's Function for the Laplacian

With the machinery developed above, we can easily obtain the (indefinite) Green's function for the Laplacian in  $m$  dimensions. We will ignore questions of BCs and simply develop a function that satisfies  $\nabla^2 G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y})$ . Without loss of generality we let  $\mathbf{y} = 0$ ; that is, we translate the axes so that  $\mathbf{y}$  becomes the new origin. Then we have  $\nabla^2 G(\mathbf{x}) = \delta(\mathbf{x})$ . In spherical coordinates this becomes

$$\nabla^2 G(\mathbf{x}) = \frac{\delta(r)}{\Omega_m r^{m-1}}. \quad (21.18)$$

Since the RHS is a function of  $r$  only, we expect  $G$  to behave in the same way. We now have to express  $\nabla^2$  in terms of spherical coordinates. In general, this is difficult; however, for a function of  $r = \sqrt{x_1^2 + \cdots + x_m^2}$  alone, such as  $F(r)$ , we have

$$\frac{\partial F}{\partial x_i} = \frac{\partial F}{\partial r} \frac{\partial r}{\partial x_i} = \frac{\partial F}{\partial r} \frac{x_i}{r} \quad \text{and} \quad \frac{\partial^2 F}{\partial x_i^2} = \frac{\partial^2 F}{\partial r^2} \frac{x_i^2}{r^2} + \frac{\partial F}{\partial r} \left( \frac{1}{r} - \frac{x_i^2}{r^3} \right),$$

so that

$$\nabla^2 F(r) = \sum_{i=1}^m \frac{\partial^2 F}{\partial x_i^2} = \frac{\partial^2 F}{\partial r^2} + \frac{m-1}{r} \frac{\partial F}{\partial r} = \frac{1}{r^{m-1}} \frac{\partial}{\partial r} \left( r^{m-1} \frac{\partial F}{\partial r} \right).$$

For the Green's function, therefore, we get

$$\frac{d}{dr} \left( r^{m-1} \frac{dG}{dr} \right) = \frac{\delta(r)}{\Omega_m}. \quad (21.19)$$

The solution, for  $m \geq 3$ , is (see Problem 21.9)

$$G(r) = -\frac{\Gamma(m/2)}{2(m-2)\pi^{m/2}} \left( \frac{1}{r^{m-2}} \right) \quad \text{for } m \geq 3. \quad (21.20)$$

Green's function for the Laplacian

We can restore the vector  $\mathbf{y}$ , at which we placed the origin, by noting that  $r = |\mathbf{r}| = |\mathbf{x} - \mathbf{y}|$ . Thus, we get

$$G(\mathbf{x}, \mathbf{y}) = -\frac{\Gamma(m/2)}{2(m-2)\pi^{m/2}} \left( \frac{1}{|\mathbf{x} - \mathbf{y}|^{m-2}} \right) \\ = -\frac{\Gamma(m/2)}{2(m-2)\pi^{m/2}} \left[ \sum_{i=1}^m (x_i - y_i)^2 \right]^{-(m-2)/2} \quad \text{for } m \geq 3. \quad (21.21)$$

Similarly, for  $m = 2$  we obtain

$$G(\mathbf{x}, \mathbf{y}) = \frac{1}{2\pi} \ln |\mathbf{x} - \mathbf{y}| = \frac{1}{4\pi} \ln[(x_1 - y_1)^2 + (x_2 - y_2)^2]. \quad (21.22)$$

solution of Poisson equation in  $m$  dimensions

Having found the Green's function for the Laplacian, we can find a solution to the inhomogeneous equation, the *Poisson equation*,  $\nabla^2 u = -\rho(\mathbf{x})$ . Thus, for  $m \geq 3$ , we get

$$u(\mathbf{x}) = -\int d^m y G(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) = \frac{\Gamma(m/2)}{2(m-2)\pi^{m/2}} \int d^m y \frac{\rho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|^{m-2}}.$$

In particular, for  $m = 3$ , we obtain

$$u(\mathbf{x}) = \frac{1}{4\pi} \int d^3 y \frac{\rho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|},$$

which is the electrostatic potential due to a charge density  $\rho(\mathbf{y})$ .

### 21.3 Formal Development

The preceding section was devoted to a discussion of the Green's function for the Laplacian with no mention of the BCs. This section will develop a formalism that not only works for more general operators, but also incorporates the BCs.

#### 21.3.1 General Properties

Basic to a study of GFs is Green's identity, whose 1-dimensional version we encountered in Chapter 20. Here, we generalize it to  $m$  dimensions. Suppose there exist two differential operators,  $L_x$  and  $L_x^\dagger$ , which for any two functions  $u$  and  $v$ , satisfy the following relation:

$$v^* L_x[u] - u(L_x^\dagger[v])^* = \nabla \cdot \mathbf{Q}[u, v^*] \equiv \sum_{i=1}^m \frac{\partial Q_i}{\partial x_i} [u, v^*]. \quad (21.23)$$

The differential operator  $L_x^\dagger$  is—as in the one-dimensional case—called the formal adjoint of  $L_x$ . Integrating (21.23) over a closed domain  $D$  in  $\mathbb{R}^m$  with boundary

$\partial D$ , and using the divergence theorem, we obtain

$$\int_D d^m x \{v^* \mathbf{L}_x[u] - u(\mathbf{L}_x^\dagger[v])^*\} = \int_{\partial D} \mathbf{Q} \cdot \hat{\mathbf{e}}_n da, \tag{21.24}$$

generalized Green's identity

where  $\hat{\mathbf{e}}_n$  is an  $m$ -dimensional unit vector normal to  $\partial D$ , and  $da$  is an element of "area" of the  $m$ -dimensional hypersurface  $\partial D$ . Equation (21.24) is the **generalized Green's identity** for  $m$  dimensions. Note that the weight function is set equal to one for simplicity.

The differential operator  $\mathbf{L}_x$  is said to be formally self-adjoint if the RHS of Equation (21.24), the surface term, vanishes. In such a case, we have  $\mathbf{L}_x = \mathbf{L}_x^\dagger$  as in one dimension. This relation is a necessary condition for the surface term to vanish because  $u$  and  $v$  are, by assumption, arbitrary.  $\mathbf{L}_x$  is called self-adjoint (or, somewhat imprecisely, hermitian) if  $\mathbf{L}_x = \mathbf{L}_x^\dagger$  and the domains of the two operators, as determined by the vanishing of the surface term, are identical.

We can use Equation (21.24) to study the pair of PDEs

$$\mathbf{L}_x[u] = f(\mathbf{x}) \quad \text{and} \quad \mathbf{L}_x^\dagger[v] = h(\mathbf{x}). \tag{21.25}$$

As in one dimension, we let  $G(\mathbf{x}, \mathbf{y})$  and  $g(\mathbf{x}, \mathbf{y})$  denote the Green's functions for  $\mathbf{L}_x$  and  $\mathbf{L}_x^\dagger$ , respectively. Let us assume that the BCs are such that the surface term in Equation (21.24) vanishes. Then we get Green's identity

$$\int_D d^m x v^* \mathbf{L}_x[u] = \int_D d^m x u(\mathbf{L}_x^\dagger[v])^*. \tag{21.26}$$

If in this equation we let  $u = G(\mathbf{x}, \mathbf{t})$  and  $v = g(\mathbf{x}, \mathbf{y})$ , where  $\mathbf{t}, \mathbf{y} \in D$ , we obtain

$$\int_D d^m x g^*(\mathbf{x}, \mathbf{y}) \delta(\mathbf{x} - \mathbf{t}) = \int_D d^m x G(\mathbf{x}, \mathbf{t}) \delta(\mathbf{x} - \mathbf{y}),$$

or  $g^*(\mathbf{t}, \mathbf{y}) = G(\mathbf{y}, \mathbf{t})$ . In particular, when  $\mathbf{L}_x$  is formally self-adjoint, we have  $G^*(\mathbf{t}, \mathbf{y}) = G(\mathbf{y}, \mathbf{t})$ , or  $G(\mathbf{t}, \mathbf{y}) = G(\mathbf{y}, \mathbf{t})$ , if all the coefficient functions of  $\mathbf{L}_x$  are real. That is, the Green's function will be symmetric.

Green's functions are symmetric functions of their arguments

If we let  $v = g(\mathbf{x}, \mathbf{y})$  and use the first equation of (21.25) in (21.26), we get  $u(\mathbf{y}) = \int_D d^m x g^*(\mathbf{x}, \mathbf{y}) f(\mathbf{x})$ , which, using  $g^*(\mathbf{t}, \mathbf{y}) = G(\mathbf{y}, \mathbf{t})$  and interchanging  $\mathbf{x}$  and  $\mathbf{y}$ , becomes  $u(\mathbf{x}) = \int_D d^m y G(\mathbf{x}, \mathbf{y}) f(\mathbf{y})$ . It can similarly be shown that  $v(\mathbf{x}) = \int_D d^m y g(\mathbf{x}, \mathbf{y}) h(\mathbf{y})$ .

### 21.3.2 Fundamental (Singular) Solutions

The inhomogeneous term of the differential equation to which  $G(\mathbf{x}, \mathbf{y})$  is a solution is the delta function,  $\delta(\mathbf{x} - \mathbf{y})$ . It would be surprising if  $G(\mathbf{x}, \mathbf{y})$  did not "take notice" of this catastrophic source term and did not adapt itself to behave differently at  $\mathbf{x} = \mathbf{y}$  than at any other "ordinary" point. We noted the singular behavior of the Green's function at  $x = y$  in one dimension when we proved Theorem 20.3.4.

There we introduced  $h(x, y)$ —which was discontinuous at  $x = y$ —as a part of the Green's function. Similarly, when we discussed the Green's functions for the Laplacian in two and  $m$  dimensions earlier in this chapter, we noted that they behaved singularly at  $\mathbf{r} = 0$  or  $\mathbf{x} = \mathbf{y}$ . In this section, we study similar properties of the GFs for other differential operators.

Next to the Laplacian in difficulty is the formally self-adjoint elliptic PDO  $\mathbf{L}_x = \nabla^2 + q(\mathbf{x})$  discussed in Problem 21.10. Substituting this operator in the generalized Green's identity and using the expression for  $\mathbf{Q}$  given in Problem 21.10, we obtain

$$\int_D d^m x \{v \mathbf{L}_x[u] - u(\mathbf{L}_x[v])\} = \int_{\partial D} (v \hat{\mathbf{e}}_n \cdot \nabla u - u \hat{\mathbf{e}}_n \cdot \nabla v) da.$$

Letting  $v = G(\mathbf{x}, \mathbf{y})$  and denoting  $\hat{\mathbf{e}}_n \cdot \nabla$  by  $\partial/\partial n$  gives

$$\int_D d^m x [G \mathbf{L}_x u - u \mathbf{L}_x G] = \int_{\partial D} \left[ G \frac{\partial u}{\partial n} - u \frac{\partial G}{\partial n} \right] da. \quad (21.27)$$

We want to use this equation to find out about the behavior of  $G(\mathbf{x}, \mathbf{y})$  as  $|\mathbf{x} - \mathbf{y}| \rightarrow 0$ . Therefore, assuming that  $\mathbf{y} \in D$ , we divide the domain  $D$  into two parts: one part is a region  $D_\epsilon$  bounded by an infinitesimal hypersphere  $S_\epsilon$  with radius  $\epsilon$  and center at  $\mathbf{y}$ ; the other is the rest of  $D$ . Instead of  $D$  we use the region  $D' \equiv D - D_\epsilon$ . The following facts are easily deduced for  $D'$ : (1)  $\mathbf{L}_x G(\mathbf{x}, \mathbf{y}) = 0$  because  $\mathbf{x} \neq \mathbf{y}$  in  $D'$ ; (2)  $\int_D = \lim_{\epsilon \rightarrow 0} \int_{D'}$ ; (3)  $\partial D' = \partial D \cup S_\epsilon$ .

Suppose that we are interested in finding a solution to

$$\mathbf{L}_x[u] = [\nabla^2 + q(\mathbf{x})]u(\mathbf{x}) = f(\mathbf{x})$$

subject to certain, as yet unspecified, BCs. Using the three facts listed above, Equation (21.27) yields

$$\begin{aligned} \int_D d^m x [G \mathbf{L}_x u - u \mathbf{L}_x G] &= \lim_{\epsilon \rightarrow 0} \int_{D'} d^m x \left[ \underbrace{G \mathbf{L}_x u}_{=f} - u \underbrace{\mathbf{L}_x G}_{=0} \right] \\ &= \lim_{\epsilon \rightarrow 0} \int_{D'} d^m x G(\mathbf{x}, \mathbf{y}) f(\mathbf{x}) = \int_D d^m x G(\mathbf{x}, \mathbf{y}) f(\mathbf{x}) \\ &= \int_{\partial D} \left( G \frac{\partial u}{\partial n} - u \frac{\partial G}{\partial n} \right) da + \int_{S_\epsilon} \left( G \frac{\partial u}{\partial n} - u \frac{\partial G}{\partial n} \right) da. \end{aligned}$$

We assume that the BCs are such that the integral over  $\partial D$  vanishes. This is a generalization of the one-dimensional case (recall from Chapter 20 that this is a necessary condition for the existence of Green's functions). Moreover, for an  $m$ -dimensional sphere,  $da = r^{m-1} d\Omega_m$ , which for  $S_\epsilon$  reduces to  $\epsilon^{m-1} d\Omega_m$ . Substituting in the preceding equation yields

$$\int_D d^m x G(\mathbf{x}, \mathbf{y}) f(\mathbf{x}) = \int_{S_\epsilon} \left( G \frac{\partial u}{\partial n} - u \frac{\partial G}{\partial n} \right) \epsilon^{m-1} d\Omega_m.$$



We would like the RHS to be  $u(\mathbf{y})$ . This will be the case if

$$\lim_{\epsilon \rightarrow 0} \int_{S_\epsilon} G(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial n} \epsilon^{m-1} d\Omega_m = 0 \quad \text{and} \quad \lim_{\epsilon \rightarrow 0} \int_{S_\epsilon} u \frac{\partial G}{\partial n} \epsilon^{m-1} d\Omega_m = u(\mathbf{y})$$

for arbitrary  $u$ . This will happen only if

$$\lim_{r \rightarrow 0} G(\mathbf{y} + \mathbf{r}, \mathbf{y}) r^{m-1} = 0, \quad \lim_{r \rightarrow 0} \frac{\partial G}{\partial r}(\mathbf{y} + \mathbf{r}, \mathbf{y}) r^{m-1} = \text{const.} \quad (21.28)$$

A solution to these two equations is

$$G(\mathbf{x}, \mathbf{y}) = \begin{cases} -\frac{F(\mathbf{x}, \mathbf{y})}{2\pi_1} \ln(|\mathbf{x} - \mathbf{y}|) + H(\mathbf{x}, \mathbf{y}) & \text{if } m = 2, \\ -\frac{F(\mathbf{x}, \mathbf{y})}{(m-2)\Omega_m |\mathbf{x} - \mathbf{y}|^{m-2}} + H(\mathbf{x}, \mathbf{y}) & \text{if } m \geq 3, \end{cases} \quad (21.29)$$

where  $H(\mathbf{x}, \mathbf{y})$  and  $F(\mathbf{x}, \mathbf{y})$  are well behaved at  $\mathbf{x} = \mathbf{y}$ . The introduction of these functions is necessary because Equation (21.28) determines the behavior of  $G(\mathbf{x}, \mathbf{y})$  only when  $\mathbf{x} \approx \mathbf{y}$ . Such behavior does not uniquely determine  $G(\mathbf{x}, \mathbf{y})$ . For instance,  $e^{|\mathbf{x}-\mathbf{y}|} \ln(|\mathbf{x} - \mathbf{y}|)$  and  $\ln(|\mathbf{x} - \mathbf{y}|)$  behave in the same way as  $|\mathbf{x} - \mathbf{y}| \rightarrow 0$ .

Equation (21.29) shows that for  $\mathbf{L}_\mathbf{x} = \nabla^2 + q(\mathbf{x})$ , the Green's function consists of two parts. The first part determines the singular behavior of the Green's function as  $\mathbf{x} \rightarrow \mathbf{y}$ . The nature of this singularity (how badly the GF "blows up" as  $\mathbf{x} \rightarrow \mathbf{y}$ ) is extremely important, because it is a prerequisite for our ability to write the solution in terms of an integral representation with the Green's function as its kernel. Due to their importance in such representations, the first terms on the RHS of Equation (21.29) are called the **fundamental solution** of the differential equation, or the *singular part* of the Green's function.

fundamental solution  
is the singular part of  
GF

What about the second part of the Green's function? What role does it play in obtaining a solution? So far we have been avoiding consideration of BCs. Here  $H(\mathbf{x}, \mathbf{y})$  can help. We choose  $H(\mathbf{x}, \mathbf{y})$  in such a way that  $G(\mathbf{x}, \mathbf{y})$  satisfies the appropriate BCs. Let us discuss this in greater detail and generality.

If BCs are ignored, the Green's function for a SOPDO  $\mathbf{L}_\mathbf{x}$  cannot be determined uniquely. In particular, if  $G(\mathbf{x}, \mathbf{y})$  is a Green's function, that is, if  $\mathbf{L}_\mathbf{x}G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y})$ , then so is  $G(\mathbf{x}, \mathbf{y}) + H(\mathbf{x}, \mathbf{y})$  as long as  $H(\mathbf{x}, \mathbf{y})$  is a solution of the homogeneous equation  $\mathbf{L}_\mathbf{x}H(\mathbf{x}, \mathbf{y}) = 0$ . Thus, we can break the Green's function into two parts:

homogeneous  
solution is the  
regular part of GF

$$G = G_s + H, \quad \text{where } \mathbf{L}_\mathbf{x}G_s(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}), \quad \mathbf{L}_\mathbf{x}H(\mathbf{x}, \mathbf{y}) = 0 \quad (21.30)$$

with  $G_s$  the singular part of the Green's function.  $H$  is called the **regular part** of the Green's function. Neither  $G_s$  nor  $H$  (nor  $G$ , therefore) is unique. However, the appropriate BCs, which depend on the type of  $\mathbf{L}_\mathbf{x}$ , will determine  $G$  uniquely.

regular part of the  
Green's function

To be more specific, let us assume that we want to find a Green's function for  $\mathbf{L}_\mathbf{x}$  that vanishes at the boundary  $\partial D$ . That is, we wish to find  $G(\mathbf{x}, \mathbf{y})$  such that

$G(\mathbf{x}_b, \mathbf{y}) = 0$ , where  $\mathbf{x}_b$  is an arbitrary point of the boundary. All that is required is to find a  $G_s$  and an  $H$  satisfying Equation (21.30) with the BC  $H(\mathbf{x}_b, \mathbf{y}) = -G_s(\mathbf{x}_b, \mathbf{y})$ . The latter problem, involving a homogeneous differential equation, can be handled by the methods of Chapters 18 and 19. Since any discussion of BCs is tied to the type of PDE, we have reserved the discussion of such specifics for the next chapter.

### 21.4 Integral Equations and GFs

Integral equations are best applied in combination with Green's functions. In fact, we can use a Green's function to turn a DE into an integral equation. If this integral equation is compact or has a compact resolvent, then the problem lends itself to the methods described in Chapters 16 and 17.

Let  $L_x$  be a SOPDO in  $m$  variables. We are interested in solving the SOPDE  $L_x[u] + \lambda V(\mathbf{x})u(\mathbf{x}) = f(\mathbf{x})$  subject to some BCs. Here  $\lambda$  is an arbitrary constant, and  $V(\mathbf{x})$  is a well-behaved function on  $\mathbb{R}^m$ . Transferring the second term on the LHS to the RHS and then treating the RHS as an inhomogeneous term, we can write the "solution" to the PDE as

$$u(\mathbf{x}) = H(\mathbf{x}) + \int_D d^m y G_0(\mathbf{x}, \mathbf{y}) [f(\mathbf{y}) - \lambda V(\mathbf{y})u(\mathbf{y})],$$

where  $D$  is the domain of  $L_x$  and  $G_0$  is the Green's function for  $L_x$  with some, as yet unspecified, BCs. The function  $H$  is a solution to the homogeneous equation, and it is present to guarantee the appropriate BCs.

Combining the first term in the integral with  $H(\mathbf{x})$ , we have

$$u(\mathbf{x}) = F(\mathbf{x}) - \lambda \int_D d^m y G_0(\mathbf{x}, \mathbf{y}) V(\mathbf{y})u(\mathbf{y}). \tag{21.31}$$

Equation (21.31) is an  $m$ -dimensional Fredholm equation whose solution can be obtained in the form of a Neumann series.

**21.4.1. Example.** Consider the bound-state Schrödinger equation in one dimension:

$$-\frac{\hbar^2}{2\mu} \frac{d^2 \Psi}{dx^2} + V(x)\Psi(x) = E\Psi(x), \quad E < 0.$$

We rewrite this equation as

$$L_x[\Psi] \equiv \left( \frac{d^2}{dx^2} - \kappa^2 \right) \Psi(x) = \frac{2\mu}{\hbar^2} V(x)\Psi(x),$$

where  $\kappa^2 = -2\mu E/\hbar^2 > 0$ . Equation (21.31) gives the equivalent integral equation

$$\Psi(x) = \Psi_0(x) + \frac{2\mu}{\hbar^2} \int_{-\infty}^{\infty} G_0(x, y) V(y)\Psi(y) dy$$

where  $\Psi_0(x)$  is the solution of  $L_x[\Psi_0] = 0$ , which is easily found to be of the general form  $\Psi_0(x) = Ae^{\kappa x} + Be^{-\kappa x}$ . If we assume that  $\Psi_0(x)$  remains finite as  $x \rightarrow \pm\infty$ ,  $\Psi_0(x)$  will be zero. Furthermore, it can be shown that  $G_0(x, y) = -e^{-\kappa|x-y|}/2\kappa$  (see Problem 20.12). Therefore,

$$\Psi(x) = -\frac{\mu}{\hbar^2\kappa} \int_{-\infty}^{\infty} e^{-\kappa|x-y|} V(y) \Psi(y) dy.$$

Now consider an attractive delta-function potential with center at  $a$ :  $V(x) = -V_0\delta(x-a)$ ,  $V_0 > 0$ . For such a potential, the integral equation yields

$$\Psi(x) = \frac{\mu}{\hbar^2\kappa} \int_{-\infty}^{\infty} e^{-\kappa|x-y|} V_0\delta(y-a)\Psi(y) dy = \frac{\mu V_0}{\hbar^2\kappa} e^{-\kappa|x-a|}\Psi(a).$$

For this equation to be consistent, i.e., to get an identity when  $x = a$ , we must have

$$\frac{\mu V_0}{\hbar^2\kappa} = 1 \Rightarrow \kappa = \frac{\mu V_0}{\hbar^2} \Rightarrow E = -\frac{\mu V_0}{2\hbar^2}.$$

Therefore, there is only one bound state and one energy level for an attractive delta-function potential. ■

To find a Neumann-series solution we can substitute the expression for  $u$  given by the RHS of Equation (21.31) in the integral of that equation. The resulting equation will have two integrals, in the second of which  $u$  appears. Substituting the new  $u$  in the second integral and continuing the process  $N$  times yields

$$u(\mathbf{x}) = F(\mathbf{x}) + \sum_{n=1}^{N-1} (-\lambda)^n \int_D d^m y K^n(\mathbf{x}, \mathbf{y}) F(\mathbf{y}) \\ + (-\lambda)^N \int_D d^m y K^N(\mathbf{x}, \mathbf{y}) u(\mathbf{y}),$$

where

$$K(\mathbf{x}, \mathbf{y}) \equiv V(\mathbf{x})G_0(\mathbf{x}, \mathbf{y}), \\ K^n(\mathbf{x}, \mathbf{y}) \equiv \int_D d^m t K^{n-1}(\mathbf{x}, \mathbf{t})K(\mathbf{t}, \mathbf{y}) \quad \text{for } n \geq 2. \quad (21.32)$$

The Neumann series is obtained by letting  $N \rightarrow \infty$ :

$$u(\mathbf{x}) = F(\mathbf{x}) + \sum_{n=1}^{\infty} (-\lambda)^n \int_D d^m y K^n(\mathbf{x}, \mathbf{y}) F(\mathbf{y}). \quad (21.33)$$

Except for the fact that here the integrations are in  $m$  variables, Equation (21.33) is the same as the Neumann series derived in Section 17.1. In exact analogy, therefore, we abbreviate (21.33) as

$$|u\rangle = |F\rangle + \sum_{n=1}^{\infty} (-\lambda)^n \mathbf{K}^n |F\rangle. \quad (21.34)$$

Equations (21.33) and (21.34) have meaning only if the Neumann series converges, i.e., if

$$|\lambda| \left[ \int_D d^m y \int_D d^m x |K(\mathbf{x}, \mathbf{y})|^2 \right]^{1/2} < 1. \tag{21.35}$$

Feynman's  
diagrammatic  
representation of GF

We will briefly discuss an intuitive physical interpretation of the Neumann series due to Feynman. Although Feynman developed this *diagrammatic technique* for quantum electrodynamics, it has been useful in other areas, such as statistical and condensed matter physics. In most cases of interest, the SOPDE is homogeneous, so  $f(\mathbf{x}) = 0$ . In that case,  $L_x$  and  $V(\mathbf{x})$  are called the *free operator* and the *interacting potential*, respectively. The solution to  $L_x[u] = 0$  is called the *free solution* and denoted by  $u_f(\mathbf{x})$ .

Let us start with Equation (21.31) written as

$$u(\mathbf{x}) = u_f(\mathbf{x}) - \lambda \int_{\mathbb{R}^m} d^m y G_0(\mathbf{x}, \mathbf{y}) V(\mathbf{y}) u(\mathbf{y}), \tag{21.36}$$

where  $G_0$  stands for the Green's function for the free operator  $L_x$ . The full Green's function, that is, that for  $L_x + \lambda V$ , will be denoted by  $G$ . Moreover, as is usually the case, the region  $D$  has been taken to be all of  $\mathbb{R}^m$ . This implies that no boundary conditions are imposed on  $u$ , which in turn permits us to use the singular part of the Green's function in the integral. Because of the importance of the full Green's function, we are interested in finding a series for  $G$  in terms of  $G_0$ , which is supposed to be known. To obtain such a series we start with the abstract operator equation and write  $\mathbf{G} = \mathbf{G}_0 + \mathbf{A}$ , where  $\mathbf{A}$  is to be determined. Operating on both sides with  $\mathbf{L}$  ("inverse" of  $\mathbf{G}_0$ ), we obtain  $\mathbf{L}\mathbf{G} = \mathbf{L}\mathbf{G}_0 + \mathbf{L}\mathbf{A} = \mathbf{1} + \mathbf{L}\mathbf{A}$ . On the other hand,  $(\mathbf{L} + \lambda\mathbf{V})\mathbf{G} = \mathbf{1}$ , or  $\mathbf{L}\mathbf{G} = \mathbf{1} - \lambda\mathbf{V}\mathbf{G}$ . These two equations give

$$\mathbf{L}\mathbf{A} = -\lambda\mathbf{V}\mathbf{G} \Rightarrow \mathbf{A} = -\lambda\mathbf{L}^{-1}\mathbf{V}\mathbf{G} = -\lambda\mathbf{G}_0\mathbf{V}\mathbf{G}.$$

Therefore,

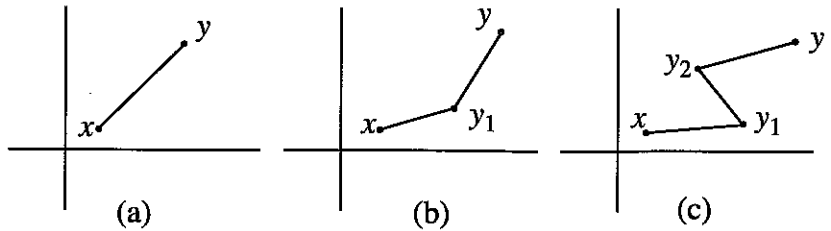
$$\mathbf{G} = \mathbf{G}_0 - \lambda\mathbf{G}_0\mathbf{V}\mathbf{G}. \tag{21.37}$$

Sandwiching both sides between  $\langle \mathbf{x} |$  and  $| \mathbf{z} \rangle$ , inserting  $\mathbf{1} = \int | \mathbf{y} \rangle \langle \mathbf{y} | d^m y$  between  $\mathbf{G}_0$  and  $\mathbf{V}$  and  $\mathbf{1} = \int | \mathbf{t} \rangle \langle \mathbf{t} | d^m t$  between  $\mathbf{V}$  and  $\mathbf{G}$ , and assuming that  $\mathbf{V}$  is local [i.e.,  $V(\mathbf{y}, \mathbf{t}) = V(\mathbf{y})\delta(\mathbf{y} - \mathbf{t})$ ], we obtain

$$G(\mathbf{x}, \mathbf{z}) = G_0(\mathbf{x}, \mathbf{z}) - \lambda \int d^m y G_0(\mathbf{x}, \mathbf{y}) V(\mathbf{y}) G(\mathbf{y}, \mathbf{z}). \tag{21.38}$$

This equation is the analogue of (21.31) and, just like that equation, is amenable to a Neumann series expansion. The result is

$$G(\mathbf{x}, \mathbf{y}) = G_0(\mathbf{x}, \mathbf{y}) + \sum_{n=1}^{\infty} (-\lambda)^n \int_{\mathbb{R}^m} d^m z G_0(\mathbf{x}, \mathbf{z}) K^n(\mathbf{z}, \mathbf{y}), \tag{21.39}$$



**Figure 21.1** Contributions to the full propagator in (a) the zeroth order, (b) the first order, and (c) the second order. At each vertex one introduces a factor of  $-\lambda V$  and integrates over all values of the variable of that vertex.

where  $K^n(\mathbf{x}, \mathbf{z})$  is as given in Equation (21.32).

Feynman's idea is to consider  $G(\mathbf{x}, \mathbf{y})$  as an interacting **propagator** between points  $\mathbf{x}$  and  $\mathbf{y}$  and  $G_0(\mathbf{x}, \mathbf{y})$  as a free propagator. The first term on the RHS of (21.39) is simply a free propagation from  $\mathbf{x}$  to  $\mathbf{y}$ . Diagrammatically, it is represented by a line joining the points  $\mathbf{x}$  and  $\mathbf{y}$  [see Figure 21.1(a)]. The second term is a free propagation from  $\mathbf{x}$  to  $\mathbf{y}_1$  (also called a **vertex**), interaction at  $\mathbf{y}_1$  with a potential  $-\lambda V(\mathbf{y}_1)$ , and subsequent free propagation to  $\mathbf{y}$  [see Figure 21.1(b)]. According to the third term, the particle or wave [represented by  $u_f(\mathbf{x})$ ] propagates freely from  $\mathbf{x}$  to  $\mathbf{y}_1$ , interacts at  $\mathbf{y}_1$  with the potential  $-\lambda V(\mathbf{y}_1)$ , propagates freely from  $\mathbf{y}_1$  to  $\mathbf{y}_2$ , interacts for a second time with the potential  $-\lambda V(\mathbf{y}_2)$ , and finally propagates freely from  $\mathbf{y}_2$  to  $\mathbf{y}$  [Figure 21.1(c)]. The interpretation of the rest of the series in (21.39) is now clear: The  $n$ th-order term of the series has  $n$  vertices between  $\mathbf{x}$  and  $\mathbf{y}$  with a factor  $-\lambda V(\mathbf{y}_k)$  and an integration over  $\mathbf{y}_k$  at vertex  $k$ . Between any two consecutive vertices  $\mathbf{y}_k$  and  $\mathbf{y}_{k+1}$  there is a factor of the free propagator  $G_0(\mathbf{y}_k, \mathbf{y}_{k+1})$ .

Feynman diagrams are used extensively in relativistic quantum field theory, for which  $m = 4$ , corresponding to the four-dimensional space-time. In this context  $\lambda$  is determined by the strength of the interaction. For quantum electrodynamics, for instance,  $\lambda$  is the fine-structure constant,  $e^2/\hbar c = 1/137$ .

## 21.5 Perturbation Theory

Few operator equations lend themselves to an exact solution, and due to the urgency of finding a solution to such equations in fundamental physics, various techniques have been developed to approximate solutions to operator equations. We have already seen instances of such techniques in, for example, the WKB method. This section is devoted to a systematic development of perturbation theory, which is one of the main tools of calculation in quantum mechanics. For a thorough treatment of perturbation theory along the lines presented here, see [Mess 66, pp. 712–720].

The starting point is the resolvent (Definition 16.8.1) of a Hamiltonian  $\mathbf{H}$ , which, using  $z$  instead of  $\lambda$ , we write as  $\mathbf{R}_z(\mathbf{H})$ . For simplicity, we assume that the eigenvalues of  $\mathbf{H}$  are discrete. This is a valid assumption if the Hamiltonian is compact or if we are interested in approximations close to one of the discrete eigenvalues. Denoting the eigenvalues of  $\mathbf{H}$  by  $\{E_i\}_{i=0}^\infty$ , we have

$$\mathbf{H}\mathbf{P}_i = E_i\mathbf{P}_i, \tag{21.40}$$

where  $\mathbf{P}_i$  is the projection operator to the  $i$ th eigenspace. We can write the resolvent in terms of the projection operators by using Equation (16.8):

$$\mathbf{R}_z(\mathbf{H}) = \sum_{i=0}^\infty \frac{\mathbf{P}_i}{E_i - z}. \tag{21.41}$$

The projection operator  $\mathbf{P}_i$  can be written as a contour integral as in Equation (16.13). Any sum of these operators can also be written as a contour integral. For instance, if  $\Gamma$  is a circle enclosing the first  $n + 1$  eigenvalues, then

$$\mathbf{P}_\Gamma \equiv \sum_{i=0}^n \mathbf{P}_i = -\frac{1}{2\pi i} \oint_\Gamma \mathbf{R}_z(\mathbf{H}) dz. \tag{21.42}$$

Multiplying Equation (21.42) by  $\mathbf{H}$  and using the definition of the resolvent, one can show that

$$\mathbf{H}\mathbf{P}_\Gamma = -\frac{1}{2\pi i} \oint_\Gamma z\mathbf{R}_z(\mathbf{H}) dz. \tag{21.43}$$

When  $\Gamma$  includes *all* eigenvalues of  $\mathbf{H}$ ,  $\mathbf{P}_\Gamma = \mathbf{1}$ , and Equation (21.43) reduces to (16.12) with  $A \rightarrow T$  and  $f(x) \rightarrow x$ .

To proceed, let us assume that  $\mathbf{H} = \mathbf{H}_0 + \lambda\mathbf{V}$  where  $\mathbf{H}_0$  is a Hamiltonian with known eigenvalues and eigenvectors, and  $\mathbf{V}$  is a *perturbing potential*;  $\lambda$  is a (small) parameter that keeps track of the order of approximation. Let us also use the abbreviations

$$\mathbf{G}(z) \equiv -\mathbf{R}_z(\mathbf{H}) \quad \text{and} \quad \mathbf{G}_0(z) \equiv -\mathbf{R}_z(\mathbf{H}_0). \tag{21.44}$$

Then a procedure very similar to that leading to Equation (21.37) yields

$$\mathbf{G}(z) = \mathbf{G}_0(z) + \lambda\mathbf{G}_0(z)\mathbf{V}\mathbf{G}(z), \tag{21.45}$$

which can be expanded in a Neumann series by iteration:

$$\mathbf{G}(z) = \sum_{n=0}^\infty \lambda^n \mathbf{G}_0(z) [\mathbf{V}\mathbf{G}_0(z)]^n. \tag{21.46}$$

Let  $\{E_a^0\}$ ,  $\{M_a^0\}$ , and  $m_a$  denote, respectively, the eigenvalues of  $\mathbf{H}_0$ , their corresponding eigenspaces, and the latter's dimensions.<sup>2</sup> In the context of perturbation

Degeneracy is the dimension of the eigenspace of the Hamiltonian.

theory,  $m_a$  is called the **degeneracy** of  $E_a^0$ , and  $E_a^0$  is called  $m_a$ -**fold degenerate**, with a similar terminology for the perturbed Hamiltonian. We assume that all eigenspaces have finite dimensions.

It is clear that eigenvalues and eigenspaces of  $\mathbf{H}$  will tend to those of  $\mathbf{H}_0$  when  $\lambda \rightarrow 0$ . So, let us collect all eigenspaces of  $\mathbf{H}$  that tend to  $\mathcal{M}_a^0$  and denote them by  $\{\mathcal{M}_i^a\}_{i=1}^{r_a}$ . Similarly, we use  $E_i^a$  and  $\mathbf{P}_i^a$  to denote, respectively, the energy eigenvalue and the projector to the eigenspace  $\mathcal{M}_i^a$ . Since dimension is a discrete quantity, it cannot depend on  $\lambda$ , and we have

$$\sum_{i=1}^{r_a} \dim \mathcal{M}_i^a = \dim \mathcal{M}_a^0 = m_a. \quad (21.47)$$

We also use the notation  $\mathbf{P}$  for the projector onto the direct sum of  $\mathcal{M}_i^a$ 's. We thus have

$$\mathbf{P} \equiv \sum_{i=1}^{r_a} \mathbf{P}_i^a \quad \text{and} \quad \lim_{\lambda \rightarrow 0} \mathbf{P} = \mathbf{P}_a^0, \quad (21.48)$$

where we have used an obvious notation for the projection operator onto  $\mathcal{M}_a^0$ .

The main task of perturbation theory is to find the eigenvalues and eigenvectors of the perturbed Hamiltonian in terms of a series in powers of  $\lambda$  of the corresponding unperturbed quantities. Since the eigenvectors—or, more appropriately, the projectors onto eigenspaces—and their corresponding eigenvalues of the perturbed Hamiltonian are related via Equation (21.40), this task reduces to writing  $\mathbf{P}$  as a series in powers of  $\lambda$  whose coefficients are operators expressible in terms of unperturbed quantities.

For sufficiently small  $\lambda$ , there exists a contour in the  $z$ -plane enclosing  $E_a^0$  and all  $E_i^a$ 's but excluding all other eigenvalues of  $\mathbf{H}$  and  $\mathbf{H}_0$ . Denote this contour by  $\Gamma_a$  and, using Equation (21.42), write

$$\mathbf{P} = \frac{1}{2\pi i} \oint_{\Gamma_a} \mathbf{G}(z) dz.$$

It follows from Equation (21.46) that

$$\mathbf{P} = \mathbf{P}_a^0 + \sum_{n=1}^{\infty} \lambda^n \mathbf{A}^{(n)}, \quad \text{where} \quad \mathbf{A}^{(n)} \equiv \frac{1}{2\pi i} \oint_{\Gamma_a} \mathbf{G}_0(z) [\mathbf{V}\mathbf{G}_0(z)]^n dz. \quad (21.49)$$

This equation shows that perturbation expansion is reduced to the calculation of  $\mathbf{A}^{(n)}$ , which is simply the residue of  $\mathbf{G}_0(z) [\mathbf{V}\mathbf{G}_0(z)]^n$ . The only singularity of the integrand in Equation (21.49) comes from  $\mathbf{G}_0(z)$ , which, by (21.44) and (21.41),

<sup>2</sup>We use the beginning letters of the Latin alphabet for the unperturbed Hamiltonian. Furthermore, we attach a superscript "0" to emphasize that the object belongs to  $\mathbf{H}_0$ .

has a pole at  $E_a^0$ . So, to calculate this residue, we simply expand  $\mathbf{G}_0(z)$  in a Laurent series about  $E_a^0$ :

$$\begin{aligned} \mathbf{G}_0(z) &= \sum_b \frac{\mathbf{P}_b^0}{z - E_b^0} = \frac{\mathbf{P}_a^0}{z - E_a^0} + \sum_{b \neq a} \frac{\mathbf{P}_b^0}{z - E_b^0} \\ &= \frac{\mathbf{P}_a^0}{z - E_a^0} + \sum_{b \neq a} \frac{\mathbf{P}_b^0}{(E_a^0 - E_b^0) \left(1 + \frac{z - E_a^0}{E_a^0 - E_b^0}\right)} \\ &= \frac{\mathbf{P}_a^0}{z - E_a^0} + \sum_{b \neq a} \sum_{k=0}^{\infty} (-1)^k \frac{(z - E_a^0)^k \mathbf{P}_b^0}{(E_a^0 - E_b^0)^{k+1}}. \end{aligned}$$

Switching the order of the two sums, and noting that our space is the Hilbert space of  $\mathbf{H}_0$  whose basis can be chosen to consist of eigenstates of  $\mathbf{H}_0$ , we can write  $\mathbf{H}_0$  instead of  $E_b^0$  in the denominator to obtain

$$\begin{aligned} \sum_{b \neq a} \frac{\mathbf{P}_b^0}{(E_a^0 - E_b^0)^{k+1}} &= \sum_{b \neq a} \frac{\mathbf{P}_b^0}{(E_a^0 - \mathbf{H}_0)^{k+1}} = \frac{\sum_{b \neq a} \mathbf{P}_b^0}{(E_a^0 - \mathbf{H}_0)^{k+1}} = \frac{\overbrace{1 - \mathbf{P}_a^0}^{\equiv \mathbf{Q}_a^0}}{(E_a^0 - \mathbf{H}_0)^{k+1}} \\ &= \frac{\mathbf{Q}_a^0}{(E_a^0 - \mathbf{H}_0)^{k+1}} \equiv \mathbf{G}_0^{k+1}(E_a^0) \mathbf{Q}_a^0 = \mathbf{Q}_a^0 \mathbf{G}_0^{k+1}(E_a^0) \mathbf{Q}_a^0, \end{aligned}$$

where we have used the completeness relation for the  $\mathbf{P}_b^0$ 's, the fact that  $\mathbf{Q}_a^0$  commutes with  $\mathbf{H}_0$  [and, therefore, with  $\mathbf{G}_0^{k+1}(E_a^0)$ ], and, in the last equality, the fact that  $\mathbf{Q}_a^0$  is a projection operator.<sup>3</sup> It follows that

$$\begin{aligned} \mathbf{G}_0(z) &= \frac{\mathbf{P}_a^0}{z - E_a^0} + \sum_{k=0}^{\infty} (-1)^k (z - E_a^0)^k \mathbf{Q}_a^0 \mathbf{G}_0^{k+1}(E_a^0) \mathbf{Q}_a^0 \\ &= \sum_{k=0}^{\infty} (-1)^k (z - E_a^0)^{k-1} \mathbf{S}^k, \end{aligned} \tag{21.50}$$

where we have introduced the notation

$$\mathbf{S}^k \equiv \begin{cases} \mathbf{P}_a^0 & \text{if } k = 0, \\ -\mathbf{Q}_a^0 \mathbf{G}_0^k(E_a^0) \mathbf{Q}_a^0 & \text{if } k \geq 1. \end{cases}$$

<sup>3</sup>Note that although  $\mathbf{G}_0(z)$  has a pole at  $E_a^0$ , the expressions in the last line of the equation above make sense because  $\mathbf{Q}_a^0$  annihilates all states with eigenvalue  $E_a^0$ . The reason for the introduction of  $\mathbf{Q}_a^0$  on both sides is to ensure that  $\mathbf{G}_0^{k+1}(E_a^0)$  will not act on an eigenstate of  $E_a^0$  on either side.



By substituting Equation (21.50) in  $\mathbf{G}_0(z)[\mathbf{V}\mathbf{G}_0(z)]^n$  we obtain a Laurent expansion whose coefficient of  $(z - E_a^0)^{-1}$  is  $\mathbf{A}^{(n)}$ . The reader may check that such a procedure yields

$$\mathbf{A}^{(n)} = (-1)^{n+1} \sum_{(n)} \mathbf{s}^{k_1} \mathbf{v} \mathbf{s}^{k_2} \mathbf{v} \dots \mathbf{v} \mathbf{s}^{k_{n+1}}, \tag{21.51}$$

where by definition,  $\sum_{(p)}$  extends over all nonnegative integers  $\{k_i\}_{i=1}^{n+1}$  such that

$$\sum_{i=1}^{n+1} k_i = p \quad \forall p \geq 0.$$

It turns out that for perturbation expansion, not only do we need the expansion of  $\mathbf{P}$  [Equations (21.49) and (21.51)], but also an expansion for  $\mathbf{H}\mathbf{P}$ . Using Equations (21.43) and (21.44), with  $\Gamma$  replaced by  $\Gamma_a$ , we have

$$\begin{aligned} \mathbf{H}\mathbf{P} &= \frac{1}{2\pi i} \oint_{\Gamma_a} z \mathbf{G}(z) dz = \frac{1}{2\pi i} \oint_{\Gamma_a} (z - E_a^0 + E_a^0) \mathbf{G}(z) dz \\ &= \frac{1}{2\pi i} \oint_{\Gamma_a} (z - E_a^0) \mathbf{G}(z) dz + E_a^0 \mathbf{P}. \end{aligned}$$

Substituting for  $\mathbf{G}(z)$  from Equation (21.46), we can rewrite this equation as

$$(\mathbf{H} - E_a^0) \mathbf{P} = \sum_{n=1}^{\infty} \lambda^n \mathbf{B}^{(n)}, \tag{21.52}$$

where

$$\mathbf{B}^{(n)} = (-1)^{n-1} \sum_{(n-1)} \mathbf{s}^{k_1} \mathbf{v} \mathbf{s}^{k_2} \mathbf{v} \dots \mathbf{v} \mathbf{s}^{k_{n+1}}. \tag{21.53}$$

Equations (21.52) and (21.53) can be used to approximate the eigenvectors and eigenvalues of the perturbed Hamiltonian in terms of those of the unperturbed Hamiltonian. It is convenient to consider two cases: the nondegenerate case in which  $m_a = 1$ , and the degenerate case in which  $m_a \geq 2$ .

### 21.5.1 The Nondegenerate Case

In the nondegenerate case, we let  $|a^0\rangle$  denote the original unperturbed eigenstate, and use Equation (21.47) to conclude that the perturbed eigenstate is also one-dimensional. In fact, it follows from (21.40) that  $\mathbf{P}|a^0\rangle$  is the desired eigenstate. Denoting the latter by  $|\psi\rangle$  and using Equation (21.49), we have

$$|\psi\rangle = \mathbf{P}|a^0\rangle = \mathbf{P}_a^0 |a^0\rangle + \sum_{n=1}^{\infty} \lambda^n \mathbf{A}^{(n)} |a^0\rangle = |a^0\rangle + \sum_{n=1}^{\infty} \lambda^n \mathbf{A}^{(n)} |a^0\rangle \tag{21.54}$$

because  $\mathbf{P}_a^0$  is the projection operator onto  $|a^0\rangle$ .

More desirable is the energy of the perturbed state  $E_a$ , which obeys the relation  $\mathbf{HP} = E_a\mathbf{P}$ . Taking the trace of this relation and noting that  $\text{tr } \mathbf{P} = \text{tr } \mathbf{P}_a^0 = 1$ , we obtain

$$\begin{aligned} E_a &= \text{tr}(\mathbf{HP}) = \text{tr} \left( E_a^0 \mathbf{P} + \sum_{n=1}^{\infty} \lambda^n \mathbf{B}^{(n)} \right) \\ &= E_a^0 + \sum_{n=1}^{\infty} \lambda^n \underbrace{\text{tr } \mathbf{B}^{(n)}}_{\equiv \varepsilon_n} = E_a^0 + \sum_{n=1}^{\infty} \lambda^n \varepsilon_n, \end{aligned} \tag{21.55}$$

where we used Equation (21.52). Since  $\lambda$  is simply a parameter to keep track of the order of perturbation, one usually includes it in the definition of the perturbing potential  $\mathbf{V}$ . The  $n$ th-order correction to the energy is then written as

$$\varepsilon_n = \text{tr } \mathbf{B}^{(n)}. \tag{21.56}$$

Since each term of  $\mathbf{B}^{(n)}$  contains  $\mathbf{P}_a^0$  at least once, and since

$$\text{tr}(\mathbf{UP}_a^0\mathbf{T}) = \text{tr}(\mathbf{TUP}_a^0)$$

for any pair of operators  $\mathbf{U}$  and  $\mathbf{T}$  (or products thereof), one can cast  $\varepsilon_n$  into the form of an expectation value of some product of operators in the unperturbed state  $|a^0\rangle$ . For example,

first-order correction to energy

$$\varepsilon_1 = \text{tr } \mathbf{B}^{(1)} = \sum_b \langle a^0 | \mathbf{P}_a^0 \mathbf{V} \mathbf{P}_a^0 | b^0 \rangle = \langle a^0 | \mathbf{V} | a^0 \rangle \tag{21.57}$$

because  $\mathbf{P}_a^0 | b^0 \rangle = 0$  unless  $b = a$ . This is the familiar expression for the **first order correction** to the energy in nondegenerate perturbation theory. Similarly,

$$\begin{aligned} \varepsilon_2 &= \text{tr } \mathbf{B}^{(2)} = -\text{tr}(\mathbf{P}_a^0 \mathbf{V} \mathbf{P}_a^0 \mathbf{V} [-\mathbf{Q}_a^0 \mathbf{G}_0^k(E_a^0) \mathbf{Q}_a^0] \\ &\quad + \mathbf{P}_a^0 \mathbf{V} [-\mathbf{Q}_a^0 \mathbf{G}_0^k(E_a^0) \mathbf{Q}_a^0] \mathbf{V} \mathbf{P}_a^0 + [-\mathbf{Q}_a^0 \mathbf{G}_0^k(E_a^0) \mathbf{Q}_a^0] \mathbf{V} \mathbf{P}_a^0 \mathbf{V} \mathbf{P}_a^0) \\ &= \langle a^0 | \mathbf{V} \mathbf{Q}_a^0 \mathbf{G}_0^k(E_a^0) \mathbf{Q}_a^0 \mathbf{V} | a^0 \rangle. \end{aligned}$$

The first and the last terms in parentheses give zero because in the trace sum,  $\mathbf{P}_a^0$  gives a nonzero contribution only if the state is  $|a^0\rangle$ , which is precisely the state annihilated by  $\mathbf{Q}_a^0$ . Using the completeness relation  $\sum_b |b^0\rangle \langle b^0| = \mathbf{1} = \sum_c |c^0\rangle \langle c^0|$  for the eigenstates of the unperturbed Hamiltonian, we can rewrite  $\varepsilon_2$  as

second-order correction to energy

$$\varepsilon_2 = \sum_{b,c} \langle a^0 | \mathbf{V} | b^0 \rangle \underbrace{\langle b^0 | \mathbf{Q}_a^0 \mathbf{G}_0^k(E_a^0) \mathbf{Q}_a^0 | c^0 \rangle}_{\delta_{bc}/(E_a^0 - E_b^0)} \langle c^0 | \mathbf{V} | a^0 \rangle = \sum_{b \neq a} \frac{|\langle a^0 | \mathbf{V} | b^0 \rangle|^2}{E_a^0 - E_b^0}.$$

This is the familiar expression for the **second-order correction** to the energy in nondegenerate perturbation theory.

### 21.5.2 The Degenerate Case

The degenerate case can also start with Equations (21.54) and (21.55). The difference is that  $\varepsilon_n$  cannot be determined as conveniently as the nondegenerate case. For example, the expression for  $\varepsilon_1$  will involve a sum over a basis of  $\mathcal{M}_a^0$  because  $\mathbf{P}_a^0 |b\rangle$  is no longer just  $|a\rangle$ , but some general vector in  $\mathcal{M}_a^0$ . Instead of pursuing this line of approach, we present a more common method, which concentrates on the way  $\mathcal{M}_a^0$  and the corresponding eigenspaces of the perturbed Hamiltonian, denoted by  $\mathcal{M}_a$ , enter in the calculation of eigenvalues and eigenvectors.

The projector  $\mathbf{P}_a^0$  acts as a unit operator when restricted to  $\mathcal{M}_a^0$ . In particular, it is invertible. In the limit of small  $\lambda$ , the projection operator  $\mathbf{P}$  is close to  $\mathbf{P}_a^0$ ; therefore, it too must be invertible, i.e.,  $\mathbf{P} : \mathcal{M}_a^0 \rightarrow \mathcal{M}_a$  is an isomorphism. Similarly,  $\mathbf{P}_a^0 : \mathcal{M}_a \rightarrow \mathcal{M}_a^0$  is also an isomorphism—not necessarily the inverse of the first one. It follows that for each vector in  $\mathcal{M}_a^0$  there is a unique vector in  $\mathcal{M}_a$  and vice versa.

The eigenvalue equation  $\mathbf{H} |E_a\rangle = E_a |E_a\rangle$  can thus be written as

$$\mathbf{H}\mathbf{P}_a |E_a^0\rangle = E_a \mathbf{P}_a |E_a^0\rangle,$$

where  $|E_a^0\rangle$  is the unique vector mapped onto  $|E_a\rangle$  by  $\mathbf{P}_a$ . Multiplying both sides by  $\mathbf{P}_a^0$ , we obtain

$$\mathbf{P}_a^0 \mathbf{H} \mathbf{P}_a |E_a^0\rangle = E_a \mathbf{P}_a^0 \mathbf{P}_a |E_a^0\rangle,$$

which is completely equivalent to the previous equation because  $\mathbf{P}_a^0$  is invertible. If we define

$$\mathbf{H}_a \equiv \mathbf{P}_a^0 \mathbf{H} \mathbf{P}_a : \mathcal{M}_a^0 \rightarrow \mathcal{M}_a^0, \quad \mathbf{K}_a \equiv \mathbf{P}_a^0 \mathbf{P}_a : \mathcal{M}_a \rightarrow \mathcal{M}_a^0, \quad (21.58)$$

the preceding equation becomes

$$\mathbf{H}_a |E_a^0\rangle = E_a \mathbf{K}_a |E_a^0\rangle. \quad (21.59)$$

As operators on  $\mathcal{M}_a^0$  both  $\mathbf{H}_a$  and  $\mathbf{K}_a$  are hermitian. In fact,  $\mathbf{K}_a$ , which can be written as the product of  $\mathbf{P}_a^0 \mathbf{P}_a$  and its hermitian conjugate, is a positive definite operator. Equation (21.59) is a generalized eigenvalue equation whose eigenvalues  $E_a$  are solutions of the equation

$$\det(\mathbf{H}_a - x \mathbf{K}_a) = 0. \quad (21.60)$$

The eigenvectors of this equation, once projected onto  $\mathcal{M}_a$  by  $\mathbf{P}_a$ , give the desired eigenvectors of  $\mathbf{H}$ .

The expansions of  $\mathbf{H}_a$  and  $\mathbf{K}_a$  are readily obtained from those of  $\mathbf{H}\mathbf{P}_a$  and  $\mathbf{P}_a$  as given in Equations (21.49) and (21.52). We give the first few terms of each expansion:

$$\begin{aligned} \mathbf{K}_a &= \mathbf{P}_a^0 - \lambda^2 \mathbf{P}_a^0 \mathbf{V} \mathbf{Q}_a^0 \mathbf{G}_0^2(E_a^0) \mathbf{Q}_a^0 \mathbf{V} \mathbf{P}_a^0 + \dots, \\ \mathbf{H}_a &= E_a^0 \mathbf{K}_a + \lambda \mathbf{P}_a^0 \mathbf{V} \mathbf{P}_a^0 + \lambda^2 \mathbf{P}_a^0 \mathbf{V} \mathbf{Q}_a^0 \mathbf{G}_0(E_a^0) \mathbf{Q}_a^0 \mathbf{V} \mathbf{P}_a^0 + \dots \end{aligned} \quad (21.61)$$

To any given order of approximation, the eigenvalues  $E_a$  are obtained by terminating the series in (21.61) at that order, plugging the resulting finite sum in Equation (21.60), and solving the determinant equation.

### 21.6 Problems

**21.1.** Show that the definitions of the three types of SOPDEs discussed in Example 21.1.8 are equivalent to the definitions based on Equation (21.5). Hint: Diagonalize the matrix of coefficients of the SOPDE:

$$a \frac{\partial^2 u}{\partial x^2} + 2b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} + F \left( x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right) = 0,$$

where  $a, b,$  and  $c$  are functions of  $x$  and  $y$ . Write the eigenvalues as  $(a + c \pm \Delta)/2$  and consider the three cases  $|\Delta| < |a + c|, |\Delta| > |a + c|,$  and  $|\Delta| = |a + c|.$

**21.2.** Find the characteristic curves for  $L_x[u] = \partial u / \partial x.$

**21.3.** Find the characteristic curves for the two-dimensional wave equation and the two-dimensional diffusion equation.

**21.4.** Solve the Cauchy problem for the two-dimensional Laplace equation subject to the Cauchy data  $u(0, y) = 0, (\partial u / \partial x)(0, y) = \epsilon \sin ky,$  where  $\epsilon$  and  $k$  are constants. Show that the solution does not vary continuously as the Cauchy data vary. In particular, show that for any  $\epsilon \neq 0$  and any preassigned  $x > 0,$  the solution  $u(x, y)$  can be made arbitrarily large by choosing  $k$  large enough.

**21.5.** Show that the  $x_i$  in Equation (21.12) describe an  $m$ -dimensional sphere of radius  $r,$  that is,  $\sum_{i=1}^m x_i^2 = r^2.$

**21.6.** Use  $J\delta(\mathbf{x} - \mathbf{a}) = \delta(\xi - \alpha)$  and the coordinate transformation from the spherical coordinate system to Cartesian coordinates to express the 3D Cartesian delta function in terms of the corresponding spherical delta function at a point  $P = (x_0, y_0, z_0) = (r_0, \theta_0, \varphi_0)$  where the Jacobian  $J$  is nonvanishing.

**21.7.** Find the volume of an  $m$ -dimensional sphere.

**21.8.** Prove Equation (21.11). First, note that the RHS of Equation (21.10) is a function of only  $k$  of the  $\alpha$ 's. This means that

$$H(\xi)|_{\xi=\alpha} = H(\alpha_1, \dots, \alpha_k).$$

(a) Rewrite Equation (21.10) by separating the integral into two parts, one involving  $\{\xi_i\}_{i=1}^k$  and the other involving  $\{\xi_i\}_{i=k+1}^m.$  Compare the RHS with the LHS and show that

$$\int J d\xi_{k+1} \cdots d\xi_m \delta(\mathbf{x} - \mathbf{a}) = \prod_{i=1}^k \delta(\xi_i - \alpha_i).$$

(b) Show that this equation implies that  $\delta(\mathbf{x} - \mathbf{a})$  is independent of  $\{\xi_i\}_{i=k+1}^m$ . Thus, one can take the delta function out of the integral.

**21.9.** Find the  $m$ -dimensional Green's function for the Laplacian as follows.

(a) Solve Equation (21.19) assuming that  $r \neq 0$  and demanding that  $G(r) \rightarrow 0$  as  $r \rightarrow \infty$  (this can be done only for  $m \geq 3$ ).

(b) Use the divergence theorem in  $m$  dimensions and (21.18) to show that

$$\iint_S \frac{dG}{dr} da = 1,$$

where  $S$  is a spherical hypersurface of radius  $r$ . Now use this and the result of part (a) to find the remaining constant of integration.

**21.10.** Consider the operator  $\mathbf{L}_x = \nabla^2 + \mathbf{b} \cdot \nabla + c$  for which  $\{b_i\}_{i=1}^m$  and  $c$  are functions of  $\{x_i\}_{i=1}^m$ .

(a) Show that  $\mathbf{L}_x^\dagger[v] = \nabla^2 v - \nabla \cdot (\mathbf{b}v) + cv$ , and

$$\mathbf{Q}[u, v^*] = \mathbf{Q}[u, v] = v \nabla u - u \nabla v + \mathbf{b}uv.$$

(b) Show that a necessary condition for  $\mathbf{L}_x$  to be self-adjoint is  $2\mathbf{b} \cdot \nabla u + u(\nabla \cdot \mathbf{b}) = 0$  for arbitrary  $u$ .

(c) By choosing some  $u$ 's judiciously, show that (b) implies that  $b_i = 0$ . Conclude that  $\mathbf{L}_x = \nabla^2 + c(\mathbf{x})$  is formally self-adjoint.

**21.11.** Solve the integral form of the Schrödinger equation for an attractive double delta-function potential

$$V(x) = -V_0[\delta(x - a_1) + \delta(x - a_2)], \quad V_0 > 0.$$

Find the eigenfunctions and obtain a transcendental equation for the eigenvalues (see Example 21.4.1).

**21.12.** Show that the integral equation associated with the damped harmonic oscillator DE  $\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0$ , having the BCs  $x(0) = x_0$ ,  $(dx/dt)_{t=0} = 0$ , can be written in either of the following forms.

$$(a) \quad x(t) = x_0 - \frac{\omega_0^2}{2\gamma} \int_0^t [1 - e^{-2\gamma(t-t')}] x(t') dt'.$$

$$(b) \quad x(t) = x_0 \cos \omega_0 t + \frac{2\gamma x_0}{\omega_0} \sin \omega_0 t - 2\gamma \int_0^t \cos[\omega_0(t-t')] x(t') dt'.$$

Hint: Take  $\omega_0^2 x$  or  $2\gamma\dot{x}$ , respectively, as the inhomogeneous term.

**21.13.** (a) Show that for scattering problems ( $E > 0$ ) the integral form of the Schrödinger equation in one dimension is

$$\Psi(x) = e^{ikx} - \frac{i\mu}{\hbar^2 k} \int_{-\infty}^{\infty} e^{ik|x-y|} V(y) \Psi(y) dy.$$

(b) Divide  $(-\infty, +\infty)$  into three regions  $R_1 = (-\infty, -a)$ ,  $R_2 = (-a, +a)$ , and  $R_3 = (a, \infty)$ . Let  $\psi_i(x)$  be  $\psi(x)$  in region  $R_i$ . Assume that the potential  $V(x)$  vanishes in  $R_1$  and  $R_3$ . Show that

$$\psi_1(x) = e^{ikx} - \frac{i\mu}{\hbar^2 k} e^{-ikx} \int_{-a}^a e^{iky} V(y) \psi_2(y) dy,$$

$$\psi_2(x) = e^{ikx} - \frac{i\mu}{\hbar^2 k} \int_{-a}^a e^{ik|x-y|} V(y) \psi_2(y) dy,$$

$$\psi_3(x) = e^{ikx} - \frac{i\mu}{\hbar^2 k} e^{ikx} \int_{-a}^a e^{-iky} V(y) \psi_2(y) dy.$$

This shows that determining the wave function in regions where there is no potential requires the wave function in the region where the potential acts.

(c) Let

$$V(x) = \begin{cases} V_0 & \text{if } |x| < a, \\ 0 & \text{if } |x| > a, \end{cases}$$

and find  $\psi_2(x)$  by the method of successive approximations. Show that the  $n$ th term is less than  $(2\mu V_0 a / \hbar^2 k)^{n-1}$  (so the Neumann series will converge) if  $(2V_0 a / \hbar v) < 1$ , where  $v$  is the velocity and  $\mu v = \hbar k$  is the momentum of the wave. Therefore, for large velocities, the Neumann series expansion is valid.

21.14. (a) Show that  $\mathbf{H}\mathbf{R}_z(\mathbf{H}) = 1 + z\mathbf{R}_z(\mathbf{H})$ . (b) Use (a) to prove Equation (21.43).

## Additional Reading

1. Folland, G. *Introduction to Partial Differential Equations*, 2nd ed., Princeton University Press, 1995. Discusses multidimensional Green's functions for various differential operators of mathematical physics.
2. Messiah, A. *Quantum Mechanics*, volume II, Wiley, 1966. A thorough treatment of perturbation theory in the style of this chapter.
3. Stakgold, I. *Green's Functions and Boundary Value Problems*, Wiley, 1979. A detailed analysis of boundary value problems in two and three dimensions.

# Multidimensional Green's Functions: Applications

The previous chapter gathered together some general properties of the GFs and their companion, the Dirac delta function. This chapter considers the Green's functions for elliptic, parabolic, and hyperbolic equations that satisfy the BCs appropriate for each type of PDE.

## 22.1 Elliptic Equations

The most general linear PDE in  $m$  variables of the elliptic type was discussed in Section 21.1.2. We will not discuss this general case, because all elliptic PDOs encountered in mathematical physics are of a much simpler nature. In fact, the self-adjoint elliptic PDO of the form  $\mathbf{L}_x = \nabla^2 + q(\mathbf{x})$  is sufficiently general for purposes of this discussion. Recall from Section 21.1.2 that the BCs associated with an elliptic PDE are of two types, Dirichlet and Neumann. Let us consider these separately.

### 22.1.1 The Dirichlet Boundary Value Problem

A Dirichlet BVP consists of an elliptic PDE together with a Dirichlet BC, such as

$$\begin{aligned}\mathbf{L}_x[u] &= \nabla^2 u + q(\mathbf{x})u = f(\mathbf{x}) && \text{for } \mathbf{x} \in D, \\ u(\mathbf{x}_b) &= g(\mathbf{x}_b) && \text{for } \mathbf{x}_b \in \partial D,\end{aligned}\tag{22.1}$$

where  $g(\mathbf{x}_b)$  is a given function defined on the closed hypersurface  $\partial D$ .

The Green's function for the Dirichlet BVP must satisfy the *homogeneous* BC, for the same reason as in the one-dimensional Green's function. Thus, the Dirichlet

Green's function, denoted by  $G_D(\mathbf{x}, \mathbf{y})$ , must satisfy

$$\mathbf{L}_x[G_D(\mathbf{x}, \mathbf{y})] = \delta(\mathbf{x} - \mathbf{y}), \quad G_D(\mathbf{x}_b, \mathbf{y}) = 0 \quad \text{for } \mathbf{x}_b \in S.$$

As discussed in Section 21.3.2, we can separate  $G_D$  into a singular part  $G_D^{(s)}$  and a regular part  $H$  where  $G_D^{(s)}$  satisfies the same DE as  $G_D$  and  $H$  satisfies the corresponding homogeneous DE and the BC  $H(\mathbf{x}_b, \mathbf{y}) = -G_D^{(s)}(\mathbf{x}_b, \mathbf{y})$ .

Using Equation (22.1) and the properties of  $G_D(\mathbf{x}, \mathbf{y})$  in Equation (21.27), we obtain

$$u(\mathbf{x}) = \int_D d^m y G_D(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) + \int_{\partial D} g(\mathbf{y}_b) \frac{\partial G_D}{\partial n_y}(\mathbf{x}, \mathbf{y}_b) da, \quad (22.2)$$

where  $\partial/\partial n_y$  indicates normal differentiation with respect to the second argument.

**Gustav Peter Lejeune Dirichlet** (1805–1859), the son of a postmaster, first attended public school, then a private school that emphasized Latin. He was precociously interested in mathematics; it is said that before the age of twelve he used his pocket money to buy mathematical books. In 1817 he entered the gymnasium in Bonn. He is reported to have been an unusually attentive and well-behaved pupil who was particularly interested in modern history as well as in mathematics.



After two years in Bonn, Dirichlet was sent to a Jesuit college in Cologne that his parents preferred. Among his teachers was the physicist Georg Simon Ohm, who gave him a thorough grounding in theoretical physics. Dirichlet completed his *Abitur* examination at the very early age of sixteen. His parents wanted him to study law, but mathematics was already his chosen field. At the time the level of pure mathematics in the German universities was at a low ebb: Except for the formidable Carl Gauss, in Göttingen, there were no outstanding mathematicians, while in Paris the firmament was studied by such luminaries as P.-S. Laplace, Adrien Legendre, Joseph Fourier, and Siméon Poisson.

Dirichlet arrived in Paris in May 1822. In the summer of 1823 he was fortunate in being appointed to a well-paid and pleasant position as tutor to the children of General Maximilien Fay, a national hero of the Napoleonic wars and then the liberal leader of the opposition in the Chamber of Deputies. Dirichlet was treated as a member of the family and met many of the most prominent figures in French intellectual life. Among the mathematicians, he was particularly attracted to Fourier, whose ideas had a strong influence upon his later works on trigonometric series and mathematical physics.

General Fay died in November 1825, and the next year Dirichlet decided to return to Germany, a plan strongly supported by Alexander von Humboldt, who was working for the strengthening of the natural sciences in Germany. Dirichlet was permitted to qualify for habilitation as *Privatdozent* at the University of Breslau; since he did not have the required doctorate, this was awarded *honoris causa* by the University of Cologne. His habilitation thesis dealt with polynomials whose prime divisors belong to special arithmetic series. A second paper from this period was inspired by Gauss's announcements on the biquadratic law of reciprocity.



Dirichlet was appointed extraordinary professor in Breslau, but the conditions for scientific work were not inspiring. In 1828 he moved to Berlin, again with the assistance of Humboldt, to become a teacher of mathematics at the military academy. Shortly afterward, at the age of twenty-three, he was appointed extraordinary (later ordinary) professor at the University of Berlin. In 1831 he became a member of the Berlin Academy of Sciences, and in the same year he married Rebecca Mendelssohn-Bartholdy, sister of Felix Mendelssohn, the composer.

Dirichlet spent twenty-seven years as a professor in Berlin and exerted a strong influence on the development of German mathematics through his lectures, through his many pupils, and through a series of scientific papers of the highest quality that he published during this period. He was an excellent teacher, always expressing himself with great clarity. His manner was modest; in his later years he was shy and at times reserved. He seldom spoke at meetings and was reluctant to make public appearances. In many ways he was a direct contrast to his lifelong friend, the mathematician Karl Gustav Jacobi.

One of Dirichlet's most important papers, published in 1850, deals with the boundary value problem, now known as *Dirichlet's boundary value problem*, in which one wishes to determine a potential function satisfying Laplace's equation and having prescribed values on a given surface, in Dirichlet's case a sphere.

In 1855, when Gauss died, the University of Göttingen was anxious to seek a successor of great distinction, and the choice fell upon Dirichlet. Dirichlet moved to Göttingen in the fall of 1855, bought a house with a garden, and seemed to enjoy the quieter life of a prominent university in a small city. He had a number of excellent pupils and relished the increased leisure for research. His work in this period was centered on general problems of mechanics. This new life, however, was not to last long. In the summer of 1858 Dirichlet traveled to a meeting in Montreux, Switzerland, to deliver a memorial speech in honor of Gauss. While there, he suffered a heart attack and was barely able to return to his family in Göttingen. During his illness his wife died of a stroke, and Dirichlet himself died the following spring.

Some special cases of (22.2) are worthy of mention. The first is  $u(\mathbf{x}_b) = 0$ , the solution to an inhomogeneous DE satisfying the homogeneous BC. We obtain this by substituting zero for  $g(\mathbf{x}_b)$  in (22.2) so that only the integration over  $D$  remains. The second special case is when the DE is homogeneous, that is, when  $f(\mathbf{x}) = 0$  but the BC is inhomogeneous. This yields an integration over the boundary  $\partial D$  alone. Finally, the solution to the homogeneous DE with the homogeneous BC is simply  $u = 0$ , referred to as the zero solution. This is consistent with physical intuition: If the function is zero on the boundary and there is no source  $f(\mathbf{x})$  to produce any "disturbance," we expect no nontrivial solution.

**22.1.1. Example.** Let us find the Green's function for the three-dimensional Laplacian  $\mathbf{L}_x = \nabla^2$  satisfying the Dirichlet BC  $G_D(\rho, y) = 0$  for  $\rho$ , on the  $xy$ -plane. Here  $D$  is the upper half-space ( $z \geq 0$ ) and  $\partial D$  is the  $xy$ -plane.

It is more convenient to use  $\mathbf{r} = (x, y, z)$  and  $\mathbf{r}' = (x', y', z')$  instead of  $\mathbf{x}$  and  $y$ , respectively. Using (21.21) as  $G_D^{(s)}$ , we can write

$$G_D(\mathbf{r}, \mathbf{r}') = -\frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|} + H(\mathbf{r}, \mathbf{r}')$$

$$= -\frac{1}{4\pi} \frac{1}{\sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}} + H(x, y, z; x', y', z').$$

The requirement that  $G_D$  vanish on the  $xy$ -plane gives

$$H(x, y, 0; x', y', z') = \frac{1}{4\pi} \frac{1}{\sqrt{(x-x')^2 + (y-y')^2 + z'^2}}.$$

This fixes the dependence of  $H$  on all variables except  $z$ . On the other hand,  $\nabla^2 H = 0$  in  $D$  implies that the form of  $H$  must be the same as that of  $G_D^{(s)}$  because except at  $\mathbf{r} = \mathbf{r}'$ , the latter does satisfy Laplace's equation. Thus, because of the symmetry of  $G_D^{(s)}$  in  $\mathbf{r}$  and  $\mathbf{r}'$  [ $G_D(\mathbf{r}, \mathbf{r}') = G_D(\mathbf{r}', \mathbf{r})$ ] and the evenness of the Laplacian in  $z$  (as well as  $x$  and  $y$ ), we have two choices for the  $z$ -dependence:  $(z-z')^2$  and  $(z+z')^2$ . The first gives  $G_D = 0$ , which is a trivial solution. Thus, we must choose

$$H(x, y, z; x', y', z') = \frac{1}{4\pi} \frac{1}{\sqrt{(x-x')^2 + (y-y')^2 + (z+z')^2}}.$$

Note that with  $\mathbf{r}'' \equiv (x', y', -z')$ , this equation satisfies  $\nabla^2 H = -\delta(\mathbf{r} - \mathbf{r}'')$ , and it may appear that  $H$  does not satisfy the homogeneous DE, as it should. However,  $\mathbf{r}''$  is outside  $D$ , and  $\mathbf{r} \neq \mathbf{r}''$  as long as  $\mathbf{r} \in D$ . So  $H$  does satisfy the homogeneous DE in  $D$ . The Green's function for the given Dirichlet BC is therefore

$$G_D(\mathbf{r}, \mathbf{r}') = -\frac{1}{4\pi} \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} - \frac{1}{|\mathbf{r} - \mathbf{r}''|} \right),$$

where  $\mathbf{r}''$  is the *reflection* of  $\mathbf{r}'$  in the  $xy$ -plane.

This result has a direct physical interpretation. If determining the solution of the Laplace equation is considered a problem in electrostatics, then  $G_D^{(s)}(\mathbf{r}, \mathbf{r}')$  is simply the potential at  $\mathbf{r}$  of a unit point charge located at  $\mathbf{r}'$ , and  $G_D(\mathbf{r}, \mathbf{r}')$  is the potential of two point charges of opposite signs, one at  $\mathbf{r}'$  and the other at the mirror image of  $\mathbf{r}'$ . The fact that the two charges are equidistant from the  $xy$ -plane ensures the vanishing of the potential in that plane. The introduction of image charges to ensure the vanishing of  $G_D$  at  $\partial D$  is common in electrostatics and is known as the **method of images**. This method reduces the Dirichlet problem for the Laplacian to finding appropriate point charges outside  $D$  that guarantee the vanishing of the potential on  $\partial D$ . For simple geometries, such as the one discussed in this example, determination of the magnitudes and locations of such image charges is easy, rendering the method extremely useful.

Having found the Green's function, we can pose the general Dirichlet BVP:  $\nabla^2 u = -\rho(\mathbf{r})$  and  $u(x, y, 0) = g(x, y)$  for  $z > 0$ . The solution is

$$\begin{aligned} u(\mathbf{r}) = & \frac{1}{4\pi} \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dy' \int_0^{\infty} dz' \rho(\mathbf{r}') \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} - \frac{1}{|\mathbf{r} - \mathbf{r}''|} \right) \\ & + \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dy' g(x', y') \left. \frac{\partial G_D}{\partial z} \right|_{z=0}, \end{aligned} \quad (22.3)$$

where  $\mathbf{r} = (x, y, z)$ ,  $\mathbf{r}' = (x', y', z')$ , and  $\mathbf{r}'' = (x', y', -z')$ .

A typical application consists in introducing a number of charges in the vicinity of an infinite conducting sheet, which is held at a constant potential  $V_0$ . If there are  $N$  charges,

$\{q_i\}_{i=1}^N$ , located at  $\{\mathbf{r}_i\}_{i=1}^N$ , then  $\rho(\mathbf{r}) = \sum_{i=1}^N q_i \delta(\mathbf{r} - \mathbf{r}_i)$ ,  $g(x, y) = \text{const} = V_0$ , and we get

$$u(\mathbf{r}) = \sum_{i=1}^N \frac{1}{4\pi} \left( \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|} - \frac{q_i}{|\mathbf{r} - \mathbf{r}'_i|} \right) + V_0 \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dy' \frac{\partial G_D}{\partial z} \Big|_{z=0}, \quad (22.4)$$

where  $\mathbf{r}_i = (x_i, y_i, z_i)$  and  $\mathbf{r}'_i = (x_i, y_i, -z_i)$ . That the double integral in Equation (22.4) is unity can be seen by direct integration or by noting that the sum vanishes when  $z = 0$ . On the other hand,  $u(x, y, 0) = V_0$ . Thus, the solution becomes

$$u(\mathbf{r}) = \sum_{i=1}^N \frac{1}{4\pi} \left( \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|} - \frac{q_i}{|\mathbf{r} - \mathbf{r}'_i|} \right) + V_0. \quad \blacksquare$$

**22.1.2. Example.** The method of images is also applicable when the boundary is a sphere. Inside a sphere of radius  $a$  with center at the origin, we wish to solve this Dirichlet BVP:  $\nabla^2 u = -\rho(r, \theta, \varphi)$  for  $r < a$ , and  $u(a, \theta, \varphi) = g(\theta, \varphi)$ . The GF satisfies

$$\begin{aligned} \nabla^2 G_D(r, \theta, \varphi; r', \theta', \varphi') &= \delta(\mathbf{r} - \mathbf{r}') \quad \text{for } r < a, \\ G_D(a, \theta, \varphi; r', \theta', \varphi') &= 0. \end{aligned} \quad (22.5)$$

Thus,  $G_D$  can again be interpreted as the potential of point charges, of which one is in the sphere and the others are outside.

We write  $G_D = G_D^{(s)} + H$  and choose  $H$  in such a way that the second equation in (22.5) is satisfied. As in the case of the  $xy$ -plane, let<sup>1</sup>  $H(\mathbf{r}, \mathbf{r}'') = -\frac{k}{4\pi|\mathbf{r} - \mathbf{r}''|}$ , where  $k$  is a constant to be determined. If  $\mathbf{r}''$  is *outside* the sphere,  $\nabla^2 H$  will vanish everywhere *inside* the sphere. The problem has been reduced to finding  $k$  and  $\mathbf{r}''$  (the location of the image charge). We want to choose  $\mathbf{r}''$  such that

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} \Big|_{r=a} = \frac{k}{|\mathbf{r} - \mathbf{r}''|} \Big|_{r=a} \Rightarrow k(|\mathbf{r} - \mathbf{r}'|)_{r=a} = (|\mathbf{r} - \mathbf{r}''|)_{r=a}.$$

This shows that  $k$  must be positive. Squaring both sides and expanding the result yields

$$k^2(a^2 + r'^2 - 2ar' \cos \gamma) = a^2 + r''^2 - 2ar'' \cos \gamma,$$

where  $\gamma$  is the angle between  $\mathbf{r}$  and  $\mathbf{r}'$ , and we have assumed that  $\mathbf{r}'$  and  $\mathbf{r}''$  are in the same direction. If this equation is to hold for arbitrary  $\gamma$ , we must have  $k^2 r' = r''$  and  $k^2(a^2 + r'^2) = a^2 + r''^2$ . Combining these two equations yields  $k^4 r'^2 - k^2(a^2 + r'^2) + a^2 = 0$ , whose positive solutions are  $k = 1$  and  $k = a/r$ . The first choice implies that  $r'' = r'$ , which is impossible because  $r''$  must be outside the sphere. We thus choose  $k = a/r'$ , which gives  $\mathbf{r}'' = (a^2/r'^2)\mathbf{r}'$ . We then have

$$G_D(\mathbf{r}, \mathbf{r}') = -\frac{1}{4\pi} \left[ \frac{1}{|\mathbf{r} - \mathbf{r}'|} - \frac{ar'}{r'^2 |\mathbf{r} - a^2 \mathbf{r}'/r'^2|} \right]. \quad (22.6)$$

<sup>1</sup>Actually, to be general, we must add an arbitrary function  $f(r'')$  to this. However, as the reader can easily verify, the following argument will show that  $f(r'') = 0$ . Besides, we are only interested in a solution, not the most general one. All simplifying assumptions that follow are made for the same reason.

Substituting this in Equation (22.2), and noting that  $\partial G/\partial n_y = (\partial G/\partial r')_{r'=a}$ , yields

$$u(\mathbf{r}) = \frac{1}{4\pi} \int_0^a r'^2 dr' \int_0^\pi \sin \theta' d\theta' \int_0^{2\pi} \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} - \frac{ar'}{r'^2 r - a^2 r'} \right) \rho(\mathbf{r}') d\varphi' \\ + \frac{a(a^2 - r^2)}{4\pi} \int_0^{2\pi} d\varphi' \int_0^\pi \sin \theta' d\theta' \frac{g(\theta', \varphi')}{|\mathbf{r} - \mathbf{a}|^3}, \quad (22.7)$$

where  $\mathbf{a} = (a, \theta', \varphi')$  is a vector from the origin to a point on the sphere. For the Laplace equation  $\rho(\mathbf{r}') = 0$ , and only the double integral in Equation (22.7) will contribute.

It can be shown that if  $g(\theta', \varphi') = \text{const} = V_0$ , then  $u(\mathbf{r}) = V_0$ . This is the familiar fact shown in electromagnetism: If the potential on a sphere is kept constant, the potential inside the sphere will be constant and equal to the potential at the surface. ■

**22.1.3. Example.** In this example we find the Dirichlet GF for a circle of radius  $a$  centered at the origin. The GF is logarithmic [see Equation (21.22)]. Therefore,  $H$  is also logarithmic, and its most general form is

$$H(\mathbf{r}, \mathbf{r}'') = -\frac{1}{2\pi} \ln(|\mathbf{r} - \mathbf{r}''|) - \frac{1}{2\pi} \ln[f(\mathbf{r}'')] = -\frac{1}{2\pi} \ln(|\mathbf{r} - \mathbf{r}''| f(\mathbf{r}'')),$$

so that

$$G_D(\mathbf{r}, \mathbf{r}') = \frac{1}{2\pi} \ln(|\mathbf{r} - \mathbf{r}'|) - \frac{1}{2\pi} \ln(|\mathbf{r} - \mathbf{r}''| f(\mathbf{r}'')) = \frac{1}{2\pi} \ln \left| \frac{\mathbf{r} - \mathbf{r}'}{(\mathbf{r} - \mathbf{r}'') f(\mathbf{r}'')} \right|$$

For  $G_D$  to vanish at all points on the circle, we must have

$$\left| \frac{\mathbf{a} - \mathbf{r}'}{(\mathbf{a} - \mathbf{r}'') f(\mathbf{r}'')} \right| = 1 \Rightarrow |\mathbf{a} - \mathbf{r}'| = |(\mathbf{r} - \mathbf{r}'') f(\mathbf{r}'')|,$$

where  $\mathbf{a}$  is a vector from origin to a point on the circle. Assuming that  $\mathbf{r}''$  and  $\mathbf{r}'$  are in the same direction, squaring both sides of the last equation and expanding the result, we obtain  $(a^2 + r''^2 - 2ar'' \cos \gamma) f^2(\mathbf{r}'') = a^2 + r'^2 - 2ar' \cos \gamma$ , where  $\gamma$  is the angle between  $\mathbf{a}$  and  $\mathbf{r}'$  (or  $\mathbf{r}''$ ). This equation must hold for arbitrary  $\gamma$ . Hence, we have  $f^2(\mathbf{r}'') r'' = r'$  and  $f^2(\mathbf{r}'')(a^2 + r''^2) = a^2 + r'^2$ . These can be solved for  $f(\mathbf{r}'')$  and  $r''$ . The result is

$$r'' = \frac{a^2}{r'^2} r', \quad f(\mathbf{r}'') = \frac{a}{r''} = \frac{r'}{a}.$$

Substituting these formulas in the expression for  $G_D$ , we obtain

$$G_D(\mathbf{r}, \mathbf{r}') = \frac{1}{2\pi} \ln(|\mathbf{r} - \mathbf{r}'|) - \frac{1}{2\pi} \ln \left( \left| \mathbf{r} - \frac{a^2}{r'^2} \mathbf{r}' \right| \frac{r'}{a} \right).$$

To write the solution to the Dirichlet BVP, we also need  $\partial G_D/\partial n = \partial G_D/\partial r'$ . Using polar coordinates, we express  $G_D$  as

$$G_D(\mathbf{r}, \mathbf{r}') = \frac{1}{4\pi} \ln \left| \frac{r^2 + r'^2 - 2rr' \cos(\theta - \theta')}{r^2 r'^2 / a^2 + a^2 - 2rr' \cos(\theta - \theta')} \right|.$$

Differentiation with respect to  $r'$  yields

$$\frac{\partial G_D}{\partial n} \Big|_{r'=a} = \frac{\partial G_D}{\partial r'} \Big|_{r'=a} = \frac{1}{2\pi a} \frac{a^2 - r^2}{r^2 + a^2 - 2ra \cos(\theta - \theta')},$$

from which we can immediately write the solution to the two-dimensional Dirichlet BVP  $\nabla^2 u = \rho, u(r = a) = g(\theta')$  as

$$u(\mathbf{r}) = \int_0^{2\pi} d\theta' \int_0^a r' G_D(\mathbf{r}, \mathbf{r}') \rho(r') dr' + \frac{a^2 - r^2}{2\pi a} \int_0^{2\pi} d\theta' \frac{g(\theta')}{r^2 + a^2 - 2ra \cos(\theta - \theta')}.$$

**Poisson integral formula** In particular, for Laplace's equation  $\rho(r') = 0$ , and we get

$$u(r, \theta) = \frac{a^2 - r^2}{2\pi a} \int_0^{2\pi} d\theta' \frac{g(\theta')}{r^2 + a^2 - 2ra \cos(\theta - \theta')}. \tag{22.8}$$

Equation (22.8) is called the **Poisson integral formula**. ■

### 22.1.2 The Neumann Boundary Value Problem

The Neumann BVP is not as simple as the Dirichlet BVP because it requires the normal derivative of the solution. But the normal derivative is related to the Laplacian through the divergence theorem. Thus, the BC and the DE are tied together, and unless we impose some solvability conditions, we may have no solution at all. These points are illustrated clearly if we consider the Laplacian operator.

Consider the Neumann BVP

$$\nabla^2 u = f(\mathbf{x}) \quad \text{for } \mathbf{x} \in D, \quad \text{and} \quad \frac{\partial u}{\partial n} = g(\mathbf{x}) \quad \text{for } \mathbf{x} \in \partial D.$$

Integrating the first equation over  $D$  and using the divergence theorem, we obtain

$$\int_D f(\mathbf{x}) d^m x = \int_D \nabla \cdot (\nabla u) d^m x = \int_{\partial D} \hat{\mathbf{e}}_n \cdot \nabla u da = \int_{\partial D} \frac{\partial u}{\partial n} da.$$

It follows that we cannot arbitrarily assign values of  $\partial u / \partial n$  on the boundary. In particular, if the BC is homogeneous, as in the case of Green's functions, the RHS is zero, and we must have  $\int_D f(\mathbf{x}) d^m x = 0$ . This relation is a restriction on the DE, and is a solvability condition, as mentioned above. To satisfy this condition, it is necessary to subtract from the inhomogeneous term its average value over the region  $D$ . Thus, if  $V_D$  is the volume of the region  $D$ , then

$$\nabla^2 u = f(\mathbf{x}) - \bar{f} \quad \text{where} \quad \bar{f} = \frac{1}{V_D} \int_D f(\mathbf{x}) d^m x$$

ensures that the Neumann BVP is solvable. In particular, the inhomogeneous term for the Green's function is not simply  $\delta(\mathbf{x} - \mathbf{y})$  but  $\delta(\mathbf{x} - \mathbf{y}) - \bar{\delta}$ , where

$$\bar{\delta} = \frac{1}{V_D} \int_D \delta(\mathbf{x} - \mathbf{y}) d^m x = \frac{1}{V_D} \quad \text{if } \mathbf{y} \in D.$$

Thus, the Green's function for the Neumann BVP,  $G_N(\mathbf{x}, \mathbf{y})$ , satisfies

$$\begin{aligned}\nabla^2 G_N(\mathbf{x}, \mathbf{y}) &= \delta(\mathbf{x} - \mathbf{y}) - \frac{1}{V_D}, \\ \frac{\partial G_N}{\partial n}(\mathbf{x}, \mathbf{y}) &= 0 \quad \text{for } \mathbf{x} \in \partial D.\end{aligned}$$

Applying Green's identity, Equation (21.27), we get

$$u(\mathbf{x}) = \int_D d^m y G_N(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) - \int_{\partial D} G_N(\mathbf{x}, \mathbf{y}) \frac{\partial u}{\partial n} da + \bar{u}, \quad (22.9)$$

where  $\bar{u} = (1/V_D) \int_D u(\mathbf{x}) d^m x$  is the average value of  $u$  in  $D$ . Equation (22.9) is valid only for the Laplacian operator, although a similar result can be obtained for a general self-adjoint SOLPDO with constant coefficients. We will not pursue that result, however, since it is of little practical use.

**Carl Gottfried Neumann** (1832–1925) was the son of Franz Ernst Neumann, a professor of physics and mineralogy at Königsberg; his mother, Luise Florentine Hagen, was a sister-in-law of the astronomer Bessel. Neumann received his primary and secondary education in Königsberg, attended the university, and formed particularly close friendships with the analyst F. J. Richelot and the geometer L. O. Hesse. After passing the examination for secondary-school teaching, he obtained his doctorate in 1855; in 1858 he qualified for lecturing in mathematics at Halle, where he became Privatdozent and, in 1863, assistant professor. In the latter year he was called to Basel, and in 1865 to Tübingen. From the autumn of 1868 until his retirement in 1911 he was at the University of Leipzig. In 1864 he married Hermine Mathilde Elise Kloss; she died in 1875.



Neumann, who led a quiet life, was a successful university teacher and a productive researcher. More than two generations of future gymnasium teachers received their basic mathematical education from him. As a researcher he was especially prominent in the field of potential theory. His investigations into *boundary value problems* resulted in pioneering achievements; in 1870 he began to develop the method of the arithmetical mean for their solution. He also coined the term "logarithmic potential." The second boundary value problem of potential theory still bears his name; a generalization of it was later provided by H. Poincaré.

Neumann was a member of the Berlin Academy and the Societies of Göttingen, Munich, and Leipzig. He performed a valuable service in founding and editing the important German mathematics periodical *Mathematische Annalen*.

Throughout the discussion so far we have assumed that  $D$  is bounded; that is, we have considered points inside  $D$  with BCs on the boundary  $\partial D$  specified. This is called an **interior BVP**. In many physical situations we are interested in points

outside  $D$ . We are then dealing with an **exterior BVP**. In dealing with such a problem, we must specify the behavior of the Green's function at infinity. In most cases, the physics of the problem dictates such behavior. For instance, for the case of an exterior Dirichlet BVP, where

$$u(\mathbf{x}) = \int_D d^m y G_D(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) + \int_{\partial D} u(\mathbf{y}_b) \frac{\partial G_D}{\partial n_y}(\mathbf{x}, \mathbf{y}_b) da$$

and it is desired that  $u(\mathbf{x}) \rightarrow 0$  as  $|\mathbf{x}| \rightarrow \infty$ , the vanishing of  $G_D(\mathbf{x}, \mathbf{y})$  at infinity guarantees that the second integral vanishes, as long as  $\partial D$  is a finite hypersurface. To guarantee the disappearance of the first integral, we must demand that  $G_D(\mathbf{x}, \mathbf{y})$  tend to zero faster than  $f(\mathbf{y})d^m y$  tends to infinity. For most cases of physical interest, the calculation of the exterior Green's functions is not conceptually different from that of the interior ones. However, the algebra may be more involved.

Later we will develop general methods for finding the Green's functions for certain partial differential operators that satisfy appropriate BCs. At this point, let us simply mention what are called mixed BCs for elliptic PDEs. A general mixed BC is of the form

$$\alpha(\mathbf{x})u(\mathbf{x}) + \beta(\mathbf{x})\frac{\partial u}{\partial n}(\mathbf{x}) = \gamma(\mathbf{x}). \quad (22.10)$$

Problem 22.6 examines the conditions that the GF must satisfy in such a case.

## 22.2 Parabolic Equations

Elliptic partial differential equations arise in static problems, where the solution is independent of time. Of the two major time-dependent equations, the wave equation and the heat (or diffusion) equation,<sup>2</sup> the latter is a parabolic PDE and the former a hyperbolic PDE. This section examines the heat equation, which is of the form  $\nabla^2 u = a^2 \partial u / \partial t$ . By changing  $t$  to  $t/a^2$ , we can write the equation as  $\mathbf{L}_{\mathbf{x},t}[u] \equiv (\partial/\partial t - \nabla^2)u(\mathbf{x}, t) = 0$ . We wish to calculate the Green's function associated with  $\mathbf{L}_{\mathbf{x},t}$  and the homogeneous BCs. Because of the time variable, we must also specify the solution at  $t = 0$ . Thus, we consider the BVP

$$\begin{aligned} \mathbf{L}_{\mathbf{x},t}[u] &\equiv \left(\frac{\partial}{\partial t} - \nabla^2\right)u(\mathbf{x}, t) = 0 && \text{for } \mathbf{x} \in D, \\ u(\mathbf{x}_b, t) = 0, \quad u(\mathbf{x}, 0) &= h(\mathbf{x}) && \text{for } \mathbf{x}_b \in \partial D, \quad \mathbf{x} \in D. \end{aligned} \quad (22.11)$$

To find a solution to (22.11), we can use a method that turns out to be useful for evaluating Green's functions in general—the method of eigenfunctions. Let

<sup>2</sup>The heat equation turns into the Schrödinger equation if  $t$  is changed to  $\sqrt{-1}t$ ; thus, the following discussion incorporates the Schrödinger equation as well.

$\{u_n\}_{n=1}^{\infty}$  be the eigenfunctions of  $\nabla^2$  with eigenvalues  $\{-\lambda_n\}_{n=1}^{\infty}$ . Let the BC be  $u_n(\mathbf{x}_b) = 0$  for  $\mathbf{x}_b \in \partial D$ . Then

$$\begin{aligned}\nabla^2 u_n(\mathbf{x}) + \lambda_n u_n(\mathbf{x}) &= 0 & \text{for } n = 1, 2, \dots, \mathbf{x} \in D, \\ u_n(\mathbf{x}_b) &= 0 & \text{for } \mathbf{x}_b \in \partial D.\end{aligned}\quad (22.12)$$

Equation (22.12) constitutes a Sturm–Liouville problem in  $m$  dimensions, which we assume to have a solution with  $\{u_n\}_{n=1}^{\infty}$  as a complete orthonormal set. We can therefore write

$$u(\mathbf{x}, t) = \sum_{n=1}^{\infty} C_n(t) u_n(\mathbf{x}). \quad (22.13)$$

This is possible because at each specific value of  $t$ ,  $u(\mathbf{x}, t)$  is a function of  $\mathbf{x}$  and therefore can be written as a linear combination of the *same set*,  $\{u_n\}_{n=1}^{\infty}$ . The coefficients  $C_n(t)$  are given by

$$C_n(t) = \int_D u(\mathbf{x}, t) u_n(\mathbf{x}) d^m x. \quad (22.14)$$

To calculate  $C_n(t)$ , we differentiate (22.14) with respect to time and use (22.11) to obtain

$$\dot{C}_n(t) \equiv \frac{dC_n}{dt} = \int_D \frac{\partial u}{\partial t}(\mathbf{x}, t) u_n(\mathbf{x}) d^m x = \int_D [\nabla^2 u(\mathbf{x}, t)] u_n(\mathbf{x}) d^m x.$$

Using Green's identity for the operator  $\nabla^2$  yields

$$\int_D [u_n \nabla^2 u - u \nabla^2 u_n] d^m x = \int_{\partial D} \left( u_n \frac{\partial u}{\partial n} - u \frac{\partial u_n}{\partial n} \right) da.$$

Since both  $u$  and  $u_n$  vanish on  $\partial D$ , the RHS is zero, and we get

$$\dot{C}_n(t) = \int_D u \nabla^2 u_n d^m x = -\lambda_n \int_D u(\mathbf{x}, t) u_n(\mathbf{x}) d^m x = -\lambda_n C_n.$$

This has the solution  $C_n(t) = C_n(0) e^{-\lambda_n t}$ , where

$$C_n(0) = \int_D u(\mathbf{y}, 0) u_n(\mathbf{y}) d^m y = \int_D h(\mathbf{y}) u_n(\mathbf{y}) d^m y,$$

so that

$$C_n(t) = e^{-\lambda_n t} \int_D h(\mathbf{y}) u_n(\mathbf{y}) d^m y.$$

Substituting this in (22.13) and switching the order of integration and summation, we get

$$u(\mathbf{x}, t) = \int_D \left[ \sum_{n=1}^{\infty} e^{-\lambda_n t} u_n(\mathbf{x}) u_n(\mathbf{y}) \right] h(\mathbf{y}) d^m y$$



and read off the GF as  $\sum_{n=1}^{\infty} e^{-\lambda_n t} u_n(\mathbf{x}) u_n(\mathbf{y}) \theta(t)$ , where we also introduced the theta function to ensure that the solution vanishes for  $t < 0$ . More generally, we have

$$G(\mathbf{x}, \mathbf{y}; t - \tau) = \sum_{n=1}^{\infty} e^{-\lambda_n(t-\tau)} u_n(\mathbf{x}) u_n(\mathbf{y}) \theta(t - \tau). \tag{22.15}$$

Note the property

$$\lim_{\tau \rightarrow t} G(\mathbf{x}, \mathbf{y}; t - \tau) = \sum_{n=1}^{\infty} u_n(\mathbf{x}) u_n(\mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}),$$

which is usually written as

$$G(\mathbf{x}, \mathbf{y}; 0^+) = \delta(\mathbf{x} - \mathbf{y}). \tag{22.16}$$

The reader may also check that

$$\mathbf{L}_{\mathbf{x},t} G(\mathbf{x}, \mathbf{y}; t - \tau) = \delta(\mathbf{x} - \mathbf{y}) \delta(t - \tau) \tag{22.17}$$

This is precisely what we expect for the Green's function of an operator in the variables  $\mathbf{x}$  and  $t$ . Another property of  $G(\mathbf{x}, \mathbf{y}; t - \tau)$  is that it vanishes on  $\partial D$ , as it should.

Having found the Green's function and noted its properties, we are in a position to solve the inhomogeneous analogue of Equation (22.11), in which the RHS of the first equation is  $f(\mathbf{x}, t)$ , and the zero on the RHS of the second equation is replaced by  $g(\mathbf{x}_b, t)$ . Experience with similar but simpler problems indicates that to make any progress toward a solution, we must come up with a form of Green's identity involving  $\mathbf{L}_{\mathbf{x},t}$  and its adjoint. It is easy to show that

$$v \mathbf{L}_{\mathbf{x},t}[u] - u \mathbf{L}_{\mathbf{x},t}^\dagger[v] = \frac{\partial}{\partial t}(uv) - \nabla \cdot (v \nabla u - u \nabla v), \tag{22.18}$$

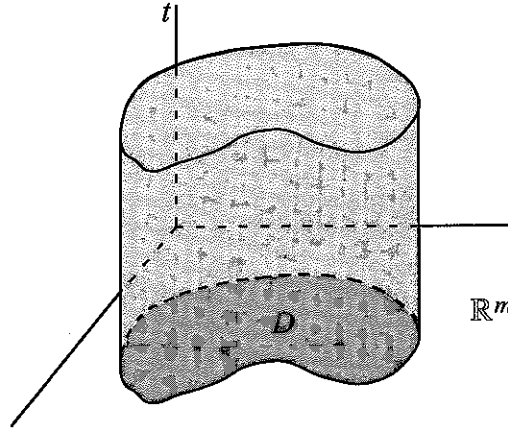
where  $\mathbf{L}_{\mathbf{x},t}^\dagger = -\partial/\partial t - \nabla^2$ .

Now consider the  $(m + 1)$ -dimensional "cylinder" one of whose bases is at  $t = \epsilon$ , where  $\epsilon$  is a small positive number. This base is barely above the  $m$ -dimensional hyperplane  $\mathbb{R}^m$ . The other base is at  $t = \tau - \epsilon$  and is a duplicate of  $D \subset \mathbb{R}^m$  (see Figure 22.1). Let  $a^\mu$ , where  $\mu = 0, 1, \dots, m$ , be the components of an  $(m + 1)$ -dimensional vector  $\mathbf{a} = (a^0, a^1, \dots, a^m)$ . Define an inner product by

$$\mathbf{a} \cdot \mathbf{b} \equiv \sum_{\mu=0}^m a^\mu b_\mu \equiv a^0 b^0 - a^1 b^1 - \dots - a^m b^m \equiv a^0 b^0 - \mathbf{a} \cdot \mathbf{b}$$

and the  $(m + 1)$ -dimensional vector  $\mathbf{Q}$  by  $Q^0 = uv$ ,  $\mathbf{Q} = v \nabla u - u \nabla v$ . Then (22.18) can be expressed as

$$v \mathbf{L}_{\mathbf{x},t}[u] - u \mathbf{L}_{\mathbf{x},t}^\dagger[v] = \sum_{\mu=0}^m \frac{\partial Q^\mu}{\partial x^\mu} \equiv \frac{\partial Q^0}{\partial x^0} - \frac{\partial Q^1}{\partial x^1} - \dots - \frac{\partial Q^m}{\partial x^m}. \tag{22.19}$$



**Figure 22.1** The “cylinder” used in evaluating the GF for the diffusion and wave equations. Note that the bases are not planes, but hyperplanes (that is, spaces such as  $\mathbb{R}^m$ ).

We recognize the RHS as a divergence in  $(m + 1)$ -dimensional space. Denoting the volume of the  $(m + 1)$ -dimensional cylinder by  $\mathcal{D}$  and its boundary by  $\partial\mathcal{D}$  and integrating (22.19) over  $\mathcal{D}$ , we obtain

$$\begin{aligned} \int_{\mathcal{D}} (v\mathbf{L}_{\mathbf{x},t}[u] - u\mathbf{L}_{\mathbf{x},t}^\dagger[v])d^{m+1}x &= \int_{\mathcal{D}} \sum_{\mu=0}^m \frac{\partial Q^\mu}{\partial x^\mu} d^{m+1}x \\ &= \int_{\partial\mathcal{D}} \sum_{\mu=0}^m Q^\mu n_\mu dS, \end{aligned} \tag{22.20}$$

where  $dS$  is an element of “area” of  $\partial\mathcal{D}$ . Note that the divergence theorem was used in the last step. The LHS is an integration over  $t$  and  $\mathbf{x}$ , which can be written as

$$\int_{\mathcal{D}} (v\mathbf{L}_{\mathbf{x},t}[u] - u\mathbf{L}_{\mathbf{x},t}^\dagger[v])d^{m+1}x = \int_\epsilon^{\tau-\epsilon} dt \int_D d^m x (v\mathbf{L}_{\mathbf{x},t}[u] - u\mathbf{L}_{\mathbf{x},t}^\dagger[v]).$$

The RHS of (22.20), on the other hand, can be split into three parts: a base at  $t = \epsilon$ , a base at  $t = \tau - \epsilon$ , and the lateral surface. The base at  $t = \epsilon$  is simply the region  $D$ , whose outward-pointing normal is in the negative  $t$  direction. Thus,  $n_0 = -1$ , and  $n_i = 0$  for  $i = 1, 2, \dots, m$ . The base at  $t = \tau - \epsilon$  is also the region  $D$ ; however, its normal is in the positive  $t$  direction. Thus,  $n_0 = 1$ , and  $n_i = 0$  for  $i = 1, 2, \dots, m$ . The element of “area” for these two bases is simply  $d^m x$ . The unit normal to the lateral surface has no time component and is simply the unit normal to the boundary of  $D$ . The element of “area” for the lateral surface is  $dt da$ , where  $da$  is an element of “area” for  $\partial D$ . Putting everything together, we

can write (22.20) as

$$\begin{aligned} & \int_{\epsilon}^{\tau-\epsilon} dt \int_D d^m x (v \mathbf{L}_{\mathbf{x},t}[u] - u \mathbf{L}_{\mathbf{x},t}^\dagger[v]) \\ &= \int_D (-Q^0)|_{t=\epsilon} d^m x + \int_D Q^0|_{t=\tau-\epsilon} d^m x - \int_{\partial D} da \int_{\epsilon}^{\tau-\epsilon} dt \mathbf{Q} \cdot \hat{\mathbf{e}}_n. \end{aligned}$$

The minus sign for the last term is due to the definition of the inner product. Substituting for  $\mathbf{Q}$  yields

$$\begin{aligned} & \int_{\epsilon}^{\tau-\epsilon} dt \int_D d^m x (v \mathbf{L}_{\mathbf{x},t}[u] - u \mathbf{L}_{\mathbf{x},t}^\dagger[v]) \\ &= - \int_D u(\mathbf{x}, \epsilon) v(\mathbf{x}, \epsilon) d^m x + \int_D u(\mathbf{x}, \tau - \epsilon) v(\mathbf{x}, \tau - \epsilon) d^m x \\ & \quad - \int_{\partial D} da \int_{\epsilon}^{\tau-\epsilon} dt \left( v \frac{\partial u}{\partial n} - u \frac{\partial v}{\partial n} \right). \end{aligned} \tag{22.21}$$

Let  $v$  be  $g(\mathbf{x}, \mathbf{y}; t - \tau)$ , the GF associated with the adjoint operator. Then Equation (22.21) gives

$$\begin{aligned} & \int_{\epsilon}^{\tau-\epsilon} dt \int_D d^m x [g(\mathbf{x}, \mathbf{y}; t - \tau) f(\mathbf{x}, t) - u(\mathbf{x}, t) \delta(\mathbf{x} - \mathbf{y}) \delta(t - \tau)] \\ &= - \int_D u(\mathbf{x}, \epsilon) g(\mathbf{x}, \mathbf{y}; \epsilon - \tau) d^m x + \int_D u(\mathbf{x}, \tau - \epsilon) g(\mathbf{x}, \mathbf{y}; -\epsilon) d^m x \\ & \quad - \int_{\partial D} da \int_{\epsilon}^{\tau-\epsilon} dt \left[ g(\mathbf{x}_b, \mathbf{y}; t - \tau) \frac{\partial u}{\partial n} - u(\mathbf{x}_b, t) \frac{\partial g}{\partial n} \right]. \end{aligned} \tag{22.22}$$

We now use the following facts:

1.  $\delta(t - \tau) = 0$  in the second integral on the LHS of Equation (22.22), because  $t$  can never be equal to  $\tau$  in the range of integration.
2. Using the symmetry property of the Green's function and the fact that  $\mathbf{L}_{\mathbf{x},t}$  is real, we have  $g(\mathbf{x}, \mathbf{y}; t - \tau) = G(\mathbf{y}, \mathbf{x}; \tau - t)$ , where we have used the fact that  $t$  and  $\tau$  are the time components of  $\mathbf{x}$  and  $\mathbf{y}$ , respectively. In particular, by (22.16),  $g(\mathbf{x}, \mathbf{y}; -\epsilon) = G(\mathbf{y}, \mathbf{x}; \epsilon) = \delta(\mathbf{x} - \mathbf{y})$ .
3. The function  $g(\mathbf{x}, \mathbf{y}; t - \tau)$  satisfies the same homogeneous BC as  $G(\mathbf{x}, \mathbf{y}; t - \tau)$ . Thus,  $g(\mathbf{x}_b, \mathbf{y}; t - \tau) = 0$  for  $\mathbf{x}_b \in \partial D$ .

Substituting all the above in (22.22), taking the limit  $\epsilon \rightarrow 0$ , and switching  $\mathbf{x}$  and  $\mathbf{y}$  and  $t$  and  $\tau$ , we obtain

$$\begin{aligned} u(\mathbf{x}, t) &= \int_0^t d\tau \int_D d^m y G(\mathbf{x}, \mathbf{y}; t - \tau) f(\mathbf{y}, \tau) + \int_D u(\mathbf{y}, 0) G(\mathbf{x}, \mathbf{y}; t) d^m y \\ & \quad - \int_0^t d\tau \int_{\partial D} u(\mathbf{y}_b, \tau) \frac{\partial G}{\partial n_y}(\mathbf{x}, \mathbf{y}_b; t - \tau) da, \end{aligned} \tag{22.23}$$

where  $\partial/\partial n_y$  in the last integral means *normal* differentiation with respect to the second argument of the Green's function.

Equation (22.23) gives the complete solution to the BVP associated with a parabolic PDE. If  $f(\mathbf{y}, \tau) = 0$  and  $u$  vanishes on the hypersurface  $\partial D$ , then Equation (22.23) gives

$$u(\mathbf{x}, t) = \int_D u(\mathbf{y}, 0)G(\mathbf{x}, \mathbf{y}; t)d^m y, \quad (22.24)$$

GF as evolution  
operator or  
propagator

which is the solution to the BVP of Equation (22.11), which led to the general Green's function of (22.15). Equation (22.24) lends itself nicely to a physical interpretation. The RHS can be thought of as an integral operator with kernel  $G(\mathbf{x}, \mathbf{y}; t)$ . This integral operator acts on  $u(\mathbf{y}, 0)$  and gives  $u(\mathbf{x}, t)$ ; that is, given the shape of the solution at  $t = 0$ , the integral operator produces the shape for all subsequent time. That is why  $G(\mathbf{x}, \mathbf{y}; t)$  is called the **evolution operator**, or **propagator**.

## 22.3 Hyperbolic Equations

The hyperbolic equation we will discuss is the wave equation

$$\mathbf{L}_{\mathbf{x},t}[u] \equiv \left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) u(\mathbf{x}, t) = 0, \quad (22.25)$$

where we have set the speed of the wave equal to unity.

We wish to calculate the Green's function for  $\mathbf{L}_{\mathbf{x},t}$  subject to appropriate BCs. Let us proceed as we did for the parabolic equation and write

$$G(\mathbf{x}, \mathbf{y}; t) = \sum_{n=1}^{\infty} C_n(\mathbf{y}; t)u_n(\mathbf{x}) \quad C_n(\mathbf{y}; t) = \int_D G(\mathbf{x}, \mathbf{y}; t)u_n(\mathbf{x})d^m x, \quad (22.26)$$

where  $u_n(\mathbf{x})$  are orthonormal eigenfunctions of  $\nabla^2$  with eigenvalues  $-\lambda_n$ , satisfying certain, as yet unspecified, BCs. As usual, we expect  $G$  to satisfy

$$\mathbf{L}_{\mathbf{x},t}[G] = \left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) G(\mathbf{x}, \mathbf{y}; t - \tau) = \delta(\mathbf{x} - \mathbf{y})\delta(t - \tau). \quad (22.27)$$

Substituting (22.26) in (22.27) with  $\tau = 0$  and using  $\nabla^2 u_n = -\lambda_n u_n$ , gives

$$\sum_{n=1}^{\infty} \left\{ \frac{\partial^2}{\partial t^2} C_n(\mathbf{y}; t) + \lambda_n C_n(\mathbf{y}; t) \right\} u_n(\mathbf{x}) = \sum_{n=1}^{\infty} [u_n(\mathbf{y})\delta(t)]u_n(\mathbf{x}),$$

where we used  $\delta(\mathbf{x} - \mathbf{y}) = \sum_{n=1}^{\infty} u_n(\mathbf{x})u_n(\mathbf{y})$  on the RHS. The orthonormality of  $u_n$  now gives  $\ddot{C}_n(\mathbf{y}; t) + \lambda_n C_n(\mathbf{y}; t) = u_n(\mathbf{y})\delta(t)$ . It follows that  $C_n(\mathbf{y}; t)$  is

separable. In fact,

$$C_n(\mathbf{y}; t) = u_n(\mathbf{y})T_n(t) \quad \text{where} \quad \left( \frac{d^2}{dt^2} + \lambda_n \right) T_n(t) = \delta(t).$$

This equation describes a one-dimensional Green's function and can be solved using the methods of Chapter 20. Assuming that  $T_n(t) = 0$  for  $t \leq 0$ , we obtain  $T_n(t) = (\sin \omega_n t / \omega_n) \theta(t)$ , where  $\omega_n^2 = \lambda_n$ . Substituting all the above results in (22.26), we obtain

$$G(\mathbf{x}, \mathbf{y}; t) = \sum_{n=1}^{\infty} u_n(\mathbf{x})u_n(\mathbf{y}) \frac{\sin \omega_n t}{\omega_n} \theta(t),$$

or, more generally,

$$G(\mathbf{x}, \mathbf{y}; t - \tau) = \sum_{n=1}^{\infty} u_n(\mathbf{x})u_n(\mathbf{y}) \frac{\sin \omega_n(t - \tau)}{\omega_n} \theta(t - \tau). \quad (22.28)$$

We note that

$$G(\mathbf{x}, \mathbf{y}; 0^+) = 0 \quad \text{and} \quad \frac{\partial G}{\partial t}(\mathbf{x}, \mathbf{y}; t) \Big|_{t \rightarrow 0^+} = \delta(\mathbf{x} - \mathbf{y}), \quad (22.29)$$

as can easily be verified.

With the Green's function for the operator  $\mathbf{L}_{\mathbf{x},t}$  of Equation (22.25) at our disposal, we can attack the BVP given by

$$\begin{aligned} \left( \frac{\partial^2}{\partial t^2} - \nabla^2 \right) u(\mathbf{x}, t) &= f(\mathbf{x}, t) \quad \text{for } \mathbf{x} \in D, \\ u(\mathbf{x}_b, t) &= h(\mathbf{x}_b, t), \quad u(\mathbf{x}, 0) = \phi(\mathbf{x}) \quad \text{for } \mathbf{x}_b \in \partial D, \quad \mathbf{x} \in D, \\ \frac{\partial u}{\partial t}(\mathbf{x}, t) \Big|_{t=0} &= \psi(\mathbf{x}) \quad \text{for } \mathbf{x} \in D. \end{aligned} \quad (22.30)$$

As in the case of the parabolic equation, we first derive an appropriate expression of Green's identity. This can be done by noting that

$$v \mathbf{L}_{\mathbf{x},t}[u] - u \mathbf{L}_{\mathbf{x},t}^\dagger[v] = \frac{\partial}{\partial t} \left( u \frac{\partial v}{\partial t} - v \frac{\partial u}{\partial t} \right) - \nabla \cdot (u \nabla v - v \nabla u).$$

Thus,  $\mathbf{L}_{\mathbf{x},t}$  is formally self-adjoint. Furthermore, we can identify

$$Q^0 = u \frac{\partial v}{\partial t} - v \frac{\partial u}{\partial t} \quad \text{and} \quad \mathbf{Q} = u \nabla v - v \nabla u.$$

Following the procedure used for the parabolic case step by step, we can easily derive a Green's identity and show that

$$\begin{aligned}
 u(\mathbf{x}, t) &= \int_0^t d\tau \int_D d^m y G(\mathbf{x}, \mathbf{y}; t - \tau) f(\mathbf{y}, \tau) \\
 &\quad + \int_D [\psi(\mathbf{y}) G(\mathbf{x}, \mathbf{y}; t) - \phi(\mathbf{y}) \frac{\partial G}{\partial t}(\mathbf{x}, \mathbf{y}; t)] d^m y \\
 &\quad - \int_0^t d\tau \int_{\partial D} h(\mathbf{y}_b, \tau) \frac{\partial G}{\partial n_y}(\mathbf{x}, \mathbf{y}_b; t - \tau) da.
 \end{aligned} \tag{22.31}$$

The details are left as Problem 22.11.

For the homogeneous PDE with the homogeneous BC  $h = 0 = \psi$ , we get

$$u(\mathbf{x}, t) = - \int_D \phi(\mathbf{y}) \frac{\partial G}{\partial t}(\mathbf{x}, \mathbf{y}; t) d^m y.$$

Note the difference between this equation and Equation (22.24). Here the propagator is the time derivative of the Green's function. There is another difference between hyperbolic and parabolic equations. When the solution to a parabolic equation vanishes on the boundary and is initially zero, and the PDE is homogeneous [ $f(\mathbf{x}, t) = 0$ ], the solution must be zero. This is clear from Equation (22.23). On the other hand, Equation (22.31) indicates that under the same circumstance, there may be a nonzero solution for a hyperbolic equation if  $\psi$  is nonzero. In such a case we obtain

$$u(\mathbf{x}, t) = \int_D \psi(\mathbf{y}) G(\mathbf{x}, \mathbf{y}; t) d^m y.$$

This difference in the two types of equations is due to the fact that hyperbolic equations have second-order time derivatives. Thus, the initial shape of a solution is not enough to uniquely specify it. The initial velocity profile is also essential. We saw examples of this in Chapter 19. The discussion of Green's functions has so far been formal. The main purpose of the remaining sections is to bridge the gap between formalism and concrete applications. Several powerful techniques are used in obtaining Green's functions, but we will focus only on two: the Fourier transform technique, and the eigenfunction expansion technique.

## 22.4 The Fourier Transform Technique

Recall that any Green's function can be written as a sum of a singular part and a regular part:  $G = G_s + H$ . Since we have already discussed homogeneous equations in detail in Chapter 19, we will not evaluate  $H$  in this section but will concentrate on the singular parts of various Green's functions.

The BCs play no role in evaluating  $G_s$ . Therefore, the Fourier transform technique (FTT), which involves integration over all space, can be utilized. The FTT

has a drawback—it does not work if the coefficient functions are not constants. For most physical applications treated in this book, however, this will not be a shortcoming.

Let us consider the most general SOLPDO with constant coefficients,

$$\mathbf{L}_x = a_0 + \sum_{j=1}^m a_j \frac{\partial}{\partial x_j} + \sum_{j,k=1}^m b_{jk} \frac{\partial^2}{\partial x_j \partial x_k}, \quad (22.32)$$

where  $a_0$ ,  $a_j$ , and  $b_{jk}$  are constants. The corresponding Green's function has a singular part that satisfies the usual PDE with the delta function on the RHS. The FTT starts with assuming a Fourier integral representation in the variable  $\mathbf{x}$  for the singular part and for the delta function:

$$G_s(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{m/2}} \int d^m k \tilde{G}_s(\mathbf{k}, \mathbf{y}) e^{i\mathbf{k}\cdot\mathbf{x}},$$

$$\delta(\mathbf{x} - \mathbf{y}) = \frac{1}{(2\pi)^m} \int d^m k e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}.$$

Substituting these equations in the PDE for the GF, we get

$$\tilde{G}_s(\mathbf{k}, \mathbf{y}) = \frac{1}{(2\pi)^{m/2}} \left( \frac{e^{-i\mathbf{k}\cdot\mathbf{y}}}{a_0 + i \sum_{j=1}^m a_j k_j - \sum_{j,l=1}^m b_{jl} k_j k_l} \right)$$

and

$$G_s(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^m} \int d^m k \frac{e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}}{a_0 + i \sum_{j=1}^m a_j k_j - \sum_{j,l=1}^m b_{jl} k_j k_l}. \quad (22.33)$$

If we can evaluate the integral in (22.33), we can find  $G$ .

The following examples apply Equation (22.33) to specific problems. Note that (22.33) indicates that  $G_s$  depends only on  $\mathbf{x} - \mathbf{y}$ . This point was mentioned in Chapter 20, where it was noted that such dependence occurs when the BCs play no part in an evaluation of the singular part of the Green's function of a DE with constant coefficients; and this is exactly the situation here.

### 22.4.1 GF for the $m$ -Dimensional Laplacian

We calculated the GF for the  $m$ -dimensional Laplacian in Section 21.2.2 using a different method. With  $a_0 = 0 = a_j$ ,  $b_{jl} = \delta_{jl}$ , and  $\mathbf{r} = \mathbf{x} - \mathbf{y}$ , Equation (22.33) reduces to

$$G_s(\mathbf{r}) = \frac{1}{(2\pi)^m} \int d^m k \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{-k^2}, \quad (22.34)$$

where  $k^2 = k_1^2 + \dots + k_m^2 = \mathbf{k} \cdot \mathbf{k}$ . To integrate (22.34), we choose spherical coordinates in the  $m$ -dimensional  $k$ -space. Furthermore, to simplify calculations

we let the  $k_m$ -axis lie along  $\mathbf{r}$  so that  $\mathbf{r} = (0, 0, \dots, |\mathbf{r}|)$  and  $\mathbf{k} \cdot \mathbf{r} = k|\mathbf{r}| \cos \theta_1$  [see Equation (21.12)]. Substituting this in (22.34) and writing  $d^m k$  in spherical coordinates yields

$$G_s(\mathbf{r}) = \frac{-1}{(2\pi)^m} \int \frac{e^{ik|\mathbf{r}| \cos \theta_1}}{k^2} k^{m-1} (\sin \theta_1)^{m-2} \dots \sin \theta_{m-2} dk d\theta_1 \dots d\theta_{m-1} \tag{22.35}$$

From Equation (21.15) we note that  $d\Omega_m = (\sin \theta_1)^{m-2} d\theta_1 d\Omega_{m-1}$ . Thus, after integrating over the angles  $\theta_2, \dots, \theta_{m-1}$ , Equation (22.35) becomes

$$G_s(\mathbf{r}) = -\frac{1}{(2\pi)^m} \Omega_{m-1} \int_0^\infty k^{m-3} dk \int_0^\pi (\sin \theta_1)^{m-2} e^{ik|\mathbf{r}| \cos \theta_1} d\theta_1.$$

The inner integral can be looked up in an integral table (see [Grad 65, p. 482]):

$$\int_0^\pi (\sin \theta_1)^{m-2} e^{ik|\mathbf{r}| \cos \theta_1} d\theta_1 = \sqrt{\pi} \left(\frac{2}{kr}\right)^{m/2-1} \Gamma\left(\frac{m-1}{2}\right) J_{m/2-1}(kr).$$

Substituting this and (21.16) in the preceding equation and using the result (see [Grad 65, p. 684])

$$\int_0^\infty x^\mu J_\nu(ax) dx = 2^\mu a^{-\mu-1} \frac{\Gamma\left(\frac{\mu+\nu+1}{2}\right)}{\Gamma\left(\frac{\mu-\nu+1}{2}\right)}$$

we obtain

$$G_s(\mathbf{r}) = -\frac{\Gamma(m/2-1)}{4\pi^{m/2}} \left(\frac{1}{r^{m-2}}\right) \quad \text{for } m > 2,$$

which agrees with (21.20) since  $\Gamma(m/2) = (m/2-1)\Gamma(m/2-1)$ .

### 22.4.2 GF for the $m$ -Dimensional Helmholtz Operator

For the Helmholtz operator  $\nabla^2 - \mu^2$ , Equation (22.33) reduces to

$$G_s(\mathbf{r}) = -\frac{1}{(2\pi)^m} \int d^m k \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\mu^2 + k^2}.$$

Following the same procedure as in the previous subsection, we find

$$\begin{aligned} G_s(\mathbf{r}) &= -\frac{\Omega_{m-1}}{(2\pi)^m} \int_0^\infty \frac{k^{m-1} dk}{\mu^2 + k^2} \int_0^\pi (\sin \theta_1)^{m-2} e^{ikr \cos \theta_1} d\theta_1 \\ &= -\frac{\Omega_{m-1}}{(2\pi)^m} \sqrt{\pi} \left(\frac{2}{r}\right)^{m/2-1} \Gamma\left(\frac{m-1}{2}\right) \int_0^\infty \frac{k^{m/2}}{\mu^2 + k^2} J_{m/2-1}(kr) dk. \end{aligned}$$



Here we can use the integral formula (see [Grad 65, pp. 686 and 952])

$$\int_0^\infty \frac{J_\nu(bx)x^{\nu+1}}{(x^2+a^2)^{\eta+1}} dx = \frac{a^{\nu-\eta}b^\eta}{2^\eta\Gamma(\eta+1)}K_{\nu-\eta}(ab),$$

where

$$K_\nu(z) = \frac{i\pi}{2}e^{i\nu\pi/2}H_\nu^{(1)}(iz),$$

to obtain

$$G_s(\mathbf{r}) = -\frac{\Omega_{m-1}}{(2\pi)^m}\sqrt{\pi}\left(\frac{2}{r}\right)^{m/2-1}\Gamma\left(\frac{m-1}{2}\right)\mu^{m/2-1}\frac{\pi}{2}e^{im\pi/4}H_{m/2-1}^{(1)}(i\mu r),$$

which simplifies to

$$G_s(\mathbf{r}) = -\frac{\pi/2}{(2\pi)^{m/2}}\left(\frac{\mu}{r}\right)^{m/2-1}e^{im\pi/4}H_{m/2-1}^{(1)}(i\mu r). \tag{22.36}$$

It can be shown (see Problem 22.8) that for  $m = 3$  this reduces to  $G_s(\mathbf{r}) = -\frac{e^{-\mu r}}{4\pi r}$ , which is the Yukawa potential due to a unit charge.

We can easily obtain the GF for  $\nabla^2 + \mu^2$  by substituting  $\pm i\mu$  for  $\mu$  in Equation (22.36). The result is

$$G_s(\mathbf{r}) = i^{m+1}\frac{\pi/2}{(2\pi)^{m/2}}\left(\frac{\mu}{r}\right)^{m/2-1}H_{m/2-1}^{(1)}(\pm i\mu r). \tag{22.37}$$

For  $m = 3$  this yields  $G_s(\mathbf{r}) = -e^{\pm i\mu r}/(4\pi r)$ . The two signs in the exponent correspond to the so-called incoming and outgoing “waves.”

**22.4.1. Example.** For a non-local potential, the time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2\mu}\nabla^2\Psi + \int_{\mathbb{R}^3} V(\mathbf{r}, \mathbf{r}')\Psi(\mathbf{r}') d^3r' = E\Psi(\mathbf{r}).$$

Then, the integral equation associated with this differential equation is (see Section 21.4)

$$\Psi(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} - \frac{\mu}{2\pi\hbar^2} \int_{\mathbb{R}^3} d^3r' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \int_{\mathbb{R}^3} d^3r'' V(\mathbf{r}', \mathbf{r}'')\Psi(\mathbf{r}''). \tag{22.38}$$

For a separable potential, for which  $V(\mathbf{r}', \mathbf{r}'') = -g^2U(\mathbf{r}')U(\mathbf{r}'')$ , we can solve (22.38) exactly. We substitute for  $V(\mathbf{r}', \mathbf{r}'')$  in (22.38) to obtain

$$\Psi(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} + \frac{\mu g^2}{2\pi\hbar^2} \int_{\mathbb{R}^3} d^3r' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} U(\mathbf{r}') \int_{\mathbb{R}^3} d^3r'' U(\mathbf{r}'')\Psi(\mathbf{r}''). \tag{22.39}$$

Non-local potentials depend not only on the observation point, but also on some other “non-local” variables.

Defining the quantities

$$Q(\mathbf{r}) \equiv \frac{\mu g^2}{2\pi \hbar^2} \int_{\mathbb{R}^3} d^3 r' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} U(\mathbf{r}'), \quad C \equiv \int_{\mathbb{R}^3} d^3 r'' U(\mathbf{r}'') \Psi(\mathbf{r}'') \quad (22.40)$$

and substituting them in (22.39) yields  $\Psi(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} + CQ(\mathbf{r})$ . Multiplying both sides of this equation by  $U(\mathbf{r})$  and integrating over  $\mathbb{R}^3$ , we get

$$\begin{aligned} C &= A \int_{\mathbb{R}^3} e^{i\mathbf{k}\cdot\mathbf{r}} U(\mathbf{r}) d^3 r + C \int_{\mathbb{R}^3} U(\mathbf{r}) Q(\mathbf{r}) d^3 r \\ &= (2\pi)^{3/2} A \tilde{U}(-\mathbf{k}) + C \int_{\mathbb{R}^3} U(\mathbf{r}) Q(\mathbf{r}) d^3 r, \end{aligned}$$

from which we obtain  $C = (2\pi)^{3/2} A \tilde{U}(-\mathbf{k}) / [1 - \int_{\mathbb{R}^3} U(\mathbf{r}) Q(\mathbf{r}) d^3 r]$ , leading to the solution

$$\Psi(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} + \frac{(2\pi)^{3/2} A \tilde{U}(-\mathbf{k})}{1 - \int_{\mathbb{R}^3} U(\mathbf{r}') Q(\mathbf{r}') d^3 r'} Q(\mathbf{r}). \quad (22.41)$$

In principle,  $\tilde{U}(-\mathbf{k})$  [the Fourier transform of  $U(\mathbf{r})$ ] and  $Q(\mathbf{r})$  can be calculated once the functional form of  $U(\mathbf{r})$  is known. Equations (22.40) and (22.41) give the solution to the Schrödinger equation in closed form. ■

When dealing with parabolic and hyperbolic equations, we will find it convenient to consider the “different” variable (usually  $t$ ) as the zeroth coordinate. In the Fourier transform we then use  $\omega = -k_0$  and write

$$\begin{aligned} G_s(\mathbf{r}, t) &= \frac{1}{(2\pi)^{(m+1)/2}} \int_{-\infty}^{\infty} d\omega \int d^m k \tilde{G}_s(\mathbf{k}, \omega) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}, \\ \delta(\mathbf{r})\delta(t) &= \frac{1}{(2\pi)^{m+1}} \int_{-\infty}^{\infty} d\omega \int d^m k e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}, \end{aligned} \quad (22.42)$$

where  $\mathbf{r}$  is the  $m$ -dimensional position vector.

### 22.4.3 GF for the $m$ -Dimensional Diffusion Operator

We substitute from (22.42) in  $(\partial/\partial t - \nabla^2)G_s(\mathbf{r}, t) = \delta(\mathbf{r})\delta(t)$  to obtain

$$G_s(\mathbf{r}, t) = \frac{1}{(2\pi)^{m+1}} \int d^m k e^{i\mathbf{k}\cdot\mathbf{r}} \int_{-\infty}^{\infty} d\omega \frac{e^{-i\omega t}}{\omega + ik^2}, \quad (22.43)$$

where as usual,  $k^2 = \sum_{i=1}^m k_i^2$ . The  $\omega$  integration can be done using the calculus of residues. The integrand has a simple pole at  $\omega = -ik^2$ , that is, in the lower half of the complex  $\omega$ -plane (LHP). To integrate, we must know the sign of  $t$ . If  $t > 0$ , the exponential factor dictates that the contour be closed in the LHP, where there is a pole and, therefore, a contribution to the residues. On the other hand, if  $t < 0$ , the contour must be closed in the UHP. The integral is then zero because

there are no poles in the upper half plane (UHP). We must therefore introduce a step function  $\theta(t)$  in the Green's function. Evaluating the residue, the  $\omega$  integration yields  $-2\pi i e^{-k^2 t}$ . (The minus sign arises because of clockwise contour integration in the LHP.) Substituting this in Equation (22.43), using spherical coordinates in which the last  $k$ -axis is along  $\mathbf{r}$ , and integrating over all angles except  $\theta_1$ , we obtain

$$\begin{aligned} G_s(\mathbf{r}, t) &= \theta(t) \frac{\Omega_{m-1}}{(2\pi)^m} \int_0^\infty k^{m-1} dk e^{-k^2 t} \int_0^\pi (\sin \theta_1)^{m-2} e^{ikr \cos \theta_1} d\theta_1 \\ &= \theta(t) \frac{\Omega_{m-1}}{(2\pi)^m} \sqrt{\pi} \left(\frac{2}{r}\right)^{m/2-1} \Gamma\left(\frac{m-1}{2}\right) \int_0^\infty k^{m/2} e^{-k^2 t} J_{m/2-1}(kr) dk. \end{aligned}$$

For the  $\theta_1$  integration, we used the result quoted in Section 22.4.1.

Using the integral formula (see [Grad 65, pp. 716 and 1058])

$$\int_0^\infty x^\mu e^{-\alpha x^2} J_\nu(\beta x) dx = \frac{\beta^\nu \Gamma\left(\frac{\mu+\nu+1}{2}\right)}{2^{\nu+1} \alpha^{(\mu+\nu+1)/2} \Gamma(\nu+1)} \Phi\left(\frac{\mu+\nu+1}{2}, \nu+1; -\frac{\beta^2}{4\alpha}\right),$$

where  $\Phi$  is the confluent hypergeometric function, we obtain

$$G_s(\mathbf{r}, t) = \theta(t) \frac{2\pi^{(m-1)/2}}{(2\pi)^m} \sqrt{\pi} \left(\frac{2}{r}\right)^{m/2-1} \frac{r^{m/2-1}}{2^{m/2} t^{m/2}} \Phi\left(\frac{m}{2}, \frac{m}{2}; -\frac{r^2}{4t}\right). \quad (22.44)$$

The power-series expansion for the confluent hypergeometric function  $\Phi$  shows that  $\Phi(\alpha, \alpha; z) = e^z$ . Substituting this result in (22.44) and simplifying, we finally obtain

$$G_s(\mathbf{r}, t) = \frac{e^{-r^2/4t}}{(4\pi t)^{m/2}} \theta(t). \quad (22.45)$$

#### 22.4.4 GF for the $m$ -Dimensional Wave Equation

The difference between this example and the preceding one is that here the time derivative is of second order. Thus, instead of Equation (22.43), we start with

$$G_s(\mathbf{r}, t) = -\frac{1}{(2\pi)^{m+1}} \int d^m k e^{i\mathbf{k}\cdot\mathbf{r}} \int_{-\infty}^\infty d\omega \frac{e^{-i\omega t}}{\omega^2 - k^2}. \quad (22.46)$$

The  $\omega$  integration can be done using the method of residues. Since the singularities of the integrand,  $\omega = \pm k$ , are on the real axis, it seems reasonable to use the principal value as the value of the integral. This, in turn, depends on the sign of  $t$ . If  $t > 0$  ( $t < 0$ ), we have to close the contour in the LHP (UHP): to avoid the explosion of the exponential. If one also insists on not including the poles inside the contour,<sup>3</sup> then one can show that

$$P \int_{-\infty}^\infty d\omega \frac{e^{-i\omega t}}{\omega^2 - k^2} = -\pi \frac{\sin kt}{k} \epsilon(t),$$

<sup>3</sup>This will determine how to (semi)circle around the poles.

where

$$\epsilon(t) \equiv \theta(t) - \theta(-t) = \begin{cases} 1 & \text{if } t > 0, \\ -1 & \text{if } t < 0. \end{cases}$$

Substituting this in (22.46) and integrating over all angles as done in the previous examples yields

$$G_s(\mathbf{r}, t) = \frac{\epsilon(t)}{2(2\pi)^{m/2} r^{m/2-1}} \int_0^\infty k^{m/2-1} J_{m/2-1}(kr) \sin kt \, dk. \quad (22.47)$$

Physics determines the contour of integration

As Problem 22.25 shows, the Green's function given by Equation (22.47) satisfies only the homogeneous wave equation with no delta function on the RHS. The reason for this is that the principal value of an integral chooses a specific contour that may not reflect the physical situation. In fact, the Green's function in (22.47) contains two pieces corresponding to the two different contours of integration, and it turns out that the physically interesting Green's functions are obtained, not from the principal value, but from giving small imaginary parts to the poles. Thus, replacing the  $\omega$  integral with a contour integral for which the two poles are pushed in the LHP and using the method of residues, we obtain

$$I_{up} \equiv \int_{-\infty}^\infty d\omega \frac{e^{-i\omega t}}{\omega^2 - k^2} = \int_{C_1} \frac{e^{-izt}}{z^2 - k^2} dz = \frac{2\pi}{k} \theta(t) \sin kt.$$

The integral is zero for  $t < 0$  because for negative values of  $t$ , the contour must be closed in the UHP, where there are no poles inside  $C_1$ . Substituting this in (22.46) and working through as before, we obtain what is called the **retarded Green's function**:

retarded Green's function

$$G_s^{(ret)}(\mathbf{r}, t) = \frac{\theta(t)}{(2\pi)^{m/2} r^{m/2-1}} \int_0^\infty k^{m/2-1} J_{m/2-1}(kr) \sin kt \, dk. \quad (22.48)$$

advanced Green's function

If the poles are pushed in the UHP we obtain the **advanced Green's function**:

$$G_s^{(adv)}(\mathbf{r}, t) = -\frac{\theta(-t)}{(2\pi)^{m/2} r^{m/2-1}} \int_0^\infty k^{m/2-1} J_{m/2-1}(kr) \sin kt \, dk. \quad (22.49)$$

Unlike the elliptic and parabolic equations discussed earlier, the integral over  $k$  is not a function but a *distribution*, as will become clear below. To find the retarded and advanced Green's functions, we write the sine term in the integral in terms of exponentials and use the following (see [Grad 65, p. 712]):

$$\int_0^\infty x^\nu e^{-\alpha x} J_\nu(\beta x) \, dx = \frac{(2\beta)^\nu \Gamma(\nu + 1/2)}{\sqrt{\pi}(\alpha^2 + \beta^2)^{\nu+1/2}} \quad \text{for } \text{Re}(\alpha) > |\text{Im}(\beta)|.$$

To ensure convergence at infinity, we add a small negative number to the exponential and define the integral

$$I_{\epsilon}^{\pm} \equiv \int_0^{\infty} k^{\nu} e^{-(\mp it + \epsilon)k} J_{\nu}(kr) dk = \frac{(2r)^{\nu} \Gamma(\nu + 1/2)}{\sqrt{\pi}} [( \mp it + \epsilon)^2 + r^2]^{-(\nu + 1/2)}.$$

For the GFs, we need to evaluate the (common) integral in (22.48) and (22.49). With  $\nu = m/2 - 1$ , we have

$$\begin{aligned} I^{(\nu)} &\equiv \int_0^{\infty} k^{\nu} J_{\nu}(kr) \sin kt dk = \frac{1}{2i} \lim_{\epsilon \rightarrow 0} (I_{\epsilon}^{+} - I_{\epsilon}^{-}) \\ &= \frac{(2r)^{\nu} \Gamma(\nu + 1/2)}{2i\sqrt{\pi}} \lim_{\epsilon \rightarrow 0} \left\{ \frac{1}{[r^2 + (-it + \epsilon)^2]^{\nu + 1/2}} - \frac{1}{[r^2 + (it + \epsilon)^2]^{\nu + 1/2}} \right\}. \end{aligned}$$

At this point, it is convenient to discuss separately the two cases of  $m$  odd and  $m$  even. Let us derive the expression for odd  $m$  (the even case is left for Problem 22.26). Define the integer  $n = (m - 1)/2 = \nu + \frac{1}{2}$  and write  $I^{(\nu)}$  as

$$I^{(n)} = \frac{(2r)^{n-1/2} \Gamma(n)}{2i\sqrt{\pi}} \lim_{\epsilon \rightarrow 0} \left\{ \frac{1}{[r^2 + (-it + \epsilon)^2]^n} - \frac{1}{[r^2 + (it + \epsilon)^2]^n} \right\}. \quad (22.50)$$

Define  $u = r^2 + (-it + \epsilon)^2$ . Then using the identity

$$\frac{1}{u^n} = \frac{(-1)^{n-1}}{(n-1)!} \frac{d^{n-1}}{du^{n-1}} \left( \frac{1}{u} \right)$$

and the chain rule,  $df/du = (1/2r)\partial f/\partial r$ , we obtain  $d/du = (1/2r)\partial/\partial r$  and

$$\frac{1}{[r^2 + (\pm it + \epsilon)^2]^n} = \frac{1}{(n-1)!} \left( -\frac{1}{2r} \frac{\partial}{\partial r} \right)^{n-1} \left[ \frac{1}{[r^2 + (\pm it + \epsilon)^2]} \right].$$

Therefore, Equation (22.50) can be written as

$$\begin{aligned} I^{(n)} &= \int_0^{\infty} k^{n-1/2} J_{n-1/2}(kr) \sin kt dk \\ &= \frac{(2r)^{n-1/2} \Gamma(n)}{2i\sqrt{\pi}} \frac{1}{(n-1)!} \\ &\quad \cdot \left( -\frac{1}{2r} \frac{\partial}{\partial r} \right)^{n-1} \left\{ \lim_{\epsilon \rightarrow 0} \left[ \frac{1}{[r^2 + (-it + \epsilon)^2]} - \frac{1}{[r^2 + (it + \epsilon)^2]} \right] \right\}. \end{aligned} \quad (22.51)$$

The limit in (22.51) is found in Problem 22.27. Using the result of that problem and  $\Gamma(n) = (n-1)!$ , we get

$$\begin{aligned} I^{(n)} &= \int_0^{\infty} k^{n-1/2} J_{n-1/2}(kr) \sin kt dk \\ &= -\frac{\sqrt{\pi}(2r)^{n-1/2}}{2} \left( -\frac{1}{2r} \frac{\partial}{\partial r} \right)^{n-1} \left\{ \frac{1}{r} [\delta(t+r) - \delta(t-r)] \right\}. \end{aligned} \quad (22.52)$$

Employing this result in (22.48) and (22.49) yields

$$\begin{aligned} G_s^{(\text{ret})}(\mathbf{r}, t) &= \frac{1}{4\pi} \left( -\frac{1}{2\pi r} \frac{\partial}{\partial r} \right)^{n-1} \left[ \frac{\delta(t-r)}{r} \right] && \text{for } n = \frac{m-1}{2}, \\ G_s^{(\text{adv})}(\mathbf{r}, t) &= \frac{1}{4\pi} \left( -\frac{1}{2\pi r} \frac{\partial}{\partial r} \right)^{n-1} \left[ \frac{\delta(t+r)}{r} \right] && \text{for } n = \frac{m-1}{2}. \end{aligned} \quad (22.53)$$

The theta functions are not needed in (22.53) because the arguments of the delta functions already meet the restrictions imposed by the theta functions.

The two functions in (22.53) have an interesting physical interpretation. Green's functions are propagators (of signals of some sort), and  $G_s^{(\text{ret})}(\mathbf{r}, t)$  is capable of propagating signals only for positive times. On the other hand,  $G_s^{(\text{adv})}(\mathbf{r}, t)$  can propagate only in the negative time direction. Thus, if initially ( $t = 0$ ) a signal is produced (by appropriate BCs), both  $G_s^{(\text{ret})}(\mathbf{r}, t)$  and  $G_s^{(\text{adv})}(\mathbf{r}, t)$  work to propagate it in their respective time directions. It may seem that  $G_s^{(\text{adv})}(\mathbf{r}, t)$  is useless because every signal propagates forward in time. This is true, however, only for classical events. In relativistic quantum field theory antiparticles are interpreted mathematically as moving in the negative time direction! Thus, we cannot simply ignore  $G_s^{(\text{adv})}(\mathbf{r}, t)$ . In fact, the correct propagator to choose in this theory is a linear combination of  $G_s^{(\text{adv})}(\mathbf{r}, t)$  and  $G_s^{(\text{ret})}(\mathbf{r}, t)$ , called the **Feynman propagator** (see [Wein 95, pp. 274–280]). The foregoing example shows a subtle difference between Green's functions for second-order differential operators in one dimension and in higher dimensions. We saw in Chapter 20 that the former are continuous functions in the interval on which they are defined. Here, we see that higher dimensional Green's functions are not only discontinuous, but that they are not even *functions* in the ordinary sense; they contain a delta function. Thus, in general, Green's functions in higher dimensions ought to be treated as distributions (generalized functions).

Feynman propagator

## 22.5 The Eigenfunction Expansion Technique

Suppose that the differential operator  $L_{\mathbf{x}}$ , defined in a domain  $D$  with boundary  $\partial D$ , has discrete eigenvalues  $\{\lambda_n\}_{n=1}^{\infty}$  with corresponding orthonormal eigenfunctions  $\{u_n(\mathbf{x})\}_{n=1}^{\infty}$ . These two sets may not be in one-to-one correspondence. Assume that the  $u_n(\mathbf{x})$ 's satisfy the same BCs as the Green's function to be defined below.

Now consider the operator  $L_{\mathbf{x}} - \lambda 1$ , where  $\lambda$  is different from all  $\lambda_n$ 's. Then, as in the one-dimensional case, this operator is invertible, and we can define its Green's function by  $(L_{\mathbf{x}} - \lambda)G_{\lambda}(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y})$  where the weight function is set equal to one. The completeness of  $\{u_n(\mathbf{x})\}_{n=1}^{\infty}$  implies that

$$\delta(\mathbf{x} - \mathbf{y}) = \sum_{n=1}^{\infty} u_n(\mathbf{x})u_n^*(\mathbf{y}) \quad \text{and} \quad G_{\lambda}(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} a_n(\mathbf{y})u_n(\mathbf{x}).$$

Substituting these two expansions in the differential equation for GF yields

$$\sum_{n=1}^{\infty} (\lambda_n - \lambda) a_n(\mathbf{y}) u_n(\mathbf{x}) = \sum_{n=1}^{\infty} u_n(\mathbf{x}) u_n^*(\mathbf{y}).$$

The orthonormality of the  $u_n$ 's gives  $a_n(\mathbf{y}) = u_n^*(\mathbf{y})/(\lambda_n - \lambda)$ . Therefore,

$$G_\lambda(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} \frac{u_n(\mathbf{x}) u_n^*(\mathbf{y})}{\lambda_n - \lambda}. \quad (22.54)$$

In particular, if zero is not an eigenvalue of  $\mathbf{L}_x$ , its Green's function can be written as

$$G(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} \frac{u_n(\mathbf{x}) u_n^*(\mathbf{y})}{\lambda_n}. \quad (22.55)$$

This is an expansion of the Green's function in terms of the eigenfunctions of  $\mathbf{L}_x$ .

It is instructive to consider a formal interpretation of Equation (22.55). Recall that the spectral decomposition theorem permits us to write  $f(\mathbf{A}) = \sum_i f(\lambda_i) \mathbf{P}_i$  for an operator  $\mathbf{A}$  with (distinct) eigenvalues  $\lambda_i$  and projection operators  $\mathbf{P}_i$ . Allowing repetition of eigenvalues in the sum, we may write  $f(\mathbf{A}) = \sum_n f(\lambda_n) |u_n\rangle \langle u_n|$ , where  $n$  counts *all* the eigenfunctions corresponding to eigenvalues. Now, let  $f(\mathbf{A}) = \mathbf{A}^{-1}$ . Then

$$\mathbf{G} = \mathbf{A}^{-1} = \sum_n \lambda_n^{-1} |u_n\rangle \langle u_n| = \sum_n \frac{|u_n\rangle \langle u_n|}{\lambda_n},$$

or, in "matrix element" form,

$$G(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} | \mathbf{G} | \mathbf{y} \rangle = \sum_n \frac{\langle \mathbf{x} | u_n \rangle \langle u_n | \mathbf{y} \rangle}{\lambda_n} = \sum_n \frac{u_n(\mathbf{x}) u_n^*(\mathbf{y})}{\lambda_n}.$$

This last expression coincides with the RHS of Equation (22.55).

Equations (22.54) and (22.55) demand that the  $u_n(\mathbf{x})$  form a complete discrete orthonormal set. We encountered many examples of such eigenfunctions in discussing Sturm–Liouville systems in Chapter 19. All the S–L systems there were, of course, one-dimensional. Here we are generalizing the S–L system to  $m$  dimensions. This is not a limitation, however, because—for the PDEs of interest—the separation of variables reduces an  $m$ -dimensional PDE to  $m$  one-dimensional ODEs. If the BCs are appropriate, the  $m$  ODEs will all be S–L systems. A review of Chapter 19 will reveal that homogeneous BCs always lead to S–L systems. In fact, Theorem 19.1.1 guarantees this claim. Since Green's functions must also satisfy homogeneous BCs, expansions such as those of (22.54) and (22.55) become possible.

**22.5.1. Example.** As a concrete example, let us obtain an eigenfunction expansion of the GF of the two-dimensional Laplacian,  $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ , inside the rectangular region  $0 \leq x \leq a, 0 \leq y \leq b$  with Dirichlet BCs. Since the GF vanishes at the boundary, the eigenvalue problem becomes  $\nabla^2 u = \lambda u$  with  $u(0, y) = u(a, y) = u(x, 0) = u(x, b) = 0$ . The method of separation of variables gives the orthonormal eigenfunctions<sup>4</sup>

$$u_{mn}(x, y) = \frac{2}{\sqrt{ab}} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right) \quad \text{for } m, n = 1, 2, \dots,$$

whose corresponding eigenvalues are  $\lambda_{mn} = -\left[\left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2\right]$ .

Inserting the eigenfunctions and the eigenvalues in Equation (22.55), we obtain

$$\begin{aligned} G(\mathbf{r}, \mathbf{r}') &= G(x, y; x', y') = \sum_{m,n=1}^{\infty} \frac{u_{mn}(x, y)u_{mn}(x', y')}{\lambda_{mn}} \\ &= -\frac{4}{ab} \sum_{m,n=1}^{\infty} \frac{\sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{m\pi}{b}y\right) \sin\left(\frac{n\pi}{a}x'\right) \sin\left(\frac{m\pi}{b}y'\right)}{\left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2}, \end{aligned}$$

where we changed  $\mathbf{x}$  to  $\mathbf{r}$  and  $\mathbf{y}$  to  $\mathbf{r}'$ . Note that the eigenvalues are never zero; thus,  $G(\mathbf{r}, \mathbf{r}')$  is well-defined. ■

In the preceding example, zero was not an eigenvalue of  $L_{\mathbf{x}}$ . This condition must hold when a Green's function is expanded in terms of eigenfunctions. In physical applications, certain conditions (which have nothing to do with the BCs) exclude the zero eigenvalue automatically when they are applied to the Green's function. For instance, the condition that the Green's function remain finite at the origin is severe enough to exclude the zero eigenvalue.

**22.5.2. Example.** Let us consider the two-dimensional Dirichlet BVP  $\nabla^2 u = f$ , with  $u = 0$  on a circle of radius  $a$ . If we consider only the BCs and ask whether zero is an eigenvalue of  $\nabla^2$ , the answer will be yes, as the following argument shows.

The most general solution to the zero-eigenvalue equation,  $\nabla^2 u = 0$ , in polar coordinates can be obtained by the method of separation of variables:

$$u(\rho, \varphi) = A + B \ln \rho + \sum_{n=1}^{\infty} (b_n \rho^n + b'_n \rho^{-n}) \cos n\varphi + \sum_{n=1}^{\infty} (c_n \rho^n + c'_n \rho^{-n}) \sin n\varphi. \tag{22.56}$$

Invoking the BC gives

$$0 = u(a, \varphi) = A + B \ln a + \sum_{n=1}^{\infty} (b_n a^n + b'_n a^{-n}) \cos n\varphi + \sum_{n=1}^{\infty} (c_n a^n + c'_n a^{-n}) \sin n\varphi,$$

which holds for arbitrary  $\varphi$  if and only if

$$A = -B \ln a, \quad b'_n = -b_n a^{2n}, \quad c'_n = -c_n a^{2n}.$$

<sup>4</sup>The inner product is defined as a double integral over the rectangle.



Substituting in (22.56) gives

$$u(\rho, \varphi) = B \ln \left( \frac{\rho}{a} \right) + \sum_{n=1}^{\infty} \left( \rho^n - \frac{a^{2n}}{\rho^n} \right) (b_n \cos n\varphi + c_n \sin n\varphi). \quad (22.57)$$

Thus, if we demand nothing beyond the BCs,  $\nabla^2$  will have a nontrivial eigen-solution corresponding to the zero eigenvalue, given by Equation (22.57).

Physical reality, however, demands that  $u(\rho, \varphi)$  be well-behaved at the origin. This condition sets  $B$ ,  $b'_n$ , and  $c'_n$  of Equation (22.56) equal to zero. The BCs then make the remaining coefficients in (22.56) vanish. Thus, the demand that  $u(\rho, \varphi)$  be well-behaved at  $\rho = 0$  turns the situation completely around and ensures the nonexistence of a zero eigenvalue for the Laplacian, which in turn guarantees the existence of a GF. ■

In many cases the operator  $L_{\mathbf{x}}$  as a whole is not amenable to a full Sturm-Liouville treatment, and as such will not yield orthonormal eigenvectors in terms of which the GF can be expanded. However, it may happen that  $L_{\mathbf{x}}$  can be broken up into two pieces one of which is an S-L operator. In such a case, the GF can be found as follows: Suppose that  $L_1$  and  $L_2$  are two commuting operators with  $L_2$  an S-L operator whose eigenvalues and eigenfunctions are known. Since  $L_2$  commutes with  $L_1$ , it can be regarded as a constant as far as operations with (and on)  $L_1$  are concerned. In particular,  $(L_1 + L_2)G = 1$  can be regarded as an operator equation in  $L_1$  alone with  $L_2$  treated as a constant. Let  $\mathbf{x}_1$  denote the subset of the variables on which  $L_1$  acts, and let  $\mathbf{x}_2$  denote the remainder of the coordinates. Then we can write  $\delta(\mathbf{x} - \mathbf{y}) = \delta(\mathbf{x}_1 - \mathbf{y}_1)\delta(\mathbf{x}_2 - \mathbf{y}_2)$ . Now let  $G_1(\mathbf{x}_1, \mathbf{y}_1; k)$  denote the Green's function for  $L_1 + k$ , where  $k$  is a constant. Then it is easily verified that

$$G(\mathbf{x}, \mathbf{y}) = G_1(\mathbf{x}_1, \mathbf{y}_1; L_2)\delta(\mathbf{x}_2 - \mathbf{y}_2). \quad (22.58)$$

In fact,

$$\begin{aligned} (L_1 + L_2)G(\mathbf{x}, \mathbf{y}) &= \underbrace{[(L_1 + L_2)G_1(\mathbf{x}_1, \mathbf{y}_1; L_2)]}_{=\delta(\mathbf{x}_1 - \mathbf{y}_1) \text{ by definition of } G_1} \delta(\mathbf{x}_2 - \mathbf{y}_2). \end{aligned}$$

Once  $G_1$  is found as a function of  $L_2$ , it can operate on  $\delta(\mathbf{x}_2 - \mathbf{y}_2)$  to yield the desired Green's function. The following example illustrates the technique.

**22.5.3. Example.** Let us evaluate the Dirichlet GF for the two-dimensional Helmholtz operator  $\nabla^2 - k^2$  in the infinite strip  $0 \leq x \leq a$ ,  $-\infty < y < \infty$ . Let  $L_1 = \partial^2/\partial y^2 - k^2$  and  $L_2 = \partial^2/\partial x^2$ . Then,

$$G(\mathbf{r}, \mathbf{r}') \equiv G(x, x', y, y') = G_1(y, y'; L_2)\delta(x - x'),$$

where  $(d^2/dy^2 - \mu^2)G_1 = \delta(y - y')$ ,  $\mu^2 \equiv k^2 - L_2$ , and  $G_1(y = -\infty) = G_1(y = \infty) = 0$ . The GF  $G_1$  can be readily found (see Problem 20.12):

$$G_1(y, y'; L_2) = -\frac{e^{-\mu|y-y'|}}{2\mu} = -\frac{e^{-\sqrt{k^2 - L_2}|y-y'|}}{2\sqrt{k^2 - L_2}}.$$

The full GF is then

$$G(\mathbf{r}, \mathbf{r}') = \left( -\frac{e^{-\sqrt{k^2 - \mathbf{L}_2}|y-y'|}}{2\sqrt{k^2 - \mathbf{L}_2}} \right) \delta(x - x'). \quad (22.59)$$

The operator  $\mathbf{L}_2$  constitutes an S-L system with eigenvalues  $\lambda_n = -(n\pi x/a)^2$  and eigenfunctions  $u_n(x) = \sqrt{2/a} \sin(n\pi x/a)$  where  $n = 1, 2, \dots$ . Therefore, the delta function  $\delta(x - x')$  can be expanded in terms of these eigenfunctions:

$$\delta(x - x') = \frac{2}{a} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{n\pi}{a}x'\right).$$

As  $\mu = \sqrt{k^2 - \mathbf{L}_2}$  acts on the delta function,  $\mathbf{L}_2$  operates on the first factor in the above expansion and gives  $\lambda_n$ . Thus,  $\mathbf{L}_2$  in Equation (22.59) can be replaced by  $-(n\pi x/a)^2$ , and we have

$$G(\mathbf{r}, \mathbf{r}') = -\frac{1}{a} \sum_{n=1}^{\infty} \left( -\frac{e^{-\sqrt{k^2 + (n\pi x/a)^2}|y-y'|}}{\sqrt{k^2 + (n\pi x/a)^2}} \right) \sin\left(\frac{n\pi}{a}x\right) \sin\left(\frac{n\pi}{a}x'\right). \quad \blacksquare$$

Sometimes it is convenient to break an operator into more than two parts. In fact, in some cases it may be advantageous to define a set of commuting self-adjoint (differential) operators  $\{\mathbf{M}_j\}$  such that the full operator  $\mathbf{L}$  can be written as  $\mathbf{L} = \sum_j \mathbf{L}_j \mathbf{M}_j$  where the differential operators  $\{\mathbf{L}_j\}$  act on variables on which the  $\mathbf{M}_j$  have no action. Since the  $\mathbf{M}_j$ 's commute among themselves, one can find simultaneous eigenfunctions for all of them. Then one expands part of the delta function in terms of these eigenfunctions in the hope that the ensuing problem becomes more manageable. The best way to appreciate this approach is via an example.

**22.5.4. Example.** Let us consider the Laplacian in spherical coordinates,

$$\nabla^2 u = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \left[ \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{\partial^2 u}{\partial \varphi^2} \right].$$

If we introduce

$$\begin{aligned} \mathbf{M}_1 u &= u, & \mathbf{L}_1 u &= \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial u}{\partial r} \right), \\ \mathbf{M}_2 u &= \frac{1}{\sin \theta} \left[ \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{\partial^2 u}{\partial \varphi^2} \right], & \mathbf{L}_2 u &= \frac{1}{r^2} u, \end{aligned} \quad (22.60)$$

the Laplacian becomes  $\nabla^2 = \mathbf{L}_1 \mathbf{M}_1 + \mathbf{L}_2 \mathbf{M}_2$ . The mutual eigenfunctions of  $\mathbf{M}_1$  and  $\mathbf{M}_2$  are simply those of  $\mathbf{M}_2$ , which is (the negative of) the angular momentum operator discussed in Chapter 12, whose eigenfunctions are the spherical harmonics. We thus have  $\mathbf{M}_2 Y_{lm}(\theta, \varphi) = -l(l+1)Y_{lm}(\theta, \varphi)$ .

Let us expand the Green's function in terms of the spherical harmonics:

$$G(\mathbf{r}, \mathbf{r}') = \sum_{l,m} g_{lm}(r; r', \theta', \varphi') Y_{lm}(\theta, \varphi).$$

We also write the delta function as

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{\delta(r - r')\delta(\theta - \theta')\delta(\varphi - \varphi')}{r'^2 \sin \theta'} = \frac{\delta(r - r')}{r'^2} \sum_{l,m} Y_{lm}(\theta, \varphi) Y_{lm}^*(\theta', \varphi'),$$

where we have used the completeness of the spherical harmonics. Substituting all of the above in  $\nabla^2 G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')$ , we obtain

$$\begin{aligned} \nabla^2 G(\mathbf{r}, \mathbf{r}') &= (\mathbf{L}_1 \mathbf{M}_1 + \mathbf{L}_2 \mathbf{M}_2) \sum_{l,m} g_{lm}(r; r', \theta', \varphi') Y_{lm}(\theta, \varphi) \\ &= \sum_{l,m} \{[(\mathbf{L}_1 - l(l+1)\mathbf{L}_2)]g_{lm}(r; r', \theta', \varphi')\} Y_{lm}(\theta, \varphi) \\ &= \frac{\delta(r - r')}{r'^2} \sum_{l,m} Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi). \end{aligned}$$

The orthogonality of the  $Y_{lm}(\theta, \varphi)$  yields

$$[(\mathbf{L}_1 - l(l+1)\mathbf{L}_2)]g_{lm}(r; r', \theta', \varphi') = \frac{\delta(r - r')}{r'^2} Y_{lm}^*(\theta', \varphi').$$

This shows that the angular part of  $g_{lm}$  is simply  $Y_{lm}^*(\theta', \varphi')$ . Separating this from the dependence on  $\mathbf{r}$  and  $\mathbf{r}'$  and substituting for  $\mathbf{L}_1$  and  $\mathbf{L}_2$ , we obtain

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dg_{lm}}{dr} \right) - \frac{l(l+1)}{r^2} g_{lm} = \frac{\delta(r - r')}{r^2}, \quad (22.61)$$

where this last  $g_{lm}$  is a function of  $r$  and  $r'$  only. The techniques of Chapter 20 can be employed to solve Equation (22.61) (see Problem 22.29). ■

The separation of the full operator into two “smaller” operators can also be used for cases in which both operators have eigenvalues and eigenvectors. The result of such an approach will, of course, be equivalent to the eigenfunction-expansion approach. However, there will be an arbitrariness in the operator approach: Which operator are we to choose as our  $\mathbf{L}_1$ ? While in Example 22.5.3 the choice was clear (the operator that had no eigenfunctions), here either operator can be chosen as  $\mathbf{L}_1$ . The ensuing GFs will be equivalent, and the series representing them will be convergent, of course. However, the *rate* of convergence may be different for the two. It turns out, for example, that if we are interested in  $G(x, y; x', y')$  for the two-dimensional Laplacian at points  $(x, y)$  whose  $y$ -coordinates are far from  $y'$ , then the appropriate expansion is obtained by letting  $\mathbf{L}_1 = \partial^2/\partial y^2$ , that is, an expansion in terms of  $x$  eigenfunctions. On the other hand, if the Green's function is to be calculated for a point  $(x, y)$  whose  $x$ -coordinate is far away from the singular point  $(x', y')$ , then the appropriate expansion is obtained by letting  $\mathbf{L}_1 = \partial^2/\partial x^2$ .

## 22.6 Problems

**22.1.** Find the GF for the Dirichlet BVP in two dimensions if  $D$  is the UHP and  $\partial D$  is the  $x$ -axis.

**22.2.** Add  $f(r'')$  to  $H(\mathbf{r}, r'')$  in Example 22.1.2 and retrace the argument given there to show that  $f(r'') = 0$ .

**22.3.** Use the method of images to find the GF for the Laplacian in the exterior region of a "sphere" of radius  $a$  in two and three dimensions.

**22.4.** Derive Equation (22.7) from Equation (22.6).

**22.5.** Using Equation (22.7) with  $\rho = 0$ , show that if  $g(\theta', \varphi') = V_0$ , the potential at any point inside the sphere is  $V_0$ .

**22.6.** Find the BC that the GF must satisfy in order for the solution  $u$  to be representable in terms of the GF when the BC on  $u$  is mixed, as in Equation (22.10). Assume a self-adjoint SOLPDO of the elliptic type, and consider the two cases  $\alpha(\mathbf{x}) \neq 0$  and  $\beta(\mathbf{x}) \neq 0$  for  $\mathbf{x} \in \partial D$ . Hint: In each case, divide the mixed BC equation by the nonzero coefficient, substitute the result in the Green's identity, and set the coefficient of the  $u$  term in the  $\partial D$  integral equal to zero.

**22.7.** Show that the diffusion operator satisfies

$$\mathbf{L}_{\mathbf{x},t} G(\mathbf{x}, \mathbf{y}; t - \tau) = \delta(\mathbf{x} - \mathbf{y})\delta(t - \tau).$$

Hint: Use

$$\frac{\partial \theta}{\partial t}(t - \tau) = \delta(t - \tau).$$

**22.8.** Show that for  $m = 3$  the expression for  $G_s(\mathbf{r})$  given by Equation (22.36) reduces to  $G_s(\mathbf{r}) = -e^{-\mu r}/(4\pi r)$ .

**22.9.** The time-independent Schrödinger equation can be rewritten as

$$(\nabla^2 + k^2)\Psi - \frac{2\mu}{\hbar^2}V(\mathbf{r})\Psi = 0,$$

where  $k^2 = 2\mu E/\hbar^2$  and  $\mu$  is the mass of the particle.

(a) Use techniques of Section 21.4 to write an integral equation for  $\Psi$ .

(b) Show that the Neumann series solution of the integral equation converges only if

$$\int_{\mathbb{R}^3} |V(\mathbf{r})|^2 d^3r < \frac{2\pi\hbar^4 \operatorname{Im} k}{\mu^2}.$$

(c) Assume that the potential is of Yukawa type:  $V(\mathbf{r}) = g^2 e^{-\kappa r}/r$ . Find a condition between the (bound state) energy and the potential strength  $g$  that ensures convergence of the Neumann series.

**22.10.** Derive Equation (22.29).

**22.11.** Derive Equation (22.31) using the procedure outlined for parabolic equations.

**22.12.** (a) Show that the GF for the Helmholtz operator  $\nabla^2 + \mu^2$  in two dimensions is

$$G(\mathbf{r}, \mathbf{r}') = -\frac{i}{4} H_0^{(1)}(\mu|\mathbf{r} - \mathbf{r}'|) + H(\mathbf{r}, \mathbf{r}'),$$

where  $H(\mathbf{r}, \mathbf{r}')$  satisfies the homogeneous Helmholtz equation.

(b) Separate the variables and use the fact that  $H$  is regular at  $\mathbf{r} = \mathbf{r}'$  to show that  $H$  can be written as

$$H(\mathbf{r}, \mathbf{r}') = \sum_{n=0}^{\infty} J_n(\mu r) [a_n(\mathbf{r}') \cos n\theta + b_n(\mathbf{r}') \sin n\theta].$$

(c) Now assume a circular boundary of radius  $a$  and the BC  $G(\mathbf{a}, \mathbf{r}') = 0$ , in which  $\mathbf{a}$  is a vector from the origin to the circular boundary. Using this BC, show that

$$\begin{aligned} a_0(\mathbf{r}') &= \frac{i}{8\pi J_0(\mu a)} \int_0^{2\pi} H_0^{(1)}(\mu\sqrt{a^2 + r'^2 - 2ar' \cos(\theta - \theta')}) d\theta, \\ a_n(\mathbf{r}') &= \frac{i}{4\pi J_n(\mu a)} \int_0^{2\pi} H_0^{(1)}(\mu\sqrt{a^2 + r'^2 - 2ar' \cos(\theta - \theta')}) \cos n\theta d\theta, \\ b_n(\mathbf{r}') &= \frac{i}{4\pi J_n(\mu a)} \int_0^{2\pi} H_0^{(1)}(\mu\sqrt{a^2 + r'^2 - 2ar' \cos(\theta - \theta')}) \sin n\theta d\theta. \end{aligned}$$

These equations completely determine  $H(\mathbf{r}, \mathbf{r}')$  and therefore  $G(\mathbf{r}, \mathbf{r}')$ .

**22.13.** Use the Fourier transform technique to find the singular part of the GF for the diffusion equation in one and three dimensions. Compare your results with that obtained in Section 22.4.3.

**22.14.** Show directly that both  $G_s^{(\text{ret})}$  and  $G_s^{(\text{adv})}$  satisfy  $\nabla^2 G = \delta(\mathbf{r})\delta(t)$  in three dimensions.

**22.15.** Consider a rectangular box with sides  $a$ ,  $b$ , and  $c$  located in the first octant with one corner at the origin. Let  $D$  denote the inside of this box.

(a) Show that zero cannot be an eigenvalue of the Laplacian operator with the Dirichlet BCs on  $\partial D$ .

(b) Find the GF for this Dirichlet BVP.

**22.16.** Find the GF for the Helmholtz equation  $(\nabla^2 + k^2)u = 0$  on the rectangle  $0 \leq x \leq a$ ,  $0 \leq y \leq b$ .

**22.17.** Find the singular part of the one-dimensional GF for the operator  $a d^2/dx^2 + b$ , where  $a > 0$  and  $b < 0$ .

**22.18.** Calculate the GF of the two-dimensional Laplacian operator appropriate for Neumann BCs on the rectangle  $0 \leq x \leq a$ ,  $0 \leq y \leq b$ .

**22.19.** Find the three-dimensional Dirichlet GF for the Helmholtz operator  $\nabla^2 - k^2$  in the half-space  $z \geq 0$ .

**22.20.** Find the three-dimensional Neumann GF for the Helmholtz operator  $\nabla^2 - k^2$  in the half-space  $z \leq 0$ .

**22.21.** Using the integral form of the Schrödinger equation in three dimensions, show that an attractive delta potential  $V(\mathbf{r}) = -V_0\delta(\mathbf{r} - \mathbf{a})$  does not have a bound state ( $E < 0$ ). Contrast this with the result of Example 21.4.1.

**22.22.** By taking the Fourier transform of both sides of the integral form of the Schrödinger equation, show that for bound-state problems ( $E < 0$ ), the equation in "momentum space" can be written as

$$\tilde{\psi}(\mathbf{p}) = -\frac{2\mu}{(2\pi)^{3/2}\hbar^2} \left( \frac{1}{\kappa^2 + p^2} \right) \int \tilde{V}(\mathbf{p} - \mathbf{q}) \tilde{\psi}(\mathbf{q}) d^3q,$$

where  $\kappa^2 = -2\mu E/\hbar^2$ .

**22.23.** Write the bound-state Schrödinger integral equation for a non-local potential, noting that  $G(\mathbf{r}, \mathbf{r}') = e^{-\kappa|\mathbf{r}-\mathbf{r}'|}/|\mathbf{r}-\mathbf{r}'|$ , where  $\kappa^2 = -2\mu E/\hbar^2$  and  $\mu$  is the mass of the bound particle. The homogeneous solution is zero, as is always the case with bound states.

(a) Assuming that the potential is of the form  $V(\mathbf{r}, \mathbf{r}') = -g^2U(\mathbf{r})U(\mathbf{r}')$ , show that a solution to the Schrödinger equation exists iff

$$\frac{\mu g^2}{2\pi\hbar^2} \int_{\mathbb{R}^3} d^3r \int_{\mathbb{R}^3} d^3r' \frac{e^{-\kappa|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} U(\mathbf{r})U(\mathbf{r}') = 1. \quad (22.62)$$

(b) Taking  $U(\mathbf{r}) = e^{-\alpha r}/r$ , show that the condition in (22.62) becomes

$$\frac{4\pi\mu g^2}{\alpha\hbar^2} \left[ \frac{1}{(\alpha + \kappa)^2} \right] = 1.$$

(c) Since  $\kappa > 0$ , prove that the equation in (b) has a unique solution only if  $g^2 > \hbar^2\alpha^2/(4\pi\mu)$ , in which case the bound-state energy is

$$E = -\frac{\hbar^2}{2\mu} \left[ \left( \frac{4\pi\mu g^2}{\alpha\hbar^2} \right)^{1/2} - \alpha \right]^2.$$

**22.24.** Repeat calculations in Sections 22.4.1 and 22.4.2 for  $m = 2$ .

**22.25.** In this problem, the dimension  $m$  is three.

(a) Derive the following identities:

$$\nabla^2 \left[ \frac{f(r)}{r} \right] = \frac{\nabla^2 f}{r} - \frac{2}{r^2} \frac{\partial f}{\partial r} + \nabla^2 \left( \frac{1}{r} \right),$$

$$\frac{d\epsilon(t)}{dt} = 2\delta(t), \quad \nabla^2\delta(t \pm r) = \delta''(t \pm r) \pm \frac{2}{r}\delta'(t \pm r),$$

where  $\epsilon(t) = \theta(t) - \theta(-t)$ .

(b) Use the results of (a) to show that the GF [Equation (22.47)] derived from the principal value of the  $\omega$  integration for the wave equation in three dimensions satisfies only the homogeneous PDE. Hint: Use  $\nabla^2(1/r) = 4\pi\delta(\mathbf{r})$ .

**22.26.** Calculate the retarded GF for the wave operator in two dimensions and show that it is equal to  $G_s^{(\text{ret})}(\mathbf{r}, t) = \theta(t)/(2\pi\sqrt{t^2 - r^2})$ . Now use this result to obtain the GF for any even number of dimensions:

$$G_s^{(\text{ret})}(\mathbf{r}, t) = \frac{\theta(t)}{2\pi} \left( -\frac{1}{2\pi r} \frac{\partial}{\partial r} \right)^{n-1} \left[ \frac{1}{\sqrt{t^2 - r^2}} \right] \quad \text{for } n = m/2.$$

**22.27.** (a) Find the singular part of the retarded GF and the advanced GF for the wave equation in three dimensions using Equations (22.48) and (22.49). Hint:  $J_{1/2}(kr) = \sqrt{2/\pi kr} \sin kr$ .

(b) Use (a) and Equation (22.51) to show that

$$\lim_{\epsilon \rightarrow 0} \left\{ \frac{1}{[r^2 + (-it + \epsilon)^2]} - \frac{1}{[r^2 + (it + \epsilon)^2]} \right\} = -\frac{i\pi}{r} [\delta(t+r) - \delta(t-r)].$$

**22.28.** Show that the eigenfunction expansion of the GF for the Dirichlet BVP for the Laplacian operator in two dimensions for which the region of interest is the interior of a circle of radius  $a$  is

$$G(\mathbf{r}, \mathbf{r}') = -\frac{2}{\pi} \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{\epsilon_n J_n(\frac{\rho}{a} x_{nm}) J_n(\frac{\rho'}{a} x_{nm}) \cos n(\varphi - \varphi')}{J_{n+1}^2(x_{nm}) x_{nm}^2},$$

where  $\epsilon_0 = \frac{1}{2}$  and  $\epsilon_n = 1$  for  $n \geq 1$ , and use has been made of Problem 14.39.

**22.29.** (a) Complete the calculations of Example 22.5.4, and

(b) find the GF for the Laplacian with Dirichlet BCs on two concentric spheres of radii  $a$  and  $b$ , with  $a < b$ .

(c) Consider the case where  $a \rightarrow 0$  and  $b \rightarrow \infty$  and compare the result with the singular part of the GF for the Laplacian.

**22.30.** Solve the Dirichlet BVP for the operator  $\nabla^2 - k^2$  in the region  $0 \leq x \leq a$ ,  $0 \leq y \leq b$ ,  $-\infty < z < \infty$ . Hint: Separate the operator into  $L_1$  and  $L_2$ .

**22.31.** Solve the problem of Example 22.5.1 using the separation of operator technique and show that the two results are equivalent.

**22.32.** Use the operator separation technique to calculate the Dirichlet GF for the two-dimensional operator  $\nabla^2 - k^2$  on the rectangle  $0 \leq x \leq a$ ,  $0 \leq y \leq b$ . Also obtain an eigenfunction expansion for this GF.

**22.33.** Use the operator separation technique to find the three-dimensional Dirichlet GF for the Laplacian in a circular cylinder of radius  $a$  and height  $h$ .

**22.34.** Calculate the singular part of the GF for the three-dimensional free Schrödinger operator

$$i\hbar \frac{\partial}{\partial t} - \frac{\hbar^2}{2\mu} \nabla^2.$$

**22.35.** Use the operator separation technique to show that

(a) the GF for the Helmholtz operator  $\nabla^2 + k^2$  in three dimensions is

$$G(\mathbf{r}, \mathbf{r}') = -ik \sum_{l=0}^{\infty} \sum_{m=-l}^l j_l(kr_{<}) h_l(kr_{>}) Y_{lm}(\theta, \varphi) Y_{lm}^*(\theta', \varphi'),$$

where  $r_{<}$  ( $r_{>}$ ) is the smaller (larger) of  $r$  and  $r'$  and  $j_l$  and  $h_l$  are the spherical Bessel and Hankel functions, respectively. No explicit BCs are assumed except that there is regularity at  $r = 0$  and that  $G(\mathbf{r}, \mathbf{r}') \rightarrow 0$  for  $|\mathbf{r}| \rightarrow \infty$ .

(b) Obtain the identity

$$\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} = ik \sum_{l=0}^{\infty} \sum_{m=-l}^l j_l(kr_{<}) h_l(kr_{>}) Y_{lm}(\theta, \varphi) Y_{lm}^*(\theta', \varphi').$$

(c) Derive the plane wave expansion [see Equation (19.26)]

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l i^l j_l(kr) Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi),$$

where  $\theta'$  and  $\varphi'$  are assumed to be the angular coordinates of  $\mathbf{k}$ . Hint: Let  $|\mathbf{r}'| \rightarrow \infty$ , and use

$$|\mathbf{r}-\mathbf{r}'| = (r'^2 + r^2 - 2\mathbf{r}\cdot\mathbf{r}')^{1/2} \rightarrow r' - \frac{\mathbf{r}'\cdot\mathbf{r}}{r'}$$

and the asymptotic formula  $h_l^{(1)}(z) \rightarrow (1/z)e^{i[z+(l+1)(\pi/2)]}$ , valid for large  $z$ .

## Additional Reading

1. Economou, E. *Green's Functions in Quantum Physics*, Springer-Verlag, 1983. Emphasizes applications of Green's function in quantum mechanics, especially solid-state physics.
2. Jackson, J. *Classical Electrodynamics*, 2nd ed., Wiley, 1975. Many examples of Green's function techniques used in electromagnetism, in particular the advanced and retarded Green's functions as used in electromagnetic radiation theory.



3. Roach, G. *Green's Functions*, Van Nostrand, 1970. Treats eigenfunction expansion of Green's functions and gives many examples.



Part VII \_\_\_\_\_

## Groups and Manifolds



## Group Theory

The tale of mathematics and physics has been one of love and hate, of harmony and discord, and of friendship and animosity. From their simultaneous inception in the shape of calculus in the seventeenth century, through an intense and interactive development in the eighteenth and most of the nineteenth century, to an estrangement in the latter part of the nineteenth and the beginning of the twentieth century, mathematics and physics have experienced the best of times and the worst of times. Sometimes, as in the case of calculus, nature dictates a mathematical dialect in which the narrative of physics is to be spoken. Other times, man, building upon that dialect, develops a sophisticated language in which—as in the case of Lagrangian and Hamiltonian interpretation of dynamics—the narrative of physics is set in the most beautiful poetry. But the happiest courtship, and the most exhilarating relationship, takes place when a discovery in physics leads to a development in mathematics that in turn feeds back into a better understanding of physics, leading to new ideas or a new interpretation of existing ideas. Such a state of affairs began in the 1930s with the advent of quantum mechanics, and, after a lull of about 30 years, revived in the late 1960s. We are fortunate to be witnesses to one of the most productive collaborations between the physics and mathematics communities in the history of both.

It is not an exaggeration to say that the single most important catalyst that has facilitated this collaboration is the idea of **symmetry** the study of which is the main topic of the theory of groups, the subject of this chapter. Although group theory, in one form or another, was known to mathematicians as early as the beginning of the nineteenth century, it found its way into physics only after the invention of quantum theory, and in particular, Dirac's interpretation of it in the language of

transformation theory. Eugene Wigner, in his seminal paper<sup>1</sup> of 1939 in which he applied group theoretical ideas to Lorentz transformations, paved the way for the marriage of group theory and quantum mechanics. Today, in every application of quantum theory, be it to atoms, molecules, solids, or elementary particles such as quarks and leptons, group-theoretical techniques are indispensable.

## 23.1 Groups

The prototype of a group is a transformation group, the set of invertible mappings of a set onto itself. Let us elaborate on this. First, we take mappings because they are the most general operations performed between sets. From a physical standpoint, mappings are essential in understanding the symmetries and other basic properties of a theory. For instance, rotations and translations are mappings of space. Second, the mappings ought to be on a single set, because we want to be able to compose any given two mappings. We cannot compose  $f : A \rightarrow B$  and  $g : A \rightarrow B$ , because, by necessity, the domain of the second must be a subset of the image of the first. With three sets, and  $A \xrightarrow{f} B$ ,  $B \xrightarrow{g} C$ , even if the composition  $f \circ g$  is defined,  $g \circ f$  will not be. Third, we want to be able to undo the mapping. Physically, this means that we should be able to retrace our path to our original position in the set. This can happen only if all mappings of interest have an inverse. Finally, we note that composing a mapping with its inverse yields identity. Therefore, the identity map must also be included in the set of mappings.

We shall come back to transformation groups frequently. In fact, almost all groups considered in this book are transformation groups. However, as in our study of vector spaces in Chapter 1, it is convenient to give a general description of (abstract) groups.

Group defined

**23.1.1. Definition.** A *group* is a set  $G$  together with an associative binary operation  $G \times G \rightarrow G$  called **multiplication**—and denoted generically by  $\star$ —having the following properties:

1. There exists a unique element<sup>2</sup>  $e \in G$  called the **identity** such that  $e \star g = g \star e = g$ .
2. For every element  $g \in G$ , there exists an element  $g^{-1}$ , called the **inverse** of  $g$ , such that  $g \star g^{-1} = g^{-1} \star g = e$ .

To emphasize the binary operation of a group, we designate it as  $(G, \star)$ .

order of a group

If the underlying set  $G$  has a finite number of elements, the group is called **finite**, and its number of elements, denoted by  $|G|$ , is called the **order** of  $G$ . We can also have an infinite group whose cardinality can be countable or continuous.

<sup>1</sup>E. P. Wigner, "On the Unitary Representations of the Inhomogeneous Lorentz Group," *Ann. of Math.* **40** (1939) 149–204.

<sup>2</sup>To distinguish between identities of different groups, we sometimes write  $e_G$  for the identity of the group  $G$ .

Given an element  $a \in G$ , we write

$$a^k \equiv \underbrace{a \star a \star \cdots \star a}_{k \text{ times}}, \quad a^{-m} \equiv \underbrace{a^{-1} \star a^{-1} \star \cdots \star a^{-1}}_{m \text{ times}}$$

and note that

$$a^i \star a^j = a^{i+j} \quad \text{for all } i, j \in \mathbb{Z}.$$

---

**Évariste Galois** (1811–1832) was definitely not the stereotypically dull mathematician, quietly creating theorems and teaching students. He was a political firebrand whose life ended in a mysterious duel when he was only 21 years old. An ardent republican, he was in the unfortunate position of having Cauchy, an ardent royalist, as the only French mathematician capable of understanding the significance of his work. His professional accomplishments (fewer than 100 pages, much of which was published posthumously) received the attention they deserved many years later. It is truly sad to realize that for decades, work from the man credited with the foundation of group theory were lost to the world of mathematics. Galois's early years were relatively happy. His father, a liberal thinker known for his wit, was director of a boarding school and later mayor of Bourg-la-Reine. Galois's mother took charge of his early education. A stubborn, eccentric woman, she mixed classical culture with a fairly stern religious upbringing. The young Galois entered the College Louis-le-Grand in 1823, but found the harsh discipline imposed by church and political authorities difficult to bear. His interest in mathematics was sparked in class by Vernier, but Galois quickly tired of the elementary character of the material, preferring instead to read the more advanced original works on his own. After a flawed attempt to solve the general fifth-order equation, Galois submitted a paper to the *Académie des Sciences* in which he described the definitive solution with the aid of group theory, of which the young Galois can be considered the creator. However, this strong initial foray into the frontiers of mathematics was accompanied by tragedy and setback. A few weeks after the paper's submission, his father committed suicide, which Galois felt was largely to be blamed on those who politically persecuted his father. A month later the young mathematician failed the entrance examination to the Ecole Polytechnique, largely due to his refusal to answer in the form demanded by the examiner. Galois did gain entrance to a less prestigious school for the preparation of secondary-school teachers. While there he read some of Abel's results (published after Abel's death) and found that they contained some of the results he had submitted to the Academy including the proof of the impossibility of solving quintics. Cauchy, assigned as the judge for Galois's paper, suggested that he revise it in light of this new information. Galois instead wrote an entirely new manuscript and submitted it in competition for the grand prix in mathematics. Tragically, the manuscript was lost on the death of Fourier, who had been assigned to examine it, leaving Galois out of the competition. These events, fueled by a later, unfair dismissal of another of his papers by Poisson, seem to have driven Galois toward political radicalism and rebellion during the renewed turmoil then plaguing France. He was arrested several times



for his agitations, although he continued work on mathematics while in custody. On May 30, 1832, he was wounded in a duel with an unknown adversary, the duel perhaps caused by an unhappy love affair. His funeral three days later sparked riots that raged through Paris in the days that followed.

The delay in recognition of the true scope of Galois's scant but amazing work stemmed partly from the originality of his ideas and the lack of competent local reviewers. Cauchy left France after seeing only the early parts of Galois's work, and much of the rest remained unnoticed until Liouville prepared the later manuscripts for publication a decade after Galois's death. Their true value wasn't appreciated for another two decades. The young mathematician himself added to the difficulty by deliberately making his writing so terse that the "established scientists" for whom he had so much disdain could not understand it. Those fortunate enough to appreciate Galois's work found fertile ground in mathematical research, in such fundamental fields as group theory and modern algebra, for decades to come.

**23.1.2. Example.** The following are examples of familiar sets that have group properties.

- (a) The set  $\mathbb{Z}$  of integers under the binary operation of addition forms a group whose identity element is 0 and the inverse of  $n$  is  $-n$ . This group is countably infinite.
- (b) The set  $\{-1, +1\}$ , under the binary operation of multiplication, forms a group whose identity element is 1 and the inverse of each element is itself. This group is finite.
- (c) The set  $\{-1, +1, -i, +i\}$ , under the binary operation of multiplication, forms a finite group whose identity element is 1.
- (d) The set  $\mathbb{R}$ , under the binary operation of addition, forms a group whose identity element is 0 and the inverse of  $r$  is  $-r$ . This group is uncountably infinite.
- (e) The set  $\mathbb{R}^+$  ( $\mathbb{Q}^+$ ) of positive real (rational) numbers, under the binary operation of multiplication, forms a group whose identity element is 1 and the inverse of  $r$  is  $1/r$ . This group is uncountably (countably) infinite.
- (f) The set  $\mathbb{C}$ , under the binary operation of addition, forms a group whose identity element is 0 and the inverse of  $z$  is  $-z$ . This group is uncountably infinite.
- (g) The uncountably infinite set  $\mathbb{C} - \{0\}$  of all complex numbers except 0, under the binary operation of multiplication, forms a group whose identity element is 1 and the inverse of  $z$  is  $1/z$ .
- (h) The uncountably infinite set  $\mathcal{V}$  of vectors in a vector space, under the binary operation of addition, forms a group whose identity element is the zero vector and the inverse of  $|a\rangle$  is  $-|a\rangle$ .
- (i) The set of invertible  $n \times n$  matrices, under the binary operation of multiplication, forms a group whose identity element is the  $n \times n$  unit matrix and the inverse of  $A$  is  $A^{-1}$ . This group is uncountably infinite.

The reader is urged to verify that each set given above is indeed a group. ■

In general, the elements of a group do not commute. Those groups whose elements do commute are so important that we give them a special name:

abelian groups  
defined

**23.1.3. Definition.** A group  $(G, \star)$  is called *abelian* or *commutative* if  $a \star b = b \star a$  for all  $a, b \in G$ . It is common to denote the binary operation of an abelian group by  $+$ .

All groups of Example 23.1.2 are abelian except the last.



**23.1.4. Example.** Let  $\mathbf{A}$  be a vector potential that gives rise to a magnetic field  $\mathbf{B}$ . The set of transformations of  $\mathbf{A}$  that give rise to the same  $\mathbf{B}$  is an abelian group. In fact, such transformations simply add the gradient of a function to  $\mathbf{A}$ . The reader can check the details. ■

symmetric or  
permutation group

The reader may also verify that the set of invertible mappings  $f : S \rightarrow S$ , i.e., the set of transformations of  $S$ , is indeed a (nonabelian) group. If  $S$  has  $n$  elements, this group is denoted by  $S_n$  and is called the **symmetric group** of  $S$ .  $S_n$  is a nonabelian (unless  $n \leq 2$ ) finite group that has  $n!$  elements. An element  $g$  of  $S_n$  is usually denoted by two rows, the top row being  $S$  itself—usually taken to be  $1, 2, \dots, n$ —and the bottom row its image under  $g$ . For example,  $g \equiv \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$  is an element of  $S_4$  such that  $g(1) = 2, g(2) = 3, g(3) = 4$ , and  $g(4) = 1$ .

Consider two groups, the set of vectors in a plane  $((x, y), +)$  and the set of complex numbers  $(\mathbb{C}, +)$ , both under addition. Although these are two different groups, the difference is superficial. We have seen similar differences in disguise in the context of vector spaces and the notion of isomorphism. The same notion applies to group theory:

homomorphism,  
isomorphism, and  
automorphism

**23.1.5. Definition.** Let  $(G, \star)$  and  $(H, \bullet)$  be groups. A **homomorphism**  $f : G \rightarrow H$  is a map such that

$$f(a \star b) = f(a) \bullet f(b) \quad \forall a, b \in G.$$

An **isomorphism** is a homomorphism that is also a bijection. Two groups are **isomorphic**, denoted by  $G \cong H$ , if there is an isomorphism  $f : G \rightarrow H$ . An isomorphism of a group onto itself is called an **automorphism**.

An immediate consequence of this definition is that  $f(e_G) = e_H$  and  $f(g^{-1}) = [f(g)]^{-1}$  (see Problem 23.9).

trivial  
homomorphism

**23.1.6. Example.** Let  $G$  be any group and  $\{1\}$  the multiplicative group consisting of the single number 1. It is straightforward to show that  $f : G \rightarrow \{1\}$ , given by (the only function available!)  $f(g) = 1$  for all  $g \in G$  is a homomorphism. This homomorphism is called the **trivial** (or sometimes, **symmetric**) homomorphism. ■

general linear group

The establishment of isomorphism  $f : \mathbb{R}^2 \rightarrow \mathbb{C}$  between  $((x, y), +)$ , and  $(\mathbb{C}, +)$  is trivial: Just write  $f(x, y) = x + iy$ . A less trivial isomorphism is the exponential map,  $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}^+, \cdot)$ . The reader may verify that this is a homomorphism (in particular, it maps addition to multiplication) and that it is one-to-one. We have noted that the set of invertible maps of a set forms a group. A very important special case of this is when the set is a vector space  $\mathcal{V}$  and the maps are all *linear*

**23.1.7. Box.** The **general linear group** of a vector space  $\mathcal{V}$ , denoted by  $GL(\mathcal{V})$ , is the set of all invertible endomorphisms of  $\mathcal{V}$ . In particular, when  $\mathcal{V} = \mathbb{C}^n$ , we usually write  $GL(n, \mathbb{C})$  instead of  $GL(\mathbb{C}^n)$  with similar notation for  $\mathbb{R}$ .

group multiplication table

It is sometimes convenient to display a finite group  $G = \{g_i\}_{i=1}^{|G|}$  as a  $|G| \times |G|$  table, called the **group multiplication table**, in which the intersection of the  $i$ th row and  $j$ th column is occupied by  $g_i \star g_j$ . Because of its trivial multiplication, the identity is usually omitted from the table.

## 23.2 Subgroups

It is customary to write  $ab$  instead of  $a \star b$ . We shall adhere to this convention, but restore the  $\star$  as necessary to avoid any possible confusion.

subgroup defined

**23.2.1. Definition.** A subset  $S$  of a group  $G$  is a **subgroup** of  $G$  if it is a group in its own right under the binary operation of  $G$ , i.e., if it contains the inverse of all its elements as well as the product of any pair of its elements.

It follows from this definition that  $e \in S$ . It is also easy to show that the intersection of two subgroups is a subgroup (Problem 23.2).

trivial subgroup

### 23.2.2. Example. EXAMPLES OF SUBGROUPS

1. For any  $G$ , the subset  $\{e\}$ , consisting of the identity alone, is a subgroup of  $G$  called the **trivial subgroup** of  $G$ .
2.  $(\mathbb{Z}, +)$  is a subgroup of  $(\mathbb{R}, +)$ .
3. The set of even integers (but not odd integers) is a subgroup of  $(\mathbb{Z}, +)$ . In fact, the set of all multiples of a positive integer  $m$ , denoted by  $\mathbb{Z}m$ , is a subgroup of  $\mathbb{Z}$ . It turns out that *all* subgroups of  $\mathbb{Z}$  are of this form.
4. The subset of  $GL(n, \mathbb{C})$  consisting of transformations that have unit determinant is a subgroup of  $GL(n, \mathbb{C})$  because the inverse of a transformation with unit determinant also has unit determinant, and the product of two transformations with unit determinants has unit determinant.

special linear group

**23.2.3. Box.** The subgroup of  $GL(n, \mathbb{C})$  consisting of elements having unit determinant is denoted by  $SL(n, \mathbb{C})$  and is called the **special linear group**.

unitary, orthogonal, special unitary, and special orthogonal groups

5. The set of unitary transformations of  $\mathbb{C}^n$ , denoted by  $U(n)$ , is a subgroup of  $GL(n, \mathbb{C})$  because the inverse of a unitary transformation is also unitary and the product of two unitary transformations is unitary.

**23.2.4. Box.** The set of unitary transformations  $U(n)$  is a subgroup of  $GL(n, \mathbb{C})$  and is called the **unitary group**. Similarly, the set of orthogonal transformations of  $\mathbb{R}^n$  is a subgroup of  $GL(n, \mathbb{R})$ . It is denoted by  $O(n)$  and called the **orthogonal group**.

Each of these groups has a special subgroup whose elements have unit determinants. These are denoted by  $SU(n)$  and  $SO(n)$ , and called **special unitary group** and **special orthogonal group**, respectively. The latter is also called the group of **rigid rotations** of  $\mathbb{R}^n$ .

6. Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , and define an inner product on  $\mathbb{R}^n$  by

$$\mathbf{x} \cdot \mathbf{y} = -x_1y_1 - \cdots - x_p y_p + x_{p+1}y_{p+1} + \cdots + x_n y_n.$$

Denote the subset of  $GL(n, \mathbb{R})$  that leaves this inner product invariant by<sup>3</sup>  $O(p, n - p)$ . Then  $O(p, n - p)$  is a subgroup of  $GL(n, \mathbb{R})$ . The set of linear transformations among  $O(p, n - p)$  that have determinant 1 is denoted by  $SO(p, n - p)$ . The special case of  $p = 0$  gives us the orthogonal and special orthogonal groups.<sup>4</sup> When  $n = 4$  and  $p = 3$ , we get the inner product of the special theory of relativity, and  $O(3, 1)$ , the set of Lorentz transformations, is called the **Lorentz group**. If one adds translations of  $\mathbb{R}^4$  to  $O(3, 1)$ , one obtains the **Poincaré group**,  $P(3, 1)$ .

Lorentz and Poincaré groups

symplectic group

7. Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{2n}$ , and  $J$  the  $2n \times 2n$  matrix  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ , where  $1$  is the  $n \times n$  unit matrix. The subset of  $GL(2n, \mathbb{R})$  that leaves  $\mathbf{x}^t J \mathbf{x}$ , called an **antisymmetric bilinear form**, invariant is a subgroup of  $GL(2n, \mathbb{R})$  called the **symplectic group** and denoted by  $Sp(2n, \mathbb{R})$ . As we shall see in Chapter 26, the symplectic group is fundamental in the formal treatment of Hamiltonian mechanics.

conjugate subgroup

8. Let  $S$  be a subgroup of  $G$  and  $g \in G$ . Then it is readily shown that the set

$$g^{-1} S g \equiv \{g^{-1} s g \mid s \in S\}$$

is also a subgroup of  $G$ , called the **subgroup conjugate to  $S$  under  $g$** , or the **subgroup  $g$ -conjugate to  $S$** . ■

When discussing vector spaces, we noted that given any subset of a vector space, one could construct a subspace out of it by taking all possible linear combinations (natural operations of the vector space) of the vectors in the subset. We called the subspace thus obtained the *span of the subset*. The same procedure is applicable in group theory as well. If  $S$  is a subset of a group  $G$ , we can generate a subgroup out of  $S$  by collecting all possible products and inverses (natural operations of the group) of the elements of  $S$ . The reader may verify that the result is indeed a subgroup of  $G$ .

subgroup generated by a subset

**23.2.5. Definition.** Let  $S$  be a subset of a group  $G$ . The **subgroup generated by  $S$** , denoted by  $\langle S \rangle$ , is the union of  $S$  and all inverses and products of the elements of  $S$ .

cyclic subgroup

In the special case for which  $S = \{a\}$ , a single element, we use  $\langle a \rangle$  instead of  $\langle \{a\} \rangle$  and call it the **cyclic subgroup** generated by  $a$ . It is simply the collection of all integer powers of  $a$ .

commutator of group elements

**23.2.6. Definition.** Let  $G$  be a group and  $a, b \in G$ . The **commutator of  $a$  and  $b$** ,

<sup>3</sup>The reader is warned that what we have denoted by  $O(p, n - p)$  is sometimes denoted by other authors by  $O(n - p, p)$  or  $O(n, p)$  or  $O(p, n)$ .

<sup>4</sup>It is customary to write  $O(n)$  and  $SO(n)$  for  $O(0, n)$  and  $SO(0, n)$

denoted by  $[a, b]$ , is

$$[a, b] \equiv aba^{-1}b^{-1}.$$

commutator  
subgroup of a group

The subgroup  $\langle \bigcup_{a,b \in G} [a, b] \rangle$  generated by all commutators of  $G$  is called the **commutator subgroup** of  $G$ . The reader may verify that a group is abelian if and only if its commutator subgroup is the trivial subgroup, i.e., consists of only the identity element.

centralizer of an  
element in  $G$  and the  
center of  $G$

**23.2.7. Definition.** Let  $x \in G$ . The set of elements of  $G$  that commute with  $x$ , denoted by  $C_G(x)$ , is called the **centralizer of  $x$  in  $G$** . The set  $Z(G)$  of elements of a group  $G$  that commute with all elements of  $G$  is called the **center of  $G$** .

**23.2.8. Theorem.**  $C_G(x)$  is a subgroup of  $G$  and  $Z(G)$  is an abelian subgroup of  $G$ . Furthermore,  $G$  is abelian if and only if  $Z(G) = G$ .

*Proof.* Proof is immediate from the definitions. □

kernel of a  
homomorphism

**23.2.9. Definition.** Let  $G$  and  $H$  be groups and let  $f : G \rightarrow H$  be a homomorphism. Define the **kernel of  $f$**  by

$$\ker f \equiv \{x \in G \mid f(x) = e \in H\}.$$

The reader may check that  $\ker f$  is a subgroup of  $G$ , and  $f(G)$  is a subgroup of  $H$ . These are the analogues of the same concepts encountered in vector spaces. In fact, if we treat a vector space as an additive group, with the zero vector as identity, then the above definition coincides with that of linear mappings and vector spaces.

Carrying the analogy further, we recall that given two subspaces  $\mathcal{U}$  and  $\mathcal{W}$  of a vector space  $\mathcal{V}$ , we denote by  $\mathcal{U} + \mathcal{W}$  all vectors of  $\mathcal{V}$  that can be written as the sum of a vector in  $\mathcal{U}$  and a vector in  $\mathcal{W}$ . There is a similar concept in group theory that is sometimes very useful.

**23.2.10. Definition.** Let  $S$  and  $T$  be subsets of a group  $(G, \star)$ . Then one defines the **product of these subsets** as

$$S \star T \equiv \{s \star t \mid s \in S \text{ and } t \in T\}.$$

In particular, if  $T$  consists of a single element  $t$ , then

$$S \star t = \{s \star t \mid s \in S\}.$$

left and right cosets

As usual, we shall drop the  $\star$  and write  $ST$  and  $St$ . If  $S$  is a subgroup, then  $St$  is called a **right coset**<sup>5</sup> of  $S$  in  $G$ . Similarly,  $tS$  is called a **left coset** of  $S$  in  $G$ . In either case,  $t$  is said to **represent the coset**.

<sup>5</sup>Some authors switch our right and left in their definition.

**23.2.11. Example.** Let  $G = \mathbb{R}^3$  treated as an additive abelian group, and let  $S$  be a plane through the origin. Then  $t + S$  is  $S$  if  $t \in S$  (see Problem 23.5); otherwise, it is a plane parallel to  $S$ . In fact,  $t + S$  is simply the translation of all points of  $S$  by  $t$ . ■

**23.2.12. Theorem.** Any two right (left) cosets of a subgroup are either disjoint or identical.

*Proof.* Let  $S$  be a subgroup of  $G$  and suppose that  $x \in Sa \cap Sb$ . Then  $x = s_1a = s_2b$  with  $s_1, s_2 \in S$ . Hence,  $ab^{-1} = s_1^{-1}s_2 \in S$ . By Problem 23.6,  $Sa = Sb$ . The left cosets can be treated in the same way. □

A more “elegant” proof starts by showing that an equivalence relation can be defined by

$$a \bowtie b \iff ab^{-1} \in S$$

and then proving that the equivalence classes of this relation are cosets of  $S$ .

One interpretation of Theorem 23.2.12 is that  $a$  and  $b$  belong to the same right coset of  $S$  if and only if  $ab^{-1} \in S$ . A second interpretation is that a coset can be represented by *any* one of its elements (why?).

All cosets (right or left) of a subgroup  $S$  have the same cardinality as  $S$  itself. This can readily be established by considering the map  $\phi : S \rightarrow Sa$  ( $\phi : S \rightarrow aS$ ) with  $\phi(s) = sa$  ( $\phi(s) = as$ ) and showing that  $\phi$  is bijective.

There are many instances both in physics and in mathematics in which a collection of points of a given set represent a single quantity. For example, it is not simply the set of ratios of integers that comprise the set of rational numbers, but the set of certain collections of such ratios: The rational number  $\frac{1}{2}$  represents  $\frac{1}{2}, \frac{2}{4}, \frac{3}{6}$ , etc. Similarly, a given magnetic field represents an infinitude of vector potentials each differing by a gradient from the others, and a physical state in quantum mechanics is an infinite number of wave functions differing from one another by a phase.

With the set of cosets constructed above, it is natural to ask whether they could be given an algebraic structure. The most natural such structure would clearly be that of a group: Given  $aS$  and  $bS$  define their product as  $abS$ . Would this operation turn the set of (left) cosets into a group? The following argument shows that it will, under an important restriction. It is clear that the identity of such a group would be  $S$  itself. It is equally clear that we should have  $(b^{-1}S)(bS) = S$ , so that  $(b^{-1}Sb)S = S$ . It follows from Problem 23.5 that we must have  $b^{-1}Sb \subset S$  for all  $b \in G$ . Now replace  $b$  with  $b^{-1}$  and note that  $bSb^{-1} \subset S$  as well. Let  $s$  be an arbitrary element of  $S$ . Then  $bsb^{-1} = s'$  for some  $s' \in S$ , and  $s = b^{-1}s'b \in b^{-1}Sb$ . It follows that  $S \subset b^{-1}Sb$  for all  $b \in G$ , and, with the reverse inclusion derived above, that  $S = b^{-1}Sb$ . This motivates the following definition.

normal subgroup  
defined

**23.2.13. Definition.** A subgroup  $N$  of a group  $G$  is called **normal** if  $N = g^{-1}Ng$  (equivalently if  $Ng = gN$ ) for all  $g \in G$ .

The preceding argument shows that the set of cosets (no specification is necessary since the right and left cosets coincide) of a normal subgroup forms a group:

quotient group **23.2.14. Theorem.** *If  $N$  is a normal subgroup of  $G$ , then the collection of all cosets of  $N$ , denoted by  $G/N$ , is a group, called the **quotient group** of  $G$  by  $N$ .*

We note that all subgroups of an abelian group are automatically normal, and that the only subgroup conjugate to a normal subgroup  $N$  is  $N$  itself (see Example 23.2.2).

**23.2.15. Example.** Let  $G = \mathbb{R}^3$  and let  $S$  be a plane through the origin as in Example 23.2.11. Since  $G$  is abelian,  $S$  is automatically normal, and  $G/S$  is the set of planes parallel to  $S$ . Let  $\hat{e}_n$  be a normal to  $S$ . Then it is readily seen that

$$G/S = \{r\hat{e}_n + S \mid r \in \mathbb{R}\}.$$

We have picked the perpendicular distance between a plane and  $S$  (with sign included) to represent that plane. The reader may check that the quotient group  $G/S$  is isomorphic to  $\mathbb{R}$ . Identifying  $S$  with  $\mathbb{R}^2$ , we can write  $\mathbb{R}^3/\mathbb{R}^2 \cong \mathbb{R}$ . The cancellation of exponents is quite accidental here!

Let  $G = \mathbb{Z}$  and  $S = \mathbb{Z}m$ , the set of multiples of the positive integer  $m$ . Since  $\mathbb{Z}$  is abelian,  $\mathbb{Z}m$  is normal, and  $\mathbb{Z}/\mathbb{Z}m$  is indeed a group, a typical element of which looks like  $k + m\mathbb{Z}$ . By adding (or subtracting) multiples of  $m$  to  $k$ , and using  $mj + m\mathbb{Z} = m\mathbb{Z}$  (see Problem 23.5), we can assume that  $0 \leq k < m$ . It follows that  $\mathbb{Z}/\mathbb{Z}m$  is a finite group. Furthermore,

$$(k_1 + m\mathbb{Z}) + (k_2 + m\mathbb{Z}) = k_1 + k_2 + m\mathbb{Z} = k + m\mathbb{Z},$$

where  $k$  is the remainder after enough multiples of  $m$  have been subtracted from  $k_1 + k_2$ . One writes  $k_1 + k_2 \equiv k \pmod{m}$ . The coset  $k + m\mathbb{Z}$  is sometimes denoted by  $\bar{k}$  and the quotient group  $\mathbb{Z}/\mathbb{Z}m$  by  $\mathbb{Z}_m$ :

$$\mathbb{Z}_m = \{\bar{0}, \bar{1}, \bar{2}, \dots, \overline{m-1}\}.$$

$\mathbb{Z}_m$  is a prototype of the finite cyclic groups. It can be shown that every cyclic group of order  $m$  is isomorphic to  $\mathbb{Z}_m$  a generator of which is  $\bar{1}$  (recall that the binary operation is addition for  $\mathbb{Z}_m$ ). ■

first isomorphism theorem **23.2.16. Theorem.** (first isomorphism theorem) *Let  $G$  and  $H$  be groups and  $f : G \rightarrow H$  a homomorphism. Then  $\ker f$  is a normal subgroup of  $G$ , and  $G/\ker f$  is isomorphic to  $f(G)$ .*

*Proof.* We have already seen that  $\ker f$  is a subgroup of  $G$ . To show that it is normal, let  $g \in G$  and  $x \in \ker f$ . Then

$$\begin{aligned} f(gxg^{-1}) &= f(g)f(x)f(g^{-1}) = f(g)e_H f(g^{-1}) = f(g)f(g^{-1}) \\ &= f(gg^{-1}) = f(e_G) = e_H. \end{aligned}$$

It follows that  $g x g^{-1} \in \ker f$ . Therefore,  $\ker f$  is normal. We leave it to the reader to show that  $\phi : G/\ker f \rightarrow f(G)$  given by  $\phi(g[\ker f]) \equiv \phi([\ker f]g) = f(g)$  is an isomorphism.<sup>6</sup>  $\square$

**23.2.17. Example.** The special linear group of  $\mathcal{V}$  is a normal subgroup of the general linear group of  $\mathcal{V}$ . To see this, note that  $\det : GL(\mathcal{V}) \rightarrow \mathbb{R}^+$  is a homomorphism whose kernel is  $SL(\mathcal{V})$ .  $\blacksquare$

conjugate and  
conjugacy class  
defined

**23.2.18. Definition.** Let  $x \in G$ . A *conjugate* of  $x$  is an element  $y$  of  $G$  that can be written as  $y = g x g^{-1}$  with  $g \in G$ . The set of all elements of  $G$  conjugate to one another is called a *conjugacy class*. The  $i$ th conjugacy class is denoted by  $K_i$ .

One can check that “ $x$  is conjugate to  $y$ ” is an equivalence relation whose classes are the conjugacy classes. In particular, two different conjugacy classes are disjoint. One can also show that each element of the center of a group constitutes a class by itself. In particular, the identity in any group is in a class by itself, and each element of an abelian group forms a (different) class.

Although a normal subgroup  $N$  contains the conjugate of each of its elements,  $N$  is not a class. The class containing any given element of  $N$  will be only a proper subset of  $N$  (unless  $N$  is trivial). The characteristic feature of a normal subgroup is that it contains the conjugacy classes of all its elements. This is not shared by other subgroups, which, in general, contain only the trivial class of the identity element.

**23.2.19. Example.** Consider the group  $SO(3)$  of rotations in three dimensions. Let us denote a rotation by  $R_{\hat{e}}(\theta)$ , where  $\hat{e}$  is the direction of the axis of rotation and  $\theta$  is the angle of rotation. A typical member of the conjugacy class of  $R_{\hat{e}}(\theta)$  is  $RR_{\hat{e}}(\theta)R^{-1}$ , where  $R$  is some rotation. Let  $\hat{e}' = R\hat{e}$  be the vector obtained by applying the rotation  $R$  on  $\hat{e}$ , and note that

$$RR_{\hat{e}}(\theta)R^{-1}\hat{e}' = RR_{\hat{e}}(\theta)R^{-1}R\hat{e} = RR_{\hat{e}}\hat{e} = R\hat{e} = \hat{e}',$$

where we used the fact that  $R_{\hat{e}}\hat{e} = \hat{e}$  because a rotation leaves its axis unchanged. This last statement, applied to the equation above, also shows that  $RR_{\hat{e}}(\theta)R^{-1}$  is a rotation about  $\hat{e}'$ . Problem 23.18 establishes that the angle of rotation associated with  $RR_{\hat{e}}(\theta)R^{-1}$  is  $\theta$ . We can summarize this as  $RR_{\hat{e}}(\theta)R^{-1} = R_{\hat{e}'}(\theta)$ . It follows that all rotations having the same angle belong to the same conjugacy class of the group of rotations in three dimensions.  $\blacksquare$

all rotations having  
the same angle  
belong to the same  
conjugacy class

### 23.2.1 Direct Products

The resolution of a vector space into a direct sum of subspaces was a useful tool in revealing its structure. The same idea can also be helpful in studying groups. Recall that the only vector common to the subspaces of a direct sum is the zero vector. Moreover, any vector of the whole space can be written as the sum of vectors taken

<sup>6</sup>Compare this theorem with the set-theoretic result obtained in Chapter 0 where the map  $X/\simeq \rightarrow f(X)$  was shown to be bijective if  $\simeq$  is the equivalence relation induced by  $f$ .

from the subspaces of the direct sum. Considering a vector space as a (abelian) group, with zero as the identity and summation as the group operation, leads to the notion of direct product.

internal direct  
product of groups

**23.2.20. Definition.** A group  $G$  is said to be the **direct product** of two of its subgroups  $H_1$  and  $H_2$ , and we write  $G = H_1 \times H_2$ , if

1. all elements of  $H_1$  commute with all elements of  $H_2$ ;
2. the group identity is the only element common to both  $H_1$  and  $H_2$ ;
3. every  $g \in G$  can be written as  $g = h_1 h_2$  with  $h_1 \in H_1$  and  $h_2 \in H_2$ .

It follows from this definition that  $h_1$  and  $h_2$  are unique, and  $H_1$  and  $H_2$  are normal. This kind of direct product is sometimes called **internal**, because the “factors”  $H_1$  and  $H_2$  are chosen from inside the group  $G$  itself. The external direct product results when we take two unrelated groups and make a group out of them:

external direct  
product of groups

**23.2.21. Proposition.** Let  $G$  and  $H$  be groups. The Cartesian product  $G \times H$ , called the **external direct product** of  $G$  and  $H$ , can be given a group structure by

$$(g, h) \star (g', h') \equiv (gg', hh').$$

Furthermore,  $G \cong G \times \{e_H\}$ ,  $H \cong \{e_G\} \times H$ ,  $G \times H \cong H \times G$ , and to within these isomorphisms,  $G \times H$  is the internal direct product of  $G \times \{e_H\}$  and  $\{e_G\} \times H$ .

The proof is left for the reader.

---

**Niels Henrik Abel** (1802–1829) was the second of seven children, son of a Lutheran minister with a small parish of Norwegian coastal islands. In school he received only average marks at first, but then his mathematics teacher was replaced by a man only seven years older than Abel. Abel’s alcoholic father died in 1820, leaving almost no financial support for his young prodigy, who became responsible for supporting his mother and family. His teacher, Holmboe, recognizing his talent for mathematics, raised money from his colleagues to enable Abel to attend Christiania (modern Oslo) University. He entered the university in 1821, 10 years after the university was founded, and soon proved himself worthy of his teacher’s accolades. His second paper, for example, contained the first example of a solution to an integral equation.



Abel then received a two-year government travel grant and journeyed to Berlin, where he met the prominent mathematician Crelle, who soon launched what was to become the leading German mathematical journal of the nineteenth century, commonly called *Crelle’s Journal*. From the start, Abel contributed important papers to Crelle’s Journal, including a classic paper on power series, the scope of which clearly reflects his desire for stringency. His most important work, also published in that journal, was a lengthy treatment of elliptic functions in



which Abel incorporated their inverse functions to show that they are a natural generalization of the trigonometric functions. In later research in this area, Abel found himself in stiff competition with another young mathematician, K. G. J. Jacobi. Abel published some papers on functional equations and integrals in 1823. In it he gives the first solution of an integral equation. In 1824 he proved the impossibility of solving algebraically the general equation of the fifth degree and published it at his own expense hoping to obtain recognition for his work.

Despite his proven intellectual success, Abel never achieved material success, not even a permanent academic position. In December of 1828, while traveling by sled to visit his fiancé for Christmas, Abel became seriously ill and died a couple of months later. Ironically, his death from tuberculosis occurred two days before Crelle wrote with the happy news of an appointment for Abel at a scientific institute in Berlin. In Abel's eulogy in his journal, Crelle wrote:

"He distinguished himself equally by the purity and nobility of his character and by a rare modesty which made his person cherished to the same degree as was his genius."

### 23.3 Group Action

The transformation groups introduced at the beginning of this chapter can be described in the language of abstract groups.

**23.3.1. Definition.** Let  $G$  be a group and  $M$  a set. The *left action* of  $G$  on  $M$  is a map  $\Phi : G \times M \rightarrow M$  such that

1.  $\Phi(e, m) = m$  for all  $m \in M$ ;
2.  $\Phi(g_1 g_2, m) = \Phi(g_1, \Phi(g_2, m))$ .

One usually denotes  $\Phi(g, m)$  by  $g \cdot m$  or more simply by  $gm$ . The *right action* is defined similarly. A subset  $N \subset M$  is called *left (right) invariant* if  $g \cdot m \in N$  ( $m \cdot g \in N$ ) for all  $g \in G$ , whenever  $m \in N$ .

**23.3.2. Example.** If we define  $f_g : M \rightarrow M$  by  $f_g(m) \equiv \Phi(g, m) = g \cdot m$ , then  $f_g$  is recognized as a transformation of  $M$ . The collection of such transformations is a *subgroup* of the set of all transformations of  $M$ . Indeed, the identity transformation is simply  $f_e$ , the inverse of  $f_g$  is  $f_{g^{-1}}$ , and the (associative) law of composition is  $f_{g_1} \circ f_{g_2} = f_{g_1 g_2}$ . There is a general theorem in group theory stating that any group is isomorphic to a subgroup of the group of transformations of an appropriate set. ■

**23.3.3. Definition.** Let  $G$  act on  $M$  and let  $m_0 \in M$ . The *orbit* of  $m_0$ , denoted by  $Gm_0$ , is

$$Gm_0 = \{m \in M \mid m = gm_0 \text{ for some } g \in G\}.$$

The action is called *transitive* if  $Gm_0 = M$ . The *stabilizer* of  $m_0$  is  $G_{m_0} = \{g \in G \mid gm_0 = m_0\}$ . The group action is called *effective* if  $gm = m$  for all  $m \in M$  implies that  $g = e$ .

left action, right action, left invariance and right invariance

any group is isomorphic to a subgroup of the group of transformations of an appropriate set

orbit, stabilizer; transitive action and effective action

Stabilizer is a subgroup.

The reader may verify that the orbit  $Gm_0$  is the smallest invariant subset of  $M$  containing  $m_0$ , and that

**23.3.4. Box.** *The stabilizer of  $m_0$  is a subgroup of  $G$ , which is sometimes called the little group of  $G$  at  $m_0$ .*

A transitive action is characterized by the fact that given any two points  $m_1, m_2 \in M$ , one can find a  $g \in G$  such that  $m_2 = gm_1$ .

**23.3.5. Example.** Let  $M = \mathbb{R}^2$  and  $G = SO(2)$ , the planar rotation group. The action is rotation of a point in the plane about the origin by an angle  $\theta$ . The orbits are circles centered at the origin. The action is effective but not transitive. The stabilizer of every point in the plane is  $\{e\}$ , except the origin, for which the whole group is the stabilizer.

Let  $M = S^1$ , the unit circle, and  $G = SO(2)$ , the rotation group in two dimensions. The action is displacement of a point on the circle. There is only one orbit, the entire circle. The action is effective and transitive. The stabilizer of every point on the circle is  $\{e\}$ .

Let  $M = G$ , a group, and let a (proper) subgroup  $S$  act on  $G$  by left multiplication. The orbits are right cosets  $Sg$  of the subgroup. The action is effective but not transitive. The stabilizer of every point in the group is  $\{e\}$ .

Let  $M = \mathbb{R} \cup \{\infty\}$ , the set of real numbers including "the point at infinity." Define an action of  $SL(2, \mathbb{R})$  on  $M$  by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot x = \frac{ax + c}{bx + d}.$$

The reader may check that this is indeed a group action with a law of multiplication identical to the matrix multiplication, and that the action is transitive, but not effective.

realization of a group

Let  $M$  be a set and  $H$  the group of transformations of  $M$ . Suppose that there is a homomorphism  $f : G \rightarrow H$  from a group  $G$  into  $H$ . Then there is a natural action of  $G$  on  $M$  given by  $g \cdot m \equiv [f(g)](m)$ . The homomorphism  $f$  is sometimes called a **realization** of  $G$ . ■

## 23.4 The Symmetric Group $S_n$

Because of its primary importance as the prototypical finite group, and because of its significance in quantum statistics, the symmetric (or permutation) group is briefly discussed in this section. It is also used extensively in the theory of representation of the general linear group and its subgroups.

A generic permutation  $\pi$  of  $n$  numbers is shown as

$$\pi = \begin{pmatrix} 1 & 2 & \dots & i & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(i) & \dots & \pi(n) \end{pmatrix}. \tag{23.1}$$

Because the mapping is bijective, no two elements can have the same image, and  $\pi(1), \pi(2), \dots, \pi(n)$  exhaust all the elements in the set  $\{i\}_{i=1}^n$ .

$e$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$
$\pi_2$	$e$	$\pi_5$	$\pi_6$	$\pi_3$	$\pi_4$
$\pi_3$	$\pi_6$	$e$	$\pi_5$	$\pi_4$	$\pi_2$
$\pi_4$	$\pi_5$	$\pi_6$	$e$	$\pi_2$	$\pi_3$
$\pi_5$	$\pi_4$	$\pi_2$	$\pi_3$	$\pi_6$	$e$
$\pi_6$	$\pi_3$	$\pi_4$	$\pi_2$	$e$	$\pi_5$

**Table 23.1** Group multiplication table for  $S_3$ .

We can display the product  $\pi_2 \circ \pi_1$  of two permutations using  $\pi_2 \circ \pi_1(i) \equiv \pi_2(\pi_1(i))$ . For instance, if

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}, \quad (23.2)$$

then the product  $\pi_2 \circ \pi_1$  takes 1 to 3, etc., because  $\pi_2 \circ \pi_1(1) \equiv \pi_2(\pi_1(1)) = \pi_2(3) = 3$ , etc. We display  $\pi_2 \circ \pi_1$  as

$$\pi_2 \circ \pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}.$$

**23.4.1. Example.** Let us construct the multiplication table for  $S_3$ . Denote the elements of  $S_3$  as follows:

$$\begin{aligned} e &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, & \pi_2 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, & \pi_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \\ \pi_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, & \pi_5 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, & \pi_6 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}. \end{aligned}$$

We give only one sample evaluation of the entries and leave the straightforward—but instructive—calculation of the other entries to the reader. Consider  $\pi_4 \circ \pi_5$ , and note that  $\pi_5(1) = 3$  and  $\pi_4(3) = 2$ ; so  $\pi_4 \circ \pi_5(1) = 2$ . Similarly,  $\pi_4 \circ \pi_5(2) = 1$  and  $\pi_4 \circ \pi_5(3) = 3$ . Thus

$$\pi_4 \circ \pi_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \pi_2.$$

The entire multiplication table is given in Table 23.1. ■

Note that both the rows and columns of the group multiplication table include all elements of the group, and no element is repeated in a row or a column. This is because left-multiplication of elements of a group by a single fixed element of the group simply permutes the group elements. Stated differently, the left multiplication map  $L_g : G \rightarrow G$ , given by  $L_g(x) = gx$ , is bijective, as the reader may verify.

Because we are dealing with finite numbers, repeated application of a permutation to an integer in the set  $\{i\}_{i=1}^n$  eventually produces the initial integer. This leads to the following definition.

cycles of symmetric group

**23.4.2. Definition.** Let  $\pi \in S_n, i \in \{1, 2, \dots, n\}$ , and let  $r$  be the smallest positive integer such that  $\pi^r(i) = i$ . Then the set of  $r$  distinct elements  $\{\pi^k(i)\}_{k=0}^{r-1}$  is called a cycle of  $\pi$  of length  $r$  or an  $r$ -cycle generated by  $i$ .

Start with 1 and apply  $\pi$  to it repeatedly until you obtain 1 again. The collection of elements so obtained forms a cycle in which 1 is contained. Then we select a second number that is not in this cycle and apply  $\pi$  to it repeatedly until the original number is obtained again. Continuing in this way, we produce a set of disjoint cycles that exhausts all elements of  $\{1, 2, \dots, n\}$ .

**23.4.3. Proposition.** Any permutation can be broken up into disjoint cycles.

It is customary to write elements of each cycle in some specific order within parentheses starting with the first element, say  $i$ , on the left, then  $\pi(i)$  immediately to its right, followed by  $\pi^2(i)$ , and so on. For example, the permutations  $\pi_1$  and  $\pi_2$  of Equation (23.2) and their product have the cycle structures  $\pi_1 = (13)(24)$ ,  $\pi_2 = (124)(3)$ , and  $\pi_2 \circ \pi_1 = (132)(4)$ , respectively.

**23.4.4. Example.** Let  $\pi_1, \pi_2 \in S_8$  be given by

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 5 & 7 & 1 & 2 & 8 & 4 & 6 \end{pmatrix}, \quad \pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 5 & 6 & 8 & 1 & 7 & 4 & 3 \end{pmatrix}.$$

The reader may verify that

$$\pi_2 \circ \pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 1 & 4 & 2 & 5 & 3 & 8 & 7 \end{pmatrix}$$

and that

$$\pi_1 = (1374)(25)(68), \quad \pi_2 = (125)(36748), \quad \pi_2 \circ \pi_1 = (16342)(5)(78).$$

In general, permutations do not commute. The product in reverse order is

$$\pi_1 \circ \pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 2 & 8 & 6 & 3 & 4 & 1 & 7 \end{pmatrix} = (15387)(2)(46),$$

which differs from  $\pi_2 \circ \pi_1$ . However, note that it has the same cycle structure as  $\pi_2 \circ \pi_1$ , in that cycles of equal length appear in both. This is a general property of all permutations. ■

cyclic permutations defined

**23.4.5. Definition.** If  $\pi \in S_n$  has a cycle of length  $r$  and all other cycles of  $\pi$  have only one element, then  $\pi$  is called a cyclic permutation of length  $r$ .

It follows that  $\pi_2 \in S_4$  as defined earlier is a cyclic permutation of length 3. Similarly,

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 2 & 1 & 3 & 5 & 4 \end{pmatrix}$$

is a cyclic permutation of length 4 (verify this).

transpositions  
defined

**23.4.6. Definition.** A cyclic permutation of length 2 is called a *transposition*.

A transposition  $(ij)$  simply switches  $i$  and  $j$ .

**23.4.7. Example.** Products of (not necessarily disjoint) cycles may be associated with a permutation whose action on  $i$  is obtained by starting with the first cycle (at the extreme right), locating the first occurrence of  $i$ , and keeping track of what each cycle does to it or its image under the preceding cycle. For example, let  $\pi_1 \in S_6$  be given as a product of cycles by  $\pi_1 = (143)(24)(456)$ . To find the permutation, we start with 1 and follow the action of the cycles on it, starting from the right. The first and second cycles leave 1 alone, and the last cycle takes it to 4. Thus,  $\pi_1(1) = 4$ . For 2 we note that the first cycle leaves it alone, the second cycle takes it to 4, and the last cycle takes 4 to 3. Thus,  $\pi_1(2) = 3$ . Similarly,  $\pi_1(3) = 1, \pi_1(4) = 5, \pi_1(5) = 6$ , and  $\pi_1(6) = 2$ . Therefore,

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 1 & 5 & 6 & 2 \end{pmatrix}.$$

We note that  $\pi_1$  is a cyclic permutation of length 6.

It is left to the reader to show that the permutation  $\pi_2 \in S_5$  given by the product  $\pi_2 = (13)(15)(12)(14)$  is cyclic:  $\pi_2 = (14253)$ . ■

The square of any transposition is the identity. Therefore, we can include it in any product of permutations without changing anything.

**23.4.8. Proposition.** An  $r$ -cycle  $(i_1, i_2, \dots, i_r)$  can be decomposed into the product of  $r - 1$  transpositions:

$$(i_1, i_2, \dots, i_r) = (i_1 i_r)(i_1 i_{r-1}) \cdots (i_1 i_3)(i_1 i_2).$$

*Proof.* The proof involves keeping track of what happens to each symbol when acted upon by the RHS and the LHS and showing that the two give the same result. This is left as an exercise for the reader. □

parity of a  
permutation defined

Although the decomposition of Proposition 23.4.8 is not unique, it can be shown that the **parity** of the decomposition (whether the number of factors is even or odd) is unique. For instance, it is easy to verify that

$$(1234) = (14)(13)(12) = (14) \underbrace{(34)(34)}_1 \underbrace{(23)(12)(23)}_1 (13)(12).$$

That is,  $(1234)$  is written as a product of 3 or 9 transpositions, both of which are odd.

We have already seen that any permutation can be written as a product of cycles. In addition, Proposition 23.4.8 says that these cycles can be further broken down into products of transpositions. This implies the following (see [Rotm 84, p. 38]):

parity of a  
permutation is  
unique

**23.4.9. Proposition.** Any permutation can be decomposed as a product of transpositions. The parity of the decomposition is unique.

even and odd permutations

**23.4.10. Definition.** A permutation is *even (odd)* if it can be expressed as a product of an even (odd) number of transpositions.

The parity of a permutation can be determined from its cycle structure and Proposition 23.4.8.

The reader may verify that the mapping from  $S_n$  to the multiplicative group of  $\{+1, -1\}$  that assigns  $+1$  to even permutations and  $-1$  to odd permutations is a group homomorphism. It follows from the first isomorphism theorem (Theorem 23.2.16) that

**23.4.11. Box.** The set of even permutations, denoted by  $A_n$ , forms a normal subgroup of  $S_n$ .

This homomorphism is usually denoted by  $\epsilon$ . We therefore define

$$\epsilon(\pi) \equiv \epsilon_\pi = \begin{cases} +1 & \text{if } \pi \text{ is even,} \\ -1 & \text{if } \pi \text{ is odd.} \end{cases} \tag{23.3}$$

Sometimes  $\delta(\pi)$  or  $\delta_\pi$  as well as  $\text{sgn}(\pi)$  is also used. The symbol,  $\epsilon_{i_1 i_2 \dots i_n}$  used in the definition of determinants, is closely related to  $\epsilon(\pi)$ . In fact,

$$\epsilon_{\pi(1)\pi(2)\dots\pi(n)} \equiv \epsilon_\pi.$$

Suppose  $\pi, \sigma \in S_n$ , and note that<sup>7</sup>  $\sigma(i) \xrightarrow{\sigma^{-1}} i \xrightarrow{\pi} \pi(i) \xrightarrow{\sigma} \sigma \circ \pi(i)$ , i.e., the composite  $\sigma \circ \pi \circ \sigma^{-1}$  of the three permutations takes  $\sigma(i)$  to  $\sigma \circ \pi(i)$ . This composite can be thought of as the permutation obtained by applying  $\sigma$  to the two rows of  $\pi = (\begin{smallmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{smallmatrix})$ :

$$\sigma \circ \pi \circ \sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ \sigma \circ \pi(1) & \sigma \circ \pi(2) & \dots & \sigma \circ \pi(n) \end{pmatrix}.$$

In particular, the cycles of  $\sigma \circ \pi \circ \sigma^{-1}$  are obtained by applying  $\sigma$  to the symbols in the cycles of  $\pi$ . Since  $\sigma$  is bijective, the cycles so obtained will remain disjoint. It follows that  $\sigma \circ \pi \circ \sigma^{-1}$ , a conjugate of  $\pi$ , has the same cycle structure as  $\pi$  itself. In fact, we have the following:

**23.4.12. Theorem.** Two permutations are conjugate if and only if they have the same cycle structure.

To find the distinct conjugacy classes of  $S_n$ , one has to construct distinct cycle structures of  $S_n$ . This in turn is equivalent to partitioning the numbers from 1 to  $n$

<sup>7</sup>Recall from Chapter 0 that  $x \xrightarrow{f} y$  means  $y = f(x)$ .

into sets of various lengths. Let  $\nu_k$  be the number of  $k$ -cycles in a permutation. The cycle structure of this permutation is denoted by  $(1^{\nu_1}, 2^{\nu_2}, \dots, n^{\nu_n})$ . Since the total number of symbols is  $n$ , we must have  $\sum_{k=1}^n k\nu_k = n$ . Defining  $\lambda_j \equiv \sum_{k=j}^n \nu_k$ , we have

$$\lambda_1 + \lambda_2 + \dots + \lambda_n = n, \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0. \tag{23.4}$$

partition of  $n$

The splitting of  $n$  into nonnegative integers  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  as in Equation (23.4) is called a **partition of  $n$** . There is a 1-1 correspondence between partitions of  $n$  and the cycle structure of  $S_n$ . We saw how  $\nu_k$ 's gave rise to  $\lambda$ 's. Conversely, given a partition of  $n$ , we can construct a cycle structure by  $\nu_k = \lambda_k - \lambda_{k+1}$ . For example, the partition (32000) of  $S_5$  corresponds to  $\nu_1 = 3 - 2 = 1$ ,  $\nu_2 = 2 - 0 = 2$ , i.e., one 1-cycle and two 2-cycles. One usually omits the zeros and writes (32) instead of (32000). When some of the  $\lambda$ 's are repeated, the number of occurrences is indicated by a power of the corresponding  $\lambda$ ; the partition is then written as  $(\mu_1^{n_1}, \mu_2^{n_2}, \dots, \mu_r^{n_r})$ , where it is understood that  $\lambda_1$  through  $\lambda_{n_1}$  have the common value  $\mu_1$ , etc. For example,  $(3^21)$  corresponds to a partition of 7 with  $\lambda_1 = 3$ ,  $\lambda_2 = 3$ , and  $\lambda_3 = 1$ . The corresponding cycle structure is  $\nu_1 = 0$ ,  $\nu_2 = 2$ , and  $\nu_3 = 1$ , i.e., two 2-cycles and one 3-cycle. The partitions of length 0 are usually ignored. Since  $\sum \lambda_i = n$ , no confusion will arise as to which symmetric group the partition belongs to. Thus (32000) and (332000) are written as (32) and  $(3^22)$ , and it is clear that (32) belongs to  $S_5$  and  $(3^22)$  to  $S_8$ .

Conjugacy classes of  $S_n$  are related to the partition of  $n$ .

**23.4.13. Example.** Let us find the different cycle structures of  $S_4$ . This corresponds to different partitions of 4. We can take  $\lambda_1 = 4$  and the rest of the  $\lambda$ 's zero. This gives the partition (4). Next, we let  $\lambda_1 = 3$ ; then  $\lambda_2$  must be 1, giving the partition (31). With  $\lambda_1 = 2$ ,  $\lambda_2$  can be either 2 or 1. In the latter case,  $\lambda_3$  must be 1 as well, and we obtain two partitions,  $(2^2)$  and  $(21^2)$ . Finally, if  $\lambda_1 = 1$ , all other nonzero  $\lambda$ 's must be 1 as well (remember that  $\lambda_k \geq \lambda_{k+1}$ ). Therefore, the last partition is of the form  $(1^4)$ . We see that there are 5 different partitions of 4. It follows that *there are 5 different conjugacy classes in  $S_4$* . ■

## 23.5 Problems

**23.1.** Let  $S$  be a subset of a group  $G$ . Show that  $S$  is a subgroup if and only if  $ab^{-1} \in S$  whenever  $a, b \in S$ .

**23.2.** Show that the intersection of two subgroups is a subgroup.

**23.3.** Let  $X$  be a subset of a group  $G$ . A word on  $X$  is either  $e_G$  or an element  $w$  of  $G$  of the form

$$w = x_1^{e_1} x_2^{e_2} \dots x_n^{e_n},$$

where  $x_i \in X$  and  $e_i = \pm 1$ . Show that the set of all words on  $X$  is a subgroup of  $G$ .

**23.4.** Let  $[a, b]$  denote the commutator of  $a$  and  $b$ . Show that

(a)  $[a, b]^{-1} = [b, a]$ ,

(b)  $[a, a] = e$  for all  $a \in G$ , and

(c)  $ab = [a, b]ba$ . It is interesting to compare these relations with the familiar commutators of operators.

**23.5.** Show that if  $S$  is a subgroup, then  $S^2 \equiv SS = S$ , and  $tS = S$  if and only if  $t \in S$ . More generally,  $TS = S$  if and only if  $T \subset S$ .

**23.6.** Show that if  $S$  is a subgroup, then  $Sa = Sb$  if and only if  $ba^{-1} \in S$  and  $ab^{-1} \in S$  ( $aS = bS$  if and only if  $a^{-1}b \in S$  and  $b^{-1}a \in S$ ).

**23.7.** Let  $S$  be a subgroup of  $G$ . Show that  $a \triangleright b$  defined by  $ab^{-1} \in S$  is an equivalence relation.

**23.8.** Show that  $C_G(x)$  is a subgroup of  $G$ . Let  $H$  be a subgroup of  $G$  and suppose  $x \in H$ . Show that  $C_H(x)$  is a subgroup of  $C_G(x)$ .

**23.9.** (a) Show that the only element  $a$  in a group with the property  $a^2 = a$  is the identity. (b) Now use  $e_G \star e_G = e_G$  to show that any homomorphism maps identity to identity. (c) Show that if  $f : G \rightarrow H$  is a homomorphism, then  $f(g^{-1}) = [f(g)]^{-1}$ .

**23.10.** Establish a bijection between the set of right cosets and the set of left cosets of a subgroup. Hint: Define a map that takes  $S_t$  to  $t^{-1}S$ .

Lagrange theorem

**23.11.** Let  $G$  be a finite group and  $S$  one of its subgroups. Convince yourself that the union of all right cosets of  $S$  is  $G$ . Now use the fact that distinct right cosets are disjoint and that they have the same cardinality to prove that the order of  $S$  divides the order of  $G$ . In fact,  $|G| = |S| |G/S|$ , where  $|G/S|$  is the number of cosets of  $S$  (also called the index of  $S$  in  $G$ ). This is **Lagrange's theorem**.

**23.12.** Let  $f : G \rightarrow H$  be a homomorphism. Show that  $\phi : G/\ker f \rightarrow f(G)$  given by  $\phi(g[\ker f]) \equiv \phi([ \ker f ]g) = f(g)$  is an isomorphism.

**23.13.** Let  $G'$  denote the commutator subgroup of a group  $G$ . Show that  $G'$  is a *normal* subgroup of  $G$  and that  $G/G'$  is abelian.

**23.14.** Let  $M = \mathbb{R} \cup \{\infty\}$ , and define an action of  $SL(2, \mathbb{R})$  on  $M$  by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot x = \frac{ax + c}{bx + d}.$$

Show that this is indeed a group action with a law of multiplication identical to the matrix multiplication, and that the action is transitive, but not effective.

**23.15.** Show that two conjugacy classes are either disjoint or identical.

**23.16.** Show that if all conjugacy classes of a group have only one element, the group must be abelian.



**23.17.** Consider a map from the conjugacy class of  $G$  containing  $x \in G$  to the set of (left) cosets  $G/C_G(x)$  given by  $\phi(axa^{-1}) = aC_G(x)$ . Show that  $\phi$  is a bijection. In particular, show that  $|C_G(x)| = |G|/|K_x^G|$  where  $K_x^G$  is the class in  $G$  containing  $x$  and  $|K_x^G|$  its order (see Problem 23.11). Use this result and Problems 23.8 and 23.11 to show that  $|H|/|K_x^H|$  divides  $|G|/|K_x^G|$ .

**23.18.** Show that  $RR_{\hat{e}}(\theta)R^{-1}$  corresponds to a rotation of angle  $\theta$ . Hint: Consider the effect of rotation on the vectors in the plane perpendicular to  $\hat{e}$ , and note that the rotated plane is perpendicular to  $\hat{e}' \equiv R\hat{e}$ .

**23.19.** Let  $G$  act on  $M$  and let  $m_0 \in M$ . Show that  $Gm_0$  is the smallest invariant subset of  $M$  containing  $m_0$ .

**23.20.** Suppose  $G$  is the direct product of  $H_1$  and  $H_2$  and  $g = h_1h_2$ . Show that the factors  $h_1$  and  $h_2$  are unique and that  $H_1$  and  $H_2$  are normal.

**23.21.** Show that  $(g, h), (g', h') \in G \times H$  are conjugate if and only if  $g$  is conjugate to  $g'$  and  $h$  is conjugate to  $h'$ . Therefore, conjugacy classes of the direct product are obtained by pairing one conjugacy class from each factor.

**23.22.** Find the products  $\pi_1 \circ \pi_2$  and  $\pi_2 \circ \pi_1$  of the two permutations

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 6 & 5 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 6 & 5 & 4 \end{pmatrix}.$$

**23.23.** Find the inverses of the permutations

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 5 & 7 & 1 & 2 & 8 & 4 & 6 \end{pmatrix}, \quad \pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 5 & 6 & 8 & 1 & 7 & 4 & 3 \end{pmatrix}$$

and show directly that  $(\pi_1 \circ \pi_2)^{-1} = \pi_2^{-1} \circ \pi_1^{-1}$ .

**23.24.** Find the inverse of each of the following permutations:  $\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$ ,  $\pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 5 & 3 \end{pmatrix}$ ,  $\pi_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix}$ , and  $\pi_4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \end{pmatrix}$ .

**23.25.** Express each of the following products in terms of disjoint cycles. Assume that all permutations are in  $S_7$ .

$$(a) (123)(347)(456)(145). \quad (b) (34)(562)(273). \quad (c) (1345)(134)(13).$$

**23.26.** Express the following permutations as products of disjoint cycles, and determine which are cyclic.

$$(a) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 4 & 5 & 6 & 2 \end{pmatrix}. \quad (b) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 4 & 5 & 6 & 3 \end{pmatrix}. \quad (c) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix}.$$

**23.27.** Express the permutation  $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 4 & 1 & 3 & 6 & 8 & 7 & 5 \end{pmatrix}$  as a product of transpositions. Is the permutation even or odd?

**23.28.** Express the following permutations as products of transpositions, and determine whether they are even or odd.

$$(a) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 1 & 5 \end{pmatrix}.$$

$$(b) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 1 & 7 & 8 & 3 & 6 & 5 & 2 \end{pmatrix}.$$

$$(c) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 5 & 3 & 2 & 1 \end{pmatrix}.$$

$$(d) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 7 & 2 & 4 & 1 & 5 & 3 \end{pmatrix}.$$

**23.29.** Show that the product of two even or two odd permutations is always even, and the product of an even and an odd permutation is always odd.

**23.30.** Show that  $\pi$  and  $\pi^{-1}$  have the same parity (both even or both odd).

**23.31.** Find the number of distinct conjugacy classes of  $S_5$  and  $S_6$ .

## Additional Reading

1. Hamermesh, M. *Group Theory and its Application to Physical Problems*, Dover, 1989. The classic textbook on group theory written specifically for physicists.
2. Rotman, J. *An Introduction to the Theory of Groups*, 3rd ed., Allyn and Bacon, 1984. An excellent (formal, but readable) introduction to group theory with many examples and lots of explanations.
3. Wigner, E. *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*, Academic Press, 1959. Another classic written by the master of group theory himself.

# Group Representation Theory

Group action is extremely important in quantum mechanics. Suppose the Hamiltonian of a quantum system is invariant under a symmetry transformation of its independent parameters such as position, momentum, and time. This invariance will show up as certain properties of the solutions of the Schrödinger equation.

Moreover, the very act of labeling quantum-mechanical states often involves groups and their actions. For example, labeling atomic states by eigenvalues of angular momentum assumes invariance of the Hamiltonian under the action of the rotation group (see Chapter 27) on the Hilbert space of the quantum-mechanical system under consideration.

## 24.1 Definitions and Examples

In the language of group theory, we have the following situation. Put all the parameters  $x_1, \dots, x_p$  of the Hamiltonian  $\mathbf{H}$  together to form a space, say  $\mathbb{R}^p$ , and write  $\mathbf{H} = \mathbf{H}(x_1, \dots, x_p) \equiv \mathbf{H}(\mathbf{x})$ . A **group of symmetry** of  $\mathbf{H}$  is a group  $G$  whose action on  $\mathbb{R}^p$  leaves  $\mathbf{H}$  unchanged,<sup>1</sup> i.e.,  $\mathbf{H}(\mathbf{x} \cdot g) = \mathbf{H}(\mathbf{x})$ . For example, a one-dimensional harmonic oscillator, with  $\mathbf{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2$ , has, among other things, parity  $P$  (defined by  $Px = -x$ ) as a symmetry. Thus, the group  $G = \{e, P\}$  is a group of symmetry of  $\mathbf{H}$ .

The Hamiltonian  $\mathbf{H}$  of a quantum-mechanical system is an operator in a Hilbert space, such as  $\mathcal{L}^2(\mathbb{R}^3)$ , the space of square-integrable functions. The important question is, What is the proper way of transporting the action of  $G$  from  $\mathbb{R}^p$

---

<sup>1</sup>It will become clear shortly that the appropriate direction for the action is from the right.

to  $\mathcal{L}^2(\mathbb{R}^3)$ ? This is a relevant question because the solutions of the Schrödinger equation are, in general, functions of the parameters of the Hamiltonian, and as such will be affected by the symmetry operation on the Hamiltonian. The answer is provided in the following definition.

representation;  
carrier space and  
dimension of a  
representation;  
faithful and identity  
representation

**24.1.1. Definition.** Let  $G$  be a group and  $\mathcal{H}$  a Hilbert space. A **representation of  $G$  on  $\mathcal{H}$**  is a homomorphism  $T : G \rightarrow GL(\mathcal{H})$ . The representation is **faithful** if the homomorphism is 1-1. We often denote  $T(g)$  by  $\mathbf{T}_g$ .  $\mathcal{H}$  is called the **carrier space of  $T$** . The trivial homomorphism  $T : G \rightarrow \{1\}$  is also called the **identity representation**. The dimension of  $\mathcal{H}$  is called the **dimension of the representation  $T$** .

We do not want to distinguish between representations that differ only by isomorphic vector spaces, because otherwise we can generate an infinite set of representations that are trivially related to one another. A *vector space isomorphism*  $f : \mathcal{H} \rightarrow \mathcal{H}'$  induces a *group isomorphism*  $\phi : GL(\mathcal{H}) \rightarrow GL(\mathcal{H}')$  defined by

$$\phi(\mathbf{T}) = f \circ \mathbf{T} \circ f^{-1} \quad \text{for } \mathbf{T} \in GL(\mathcal{H}).$$

This motivates the following definition.

equivalent  
representations

**24.1.2. Definition.** Two representations  $T : G \rightarrow GL(\mathcal{H})$  and  $T' : G \rightarrow GL(\mathcal{H}')$  are called **equivalent** if there exists an isomorphism  $f : \mathcal{H} \rightarrow \mathcal{H}'$  such that  $\mathbf{T}'_g = f \circ \mathbf{T}_g \circ f^{-1}$  for all  $g \in G$ .

**24.1.3. Box.** Any representation  $T : G \rightarrow GL(\mathcal{H})$  defines an action of the group  $G$  on the Hilbert space  $\mathcal{H}$  by  $\Phi(g, |a\rangle) \equiv \mathbf{T}_g |a\rangle$ .

As we saw in Chapters 2 and 3, the transformation of an operator  $\mathbf{A}$  under  $\mathbf{T}_g$  would have to be defined by  $\mathbf{T}_g \mathbf{A} (\mathbf{T}_g)^{-1}$ . For a Hamiltonian with a group of symmetry  $G$ , this leads to the identity

$$\mathbf{T}_g [\mathbf{H}(\mathbf{x})] (\mathbf{T}_g)^{-1} = \mathbf{H}(\mathbf{x} \cdot g).$$

Similarly, the action of the group on a vector (function) in  $\mathcal{L}^2(\mathbb{R}^3)$  is defined by

$$(\mathbf{T}_g \psi)(\mathbf{x}) \equiv \psi(\mathbf{x} \cdot g), \tag{24.1}$$

where the parentheses around  $\mathbf{T}_g \psi$  designate it as a *new* function. One can show that if  $G$  acts on the independent variables of a function *on the right* as in Equation (24.1), then the vector space of such functions is the carrier space of a representation of  $G$ . In fact,

$$(\mathbf{T}_{g_1 g_2} \psi)(\mathbf{x}) \equiv \psi(\mathbf{x} \cdot (g_1 g_2)) = \psi((\mathbf{x} \cdot g_1) \cdot g_2) = (\mathbf{T}_{g_2} \psi)(\mathbf{x} \cdot g_1) \equiv \varphi(\mathbf{x} \cdot g_1),$$

where we have defined the new function  $\varphi$  by the last equality. Now note that

$$\varphi(\mathbf{x} \cdot g_1) = (\mathbf{T}_{g_1} \varphi)(\mathbf{x}) = (\mathbf{T}_{g_1} (\mathbf{T}_{g_2} \psi))(\mathbf{x}) = (\mathbf{T}_{g_1} \mathbf{T}_{g_2} \psi)(\mathbf{x}).$$

It follows from the last two equations that

$$\mathbf{T}_{g_1 g_2} \psi = \mathbf{T}_{g_1} \mathbf{T}_{g_2} \psi.$$

Since this holds for arbitrary  $\psi$ , we must have  $\mathbf{T}_{g_1 g_2} = \mathbf{T}_{g_1} \mathbf{T}_{g_2}$ , i.e., that  $T$  is a representation. When the action of a group is “naturally” from the left, such as the action of a matrix on a column vector, we replace  $\mathbf{x} \cdot g$  with  $g^{-1} \cdot \mathbf{x}$ . The reader can check that  $T : G \rightarrow GL(\mathcal{H})$ , given by  $\mathbf{T}_g \psi(\mathbf{x}) = \psi(g^{-1} \cdot \mathbf{x})$ , is indeed a representation.

**24.1.4. Example.** Let the Hamiltonian of the time-independent Schrödinger equation  $\mathbf{H}|\psi\rangle = E|\psi\rangle$  be invariant under the action of a group  $G$ . This means that

$$\mathbf{T}_g \mathbf{H} \mathbf{T}_g^{-1} = \mathbf{H} \Rightarrow [\mathbf{H}, \mathbf{T}_g] = \mathbf{0},$$

i.e., that  $\mathbf{H}$  and  $\mathbf{T}_g$  are simultaneously diagonalizable (Theorem 4.4.15). It follows that we can choose the energy eigenstates to be eigenstates of  $\mathbf{T}_g$  as well, and we can label the states not only by the energy “quantum numbers”—eigenvalues of  $\mathbf{H}$ —but also by the eigenvalues of  $\mathbf{T}_g$ . For example, if the Hamiltonian is invariant under the action of parity  $P$ , then we can choose the states to be *even*, corresponding to parity eigenvalue of  $+1$ , or *odd*, corresponding to parity eigenvalue of  $-1$ . Similarly, if  $G$  is the rotation group, then the states can be labeled by the eigenvalues of the rotation operators, which are, as we shall see, equivalent to the angular momentum operators discussed in Chapter 12.

In crystallography and solid-state physics, the Hamiltonian of an (infinite) lattice is invariant under translation by an integer multiple of each so-called *primitive lattice translation*, the three noncoplanar vectors that define a primitive cell of the crystal. The preceding argument shows that the energy eigenstates can be taken to be the eigenstates of the translation operator as well. ■

Energy eigenstates can be labeled by eigenvalues of the symmetry operators as well.

It is common to choose a basis and represent all  $\mathbf{T}_g$ ’s in terms of matrices. Then one gets a **matrix representation** of the group  $G$ .

matrix representations

**24.1.5. Example.** Consider the action of the 2D rotation group  $SO(2)$  (rotation about the  $z$ -axis) on  $\mathbb{R}^3$ :

$$\begin{aligned} x' &= x \cos \theta - y \sin \theta, \\ \mathbf{r}' = R_z(\theta) \mathbf{r} &\Rightarrow y' = x \sin \theta + y \cos \theta, \\ z' &= z. \end{aligned}$$

For a Hilbert space, also choose  $\mathbb{R}^3$ . Define the homomorphism  $T : G \rightarrow GL(\mathcal{H})$  to be the identity map, so that  $T(R_z(\theta)) \equiv \mathbf{T}_\theta = R_z(\theta)$ . The operator  $\mathbf{T}_\theta$  transforms the standard basis vectors of  $\mathcal{H}$  as

$$\begin{aligned} \mathbf{T}_\theta \hat{\mathbf{e}}_1 &= \mathbf{T}_\theta(1, 0, 0) = (\cos \theta, \sin \theta, 0) = \cos \theta \hat{\mathbf{e}}_1 + \sin \theta \hat{\mathbf{e}}_2 + 0 \hat{\mathbf{e}}_3, \\ \mathbf{T}_\theta \hat{\mathbf{e}}_2 &= \mathbf{T}_\theta(0, 1, 0) = (-\sin \theta, \cos \theta, 0) = -\sin \theta \hat{\mathbf{e}}_1 + \cos \theta \hat{\mathbf{e}}_2 + 0 \hat{\mathbf{e}}_3, \\ \mathbf{T}_\theta \hat{\mathbf{e}}_3 &= \mathbf{T}_\theta(0, 0, 1) = (0, 0, 1) = 0 \hat{\mathbf{e}}_1 + 0 \hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3. \end{aligned}$$

It follows that the *matrix representation* of  $SO(2)$  in the standard basis of  $\mathcal{H}$  is

$$\mathbf{T}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that  $SO(2)$  is an *infinite* group; its cardinality is determined by the “number” of  $\theta$ 's. ■

**24.1.6. Example.** Let  $S_3$  act on  $\mathbb{R}^3$  on the right by shuffling components:

$$(x_1, x_2, x_3) \cdot \pi = (x_{\pi(1)}, x_{\pi(2)}, x_{\pi(3)}), \quad \pi \in S_3.$$

For the carrier space, choose  $\mathbb{R}^3$  as well. Let  $T : S_3 \rightarrow GL(\mathbb{R}^3)$  be given as follows:  $T(\pi)$  is the matrix that takes the column vector  $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$  to  $\begin{pmatrix} x_{\pi(1)} \\ x_{\pi(2)} \\ x_{\pi(3)} \end{pmatrix}$ . As a specific illustration, consider  $\pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$  and write  $\mathbf{T}_\pi$  for  $T(\pi)$ . Then

$$\begin{aligned} \mathbf{T}_\pi(\hat{\mathbf{e}}_1) &= \mathbf{T}_\pi(1, 0, 0) = (1, 0, 0) \cdot \pi = (0, 1, 0) = \hat{\mathbf{e}}_2, \\ \mathbf{T}_\pi(\hat{\mathbf{e}}_2) &= \mathbf{T}_\pi(0, 1, 0) = (0, 1, 0) \cdot \pi = (0, 0, 1) = \hat{\mathbf{e}}_3, \\ \mathbf{T}_\pi(\hat{\mathbf{e}}_3) &= \mathbf{T}_\pi(0, 0, 1) = (0, 0, 1) \cdot \pi = (1, 0, 0) = \hat{\mathbf{e}}_1, \end{aligned}$$

which give rise to the matrix

$$\mathbf{T}_\pi = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

The reader may construct the other five matrices of this representation and verify directly that it is indeed a (faithful) representation: Products and inverses of permutations are mapped onto products and inverses of the corresponding matrices. ■

The utility of a representation lies in our comfort with the structure of vector spaces. The climax of such comfort is the spectral decomposition theorems of (normal) operators on vector spaces of finite (Chapter 4) and infinite (Chapter 16) dimensions. The operators  $\mathbf{T}_g$ , relevant to our present discussion, are, in general, neither normal nor simultaneously commuting. Therefore, the complete diagonalizability of all  $\mathbf{T}_g$ 's is out of the question (unless the group happens to be abelian).

The best thing next to complete diagonalization is to see whether there are common *invariant subspaces* of the vector space  $\mathcal{H}$  carrying the representation. We already know how to construct (minimal) “invariant” subsets of  $\mathcal{H}$ : these are precisely the *orbits* of the action of the group  $G$  on  $\mathcal{H}$ . The linearity of  $\mathbf{T}_g$ 's guarantees that the span of each orbit is actually an invariant subspace, and that such subspaces are the smallest invariant subspaces containing a given vector. Our aim is to find those *minimal* invariant subspaces whose orthogonal complements are also invariant. We encountered the same situation in Chapter 4 for a single operator.

reducible and  
irreducible  
representations

**24.1.7. Definition.** A representation  $T : G \rightarrow GL(\mathcal{H})$  is called **reducible** if there exist subspaces  $\mathcal{U}$  and  $\mathcal{W}$  of  $\mathcal{H}$  such that  $\mathcal{H} = \mathcal{U} \oplus \mathcal{W}$  and both  $\mathcal{U}$  and  $\mathcal{W}$  are invariant under all  $T_g$ 's. If no such subspaces exist,  $\mathcal{H}$  is said to be **irreducible**.

In most cases of physical interest, where  $\mathcal{H}$  is a Hilbert space,  $\mathcal{W} = \mathcal{U}^\perp$ . Then, in the language of Definition 4.2.1, a representation is reducible if a proper subspace of  $\mathcal{H}$  reduces all  $T_g$ 's.

**24.1.8. Example.** Let  $S_3$  act on  $\mathbb{R}^3$  as in Example 24.1.6. For the carrier space  $\mathcal{H}$ , choose the space of functions on  $\mathbb{R}^3$ , and for  $T$ , the homomorphism  $T : G \rightarrow GL(\mathcal{H})$ , given by  $T_g \psi(\mathbf{x}) = \psi(\mathbf{x} \cdot g)$ , for  $\psi \in \mathcal{H}$ . Any  $\psi$  that is symmetric in  $x, y, z$ , such as  $xyz, x + y + z$ , or  $x^2 + y^2 + z^2$ , defines a one-dimensional invariant subspace of  $\mathcal{H}$ . To obtain another invariant subspace, consider  $\psi_1(x, y, z) \equiv xy$  and let  $\{\pi_i\}_{i=1}^6$  be as given in Example 23.4.1. Then, denoting  $T_{\pi_i}$  by  $T_i$ , the reader may check that

$$\begin{aligned} [T_1 \psi_1](x, y, z) &= \psi_1((x, y, z) \cdot \pi_1) = \psi_1(x, y, z) = xy = \psi_1(x, y, z), \\ [T_2 \psi_1](x, y, z) &= \psi_1((x, y, z) \cdot \pi_2) = \psi_1(y, x, z) = yx = \psi_1(x, y, z), \\ [T_3 \psi_1](x, y, z) &= \psi_1((x, y, z) \cdot \pi_3) = \psi_1(z, y, x) = zy \equiv \psi_2(x, y, z), \\ [T_4 \psi_1](x, y, z) &= \psi_1((x, y, z) \cdot \pi_4) = \psi_1(x, z, y) = xz \equiv \psi_3(x, y, z), \\ [T_5 \psi_1](x, y, z) &= \psi_1((x, y, z) \cdot \pi_5) = \psi_1(z, x, y) = zx = \psi_3(x, y, z), \\ [T_6 \psi_1](x, y, z) &= \psi_1((x, y, z) \cdot \pi_6) = \psi_1(y, z, x) = yz = \psi_2(x, y, z). \end{aligned}$$

This is clearly a three-dimensional invariant subspace of  $\mathcal{H}$  with  $\psi_1, \psi_2$ , and  $\psi_3$  as a convenient basis, in which the first three permutations are represented by

$$T_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It is instructive for the reader to verify these relations and to find the three remaining matrices. ■

**24.1.9. Example.** Let  $S_3$  act on  $\mathbb{R}^3$  as in Example 24.1.6. For the carrier space of representation, choose the subspace  $\mathcal{V}$  of the  $\mathcal{H}$  of Example 24.1.8 spanned by the six functions  $x, y, z, xy, xz$ , and  $yz$ . For  $T$ , choose the same homomorphism as in Example 24.1.8 restricted to  $\mathcal{V}$ . It is clear that the subspaces  $\mathcal{U}$  and  $\mathcal{W}$  spanned, respectively, by the first three and the last three functions are invariant under  $S_3$ , and that  $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$ . It follows that the representation is reducible. The matrix form of this representation is found to be of the general form  $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ , where  $B$  is one of the 6 matrices of Example 24.1.8. The matrix  $A$ , corresponding to the three functions  $x, y$ , and  $z$ , can be found similarly. ■

Let  $\mathcal{H}$  be a carrier space, finite- or infinite-dimensional. For any vector  $|a\rangle$ , the reader may check that the span of  $\{T_g |a\rangle\}_{g \in G}$  is an invariant subspace of  $\mathcal{H}$ . If  $G$  is finite, this subspace is clearly finite-dimensional. The irreducible subspace containing  $|a\rangle$ , a subspace of the span of  $\{T_g |a\rangle\}_{g \in G}$ , will also be finite-dimensional. Because of the arbitrariness of  $|a\rangle$ , it follows that every vector of  $\mathcal{H}$  lies in an irreducible subspace, and that

**24.1.10. Box.** *All irreducible representations of a finite group are finite-dimensional.*

Due to the importance and convenience of unitary operators (for example, the fact that they leave the inner product invariant), it is desirable to be able to construct a unitary representation—or a representation that is equivalent to one—of groups. The following theorem ensures that this desire can be realized for finite groups.

All representations are equivalent to unitary representations.

**24.1.11. Theorem.** *Every representation of a finite group  $G$  is equivalent to some unitary representation.*

*Proof.* We present the proof because of its simplicity and elegance. Let  $T$  be a representation of  $G$ . Consider the positive hermitian operator  $\mathbf{T} \equiv \sum_{x \in G} \mathbf{T}_x^\dagger \mathbf{T}_x$  and note that

$$\begin{aligned} \mathbf{T}_g^\dagger \mathbf{T}_g &= \sum_{x \in G} [T(g)]^\dagger [T(x)]^\dagger T(x) T(g) \\ &= \sum_{x \in G} [T(xg)]^\dagger T(xg) = \sum_{y \in G} [T(y)]^\dagger T(y) = \mathbf{T}, \end{aligned} \tag{24.2}$$

where we have used the fact that the sum over  $x$  and  $y \equiv xg$  sweep through the entire group. Now let  $\mathbf{S} = \sqrt{\mathbf{T}}$ , and multiply both sides of Equation (24.2)—with  $\mathbf{S}^2$  replacing  $\mathbf{T}$ —by  $\mathbf{S}^{-1}$  on the left and by  $\mathbf{T}_g^{-1} \mathbf{S}^{-1}$  on the right to obtain

$$\mathbf{S}^{-1} \mathbf{T}_g^\dagger \mathbf{S} = \mathbf{S} \mathbf{T}_g^{-1} \mathbf{S}^{-1} \Rightarrow (\mathbf{S} \mathbf{T}_g \mathbf{S}^{-1})^\dagger = (\mathbf{S} \mathbf{T}_g \mathbf{S}^{-1})^{-1} \quad \forall g \in G.$$

This shows that the representation  $T'$  defined by  $\mathbf{T}'_g \equiv \mathbf{S} \mathbf{T}_g \mathbf{S}^{-1}$  for all  $g \in G$  is unitary. □

There is another convenience afforded by unitary representations:

**24.1.12. Theorem.** *Let  $T : G \rightarrow GL(\mathcal{H})$  be a unitary representation and  $\mathcal{W}$  an invariant subspace of  $\mathcal{H}$ . Then,  $\mathcal{W}^\perp$  is also invariant.*

*Proof.* Suppose  $|a\rangle \in \mathcal{W}^\perp$ . We need to show that  $\mathbf{T}_g |a\rangle \in \mathcal{W}^\perp$  for all  $g \in G$ . To this end, let  $|b\rangle \in \mathcal{W}$ . Then

$$\langle b | \mathbf{T}_g |a\rangle = (\langle a | \mathbf{T}_g^\dagger |b\rangle)^* = (\langle a | \mathbf{T}_g^{-1} |b\rangle)^* = (\langle a | \mathbf{T}_{g^{-1}} |b\rangle)^* = 0,$$

because  $\mathbf{T}_{g^{-1}} |b\rangle \in \mathcal{W}$ . It follows from this equality that  $\mathbf{T}_g |a\rangle \in \mathcal{W}^\perp$  for all  $g \in G$ . □

The carrier space  $\mathcal{H}$  of a unitary representation is either irreducible or has an invariant subspace  $\mathcal{W}$ , in which case we have  $\mathcal{H} = \mathcal{W} \oplus \mathcal{W}^\perp$ , where, by Theorem 24.1.12,  $\mathcal{W}^\perp$  is also invariant. If  $\mathcal{W}$  and  $\mathcal{W}^\perp$  are not irreducible, then they too can



be written as direct sums of invariant subspaces. Continuing this process, we can decompose  $\mathcal{H}$  into irreducible invariant subspaces  $\mathcal{W}^{(k)}$  such that

$$\mathcal{H} = \mathcal{W}^{(1)} \oplus \mathcal{W}^{(2)} \oplus \mathcal{W}^{(3)} \oplus \dots$$

If the carrier space is finite-dimensional, which we assume from now on and for which we use the notation  $\mathcal{V}$ , then the above direct sum is finite and we write

$$\mathcal{V} = \mathcal{W}^{(1)} \oplus \mathcal{W}^{(2)} \oplus \dots \oplus \mathcal{W}^{(p)} \equiv \sum_{k=1}^p \mathcal{W}^{(k)}. \tag{24.3}$$

One can think of  $\mathcal{W}^{(k)}$  as the carrier space of an (irreducible) representation. The homomorphism  $T^{(k)} : G \rightarrow GL(\mathcal{W}^{(k)})$  is simply the restriction of  $T$  to the subspace  $\mathcal{W}^{(k)}$ , and we write

$$\mathbf{T}_g = \mathbf{T}_g^{(1)} \oplus \mathbf{T}_g^{(2)} \oplus \dots \oplus \mathbf{T}_g^{(r)} \equiv \sum_{k=1}^r \mathbf{T}_g^{(k)}.$$

If we identify all equivalent irreducible representations and collect them together, we may rewrite the last equation as

$$\mathbf{T}_g = m_1 \mathbf{T}_g^{(1)} \oplus m_2 \mathbf{T}_g^{(2)} \oplus \dots \oplus m_\rho \mathbf{T}_g^{(\rho)} \equiv \sum_{\alpha=1}^{\rho} \oplus m_\alpha \mathbf{T}_g^{(\alpha)}, \tag{24.4}$$

where  $\rho$  is the number of *inequivalent* irreducible representations and  $m_\alpha$  are positive integers giving the number of times an irreducible representation  $\mathbf{T}_g^{(\alpha)}$  and all its equivalents occur in a given representation.

In terms of matrices,  $\mathbf{T}_g$  will be represented in a block-diagonal form as

$$\mathbf{T}_g = \begin{pmatrix} \mathbf{T}_g^{(1)} & 0 & \dots & 0 \\ 0 & \mathbf{T}_g^{(2)} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \mathbf{T}_g^{(r)} \end{pmatrix},$$

where some of the  $\mathbf{T}_g^{(k)}$  may be equivalent.

antisymmetric  
representation of a  
permutation group

**24.1.13. Example.** A one-dimensional (and therefore irreducible) representation, defined for all groups, is the **trivial** (symmetric) representation  $T : G \rightarrow \mathbb{C}$  given by  $T(g) = 1$  for all  $g \in G$ . For the permutation group  $S_n$ , one can define another one-dimensional (thus irreducible) representation  $T : S_n \rightarrow \mathbb{C}$ , called the **antisymmetric** representation, given by  $T(\pi) = +1$  if  $\pi$  is even, and  $T(\pi) = -1$  if  $\pi$  is odd. ■

Given any (matrix) representation  $T$  of  $G$ , one can form the transpose inverse matrices  $(\mathbf{T}_g^t)^{-1}$ , and complex conjugate matrices  $\mathbf{T}_g^*$ . The reader may check that each set of these matrices forms a representation of  $G$ .

adjoint and complex  
conjugate  
representations

**24.1.14. Definition.** The set of matrices  $(\mathbf{T}_g^t)^{-1}$  and  $\mathbf{T}_g^*$  are called, respectively, the **adjoint representation**, denoted by  $\bar{T}$ , and the **complex conjugate representation**, denoted by  $T^*$ .

## 24.2 Orthogonality Properties

Homomorphisms preserve group structures. By studying a group that is more attuned to concrete manipulations, we gain insight into the structure of groups that are homomorphic to it. The group of invertible operators on a vector space, especially in their matrix representation, are particularly suited for such a study because of our familiarity with matrices and operators. The last section reduced this study to inequivalent irreducible representations. This section is devoted to a detailed study of such representations.

Schur's lemma **24.2.1. Lemma.** (Schur's lemma) *Let  $T : G \rightarrow GL(\mathcal{V})$  and  $T' : G \rightarrow GL(\mathcal{V}')$  be irreducible representations of  $G$ . If  $\mathbf{A} \in \mathcal{L}(\mathcal{V}, \mathcal{V}')$  is such that*

$$\mathbf{A}T_g = T'_g\mathbf{A} \quad \forall g \in G, \quad (24.5)$$

*then either  $\mathbf{A}$  is an isomorphism (i.e.,  $T$  is equivalent to  $T'$ ), or  $\mathbf{A} = 0$ .*

*Proof.* Let  $|a\rangle \in \ker \mathbf{A}$ . Then

$$\mathbf{A}T_g|a\rangle = T'_g\underbrace{\mathbf{A}|a\rangle}_{=0} = 0 \Rightarrow T_g|a\rangle \in \ker \mathbf{A} \quad \forall g \in G.$$

It follows that  $\ker \mathbf{A}$ , a subspace of  $\mathcal{V}$ , is invariant under  $T$ . Irreducibility of  $T$  implies that either  $\ker \mathbf{A} = \mathcal{V}$ , or  $\ker \mathbf{A} = 0$ . The first case asserts that  $\mathbf{A}$  is the zero linear transformation; the second case implies that  $\mathbf{A}$  is injective.

Similarly, let  $|b\rangle \in \mathbf{A}(\mathcal{V})$ . Then  $|b\rangle = \mathbf{A}|x\rangle$  for some  $|x\rangle \in \mathcal{V}$ :

$$T'_g|b\rangle = T'_g\mathbf{A}|x\rangle = \underbrace{\mathbf{A}T_g|x\rangle}_{\in \mathbf{A}(\mathcal{V})} \Rightarrow T'_g|b\rangle \in \mathbf{A}(\mathcal{V}) \quad \forall g \in G.$$

It follows that  $\mathbf{A}(\mathcal{V})$ , a subspace of  $\mathcal{V}'$ , is invariant under  $T'$ . Irreducibility of  $T'$  implies that either  $\mathbf{A}(\mathcal{V}) = 0$ , or  $\mathbf{A}(\mathcal{V}) = \mathcal{V}'$ . The first case is consistent with the first conclusion drawn above:  $\ker \mathbf{A} = \mathcal{V}$ . The second case asserts that  $\mathbf{A}$  is surjective. Combining the two results, we conclude that  $\mathbf{A}$  is either the zero operator or an isomorphism.  $\square$

Lemma 24.2.1 becomes extremely useful when we concentrate on a single irreducible representation, i.e., when  $T' = T$ .

**24.2.2. Lemma.** *Let  $T : G \rightarrow GL(\mathcal{V})$  be an irreducible representation of  $G$ . If  $\mathbf{A} \in \mathcal{L}(\mathcal{V})$  is such that  $\mathbf{A}T_g = T_g\mathbf{A}$  for all  $g \in G$ , then  $\mathbf{A} = \lambda\mathbf{1}$ .*

*Proof.* Replacing  $\mathcal{V}'$  with  $\mathcal{V}$  in Lemma 24.2.1, we conclude that  $\mathbf{A} = 0$  or  $\mathbf{A}$  is an isomorphism of  $\mathcal{V}$ . In the first case,  $\lambda = 0$ . In the second case,  $\mathbf{A}$  must have a nonzero eigenvalue  $\lambda$  and at least one eigenvector (see Theorem 4.3.4). It follows that the operator  $\mathbf{A} - \lambda\mathbf{1}$  commutes with all  $T_g$ 's and it is not an isomorphism (why not?). Therefore, it must be the zero operator.  $\square$

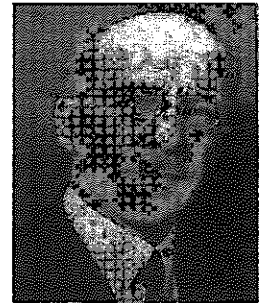
We can immediately put this lemma to good use. If  $G$  is abelian, all operators  $\{T_x\}_{x \in G}$  commute with one another. Focusing on one of these operators, say  $T_g$ , noting that it commutes with all operators of the representation, and using Lemma 24.2.2, we conclude that  $T_g = \lambda 1$ . It follows that when  $T_g$  acts on a vector, it gives a multiple of that vector. Therefore, it leaves any one-dimensional subspace of the carrier space invariant. Since this is true for all  $g \in G$ , we have the following result.

**24.2.3. Theorem.** *All irreducible representations of an abelian group are one-dimensional.*

This theorem is an immediate consequence of Schur's lemma, and is independent of the order of  $G$ . In particular, it holds for infinite groups, if Schur's lemma holds for those groups. One important class of infinite groups for which Schur's lemma holds is the Lie groups. Thus, all abelian Lie groups have 1-dimensional irreducible representations. We shall see later that the converse of Theorem 24.2.3 is also true for finite groups.

---

**Issai Schur** (1875–1945) was one of the most brilliant mathematicians active in Germany during the first third of the twentieth century. He attended the Gymnasium in Libau (now Liepaja, Latvia) and then the University of Berlin, where he spent most of his scientific career from 1911 until 1916. When he returned to Berlin, he was an assistant professor at Bonn. He became full professor at Berlin in 1919. Schur was forced to retire by the Nazi authorities in 1935 but was able to emigrate to Palestine in 1939. He died there of a heart ailment several years later. Schur had been a member of the Prussian Academy of Sciences before the Nazi purges. He married and had a son and daughter.



Schur's principal field was the representation theory of groups, founded a little before 1900 by his teacher Frobenius. Schur seems to have completed this field shortly before World War I, but he returned to the subject after 1925, when it became important for physics. Further developed by his student Richard Brauer, it is in our time experiencing an extraordinary growth through the opening of new questions. Schur's dissertation (1901) became fundamental to the representation theory of the general linear group; in fact, English mathematicians have named certain of the functions appearing in the work "S-functions" in Schur's honor. In 1905 Schur reestablished the theory of group characters—the keystone of representation theory. The most important tool involved is "Schur's lemma." Along with the representation of groups by integral linear substitutions, Schur was also the first to study representation by linear fractional substitutions, treating this more difficult problem almost completely in two works (1904, 1907). In 1906 Schur considered the fundamental problems that appear when an algebraic number field is taken as the domain; a number appearing in this connection is now called the Schur index. His works written after 1925 include a complete description of the rational and of the continuous representations of the general linear group; the foundations of this work were in his dissertation.

A lively interchange with many colleagues led Schur to contribute important memoirs to other areas of mathematics. Some of these were published as collaborations with other authors, although publications with dual authorship were almost unheard of at that time. Here we simply indicate the areas: pure group theory, matrices, algebraic equations, number theory, divergent series, integral equations, and function theory.

**24.2.4. Example.** Suppose that the Hamiltonian  $\mathbf{H}$  of a quantum mechanical system with Hilbert space  $\mathcal{H}$  has a group of symmetry with a representation  $T : G \rightarrow GL(\mathcal{H})$ . Then  $\mathbf{H}T_g = T_g\mathbf{H}$  for all  $g \in G$ . It follows that  $\mathbf{H} = \lambda\mathbf{1}$  if the representation is irreducible. Therefore,

**24.2.5. Box.** All vectors of each invariant irreducible subspace are eigenstates of the hamiltonian corresponding to the same eigenvalue, i.e., they all have the same energy. Therefore, the degeneracy of that energy state is at least as large as the dimension of the carrier space.

It is helpful to arrive at the statement above from a different perspective. Consider a vector  $|x\rangle$  in the eigenspace  $\mathcal{M}_i$  corresponding to the energy eigenvalue  $E_i$ . Since  $T_g$  and  $\mathbf{H}$  commute,  $T_g|x\rangle$  is also in  $\mathcal{M}_i$ . Therefore, an eigenspace of a Hamiltonian with a group of symmetry is invariant under all  $T_g$  for any representation  $T$  of that group. If  $T$  is one of the irreducible representations of  $G$ , say  $T^{(\alpha)}$  with dimension  $n_\alpha$ , then  $\dim \mathcal{M}_i \geq n_\alpha$ . ■

Consider two irreducible representations  $T^{(\alpha)}$  and  $T^{(\beta)}$  of a group  $G$  with carrier spaces  $\mathcal{W}^{(\alpha)}$  and  $\mathcal{W}^{(\beta)}$ , respectively. Let  $\mathbf{X}$  be any operator in  $\mathcal{L}(\mathcal{W}^{(\alpha)}, \mathcal{W}^{(\beta)})$ , and define

$$\mathbf{A} \equiv \sum_{x \in G} T_x^{(\alpha)} \mathbf{X} T_{x^{-1}}^{(\beta)} = \sum_{x \in G} T^{(\alpha)}(x) \mathbf{X} T^{(\beta)}(x^{-1}).$$

Then, we have

$$\begin{aligned} T_g^{(\alpha)} \mathbf{A} &= \sum_{x \in G} T^{(\alpha)}(g) T^{(\alpha)}(x) \mathbf{X} T^{(\beta)}(x^{-1}) T^{(\beta)}(g^{-1}) T^{(\beta)}(g) \\ &= \sum_{x \in G} T^{(\alpha)}(gx) \mathbf{X} T^{(\beta)}((gx)^{-1}) T^{(\beta)}(g) = \mathbf{A} T_g^{(\beta)}. \end{aligned}$$

=  $\mathbf{A}$  because this sum also covers all  $G$

We are interested in the two cases where  $T^{(\alpha)} = T^{(\beta)}$ , and where  $T^{(\alpha)}$  is not equivalent to  $T^{(\beta)}$ . In the first case, Lemma 24.2.2 gives  $\mathbf{A} = \lambda\mathbf{1}$ ; in the second case, Lemma 24.2.1 gives  $\mathbf{A} = 0$ . Combining these two results and labeling the constant multiplying the unit operator by  $X$ , we can write

$$\sum_{g \in G} T^{(\alpha)}(g) \mathbf{X} T^{(\beta)}(g^{-1}) = \lambda_X \delta_{\alpha\beta} \mathbf{1}. \tag{24.6}$$

The presence of the completely arbitrary operator  $\mathbf{X}$  indicates that Equation (24.6) is a powerful statement about—and a severe restriction on—the operators  $T^{(\alpha)}(g)$ . This becomes more transparent if we select a basis, represent all operators by matrices, and for  $\mathbf{X}$ , the matrix representation of  $\mathbf{X}$ , choose a matrix whose only nonzero element is 1 and occurs at the  $l$ th row and  $m$ th column. Then Equation (24.6) becomes

$$\sum_{g \in G} T_{il}^{(\alpha)}(g) T_{mj}^{(\beta)}(g^{-1}) = \lambda_{lm} \delta_{\alpha\beta} \delta_{ij},$$

where  $\lambda_{lm}$  is a constant that can be evaluated as follows. Set  $j = i$ ,  $\alpha = \beta$ , and sum over  $i$ . The RHS will give  $\lambda_{lm} \sum_i \delta_{ii} = \lambda_{lm} n_\alpha$ , where  $n_\alpha$  is the dimension of the carrier space of  $T^{(\alpha)}$ . For the LHS we get

$$\begin{aligned} LHS &= \sum_{g \in G} \sum_i T_{il}^{(\alpha)}(g) T_{mi}^{(\alpha)}(g^{-1}) = \sum_{g \in G} (T^{(\alpha)}(g^{-1}) T^{(\alpha)}(g))_{ml} \\ &= \sum_{g \in G} T_{ml}^{(\alpha)}(g^{-1}g) = \sum_{g \in G} \underbrace{T_{ml}^{(\alpha)}(e)}_{=(1)_{ml}} = |G| \delta_{ml}, \end{aligned}$$

where  $|G|$  is the order of the group. Putting everything together, we obtain

$$\sum_{g \in G} T_{il}^{(\alpha)}(g) T_{mj}^{(\beta)}(g^{-1}) = \frac{|G|}{n_\alpha} \delta_{ml} \delta_{\alpha\beta} \delta_{ij}, \tag{24.7}$$

or

$$\sum_{g \in G} T_{il}^{(\alpha)}(g) T_{jm}^{(\beta)*}(g) = \frac{|G|}{n_\alpha} \delta_{ml} \delta_{\alpha\beta} \delta_{ij} \tag{24.8}$$

if the representation is unitary.

Equations (24.7) and (24.8) depend on the basis chosen in which to express matrices. To eliminate this dependence, we first introduce the important concept of character.

character of a representation;  
simple character,  
compound character

**24.2.6. Definition.** Let  $T : G \rightarrow GL(\mathcal{V})$  be a representation of the group  $G$ . The *character* of this representation is the map  $\chi : G \rightarrow \mathbb{C}$  given by

$$\chi(g) \equiv \text{tr } T_g = \sum_i T_{ii}(g),$$

where  $T(g)$  is the matrix representation of  $T_g$  in any basis of  $\mathcal{V}$ . If  $T$  is irreducible, the character is called *simple*; otherwise, it is called *compound*.

The character of the identity element in any representation can be calculated immediately. Since a homomorphism maps identity onto identity,  $T_e = \mathbf{1}$ . Therefore,

$$\chi(e) = \text{tr}(\mathbf{1}) = \dim \mathcal{V}. \tag{24.9}$$

Recall that two elements  $x, y \in G$  belong to the same conjugacy class if there exist  $g \in G$  such that  $x = gyg^{-1}$ . This same relation holds for the operators representing the elements:  $T_x = T_g T_y T_g^{-1}$ . Taking the trace of both sides, and noting that  $T_{g^{-1}} = T_g^{-1}$ , one can show that

**24.2.7. Box.** *All elements of a group belonging to the same conjugacy class have the same character.*

Setting  $i = l$  and  $j = m$  in (24.7) and summing over  $i$  and  $j$ , we obtain

$$\sum_{g \in G} \chi^{(\alpha)}(g) \chi^{(\beta)}(g^{-1}) = \frac{|G|}{n_\alpha} \delta_{\alpha\beta} \sum_{i,j} \delta_{ji} \delta_{ij} = \frac{|G|}{n_\alpha} \delta_{\alpha\beta} \underbrace{\sum_j \delta_{jj}}_{=n_\alpha} = |G| \delta_{\alpha\beta}, \tag{24.10}$$

or

$$\sum_{g \in G} \chi^{(\alpha)}(g) \chi^{(\beta)*}(g) = |G| \delta_{\alpha\beta} \tag{24.11}$$

if the representation is unitary. This equation suggests a useful interpretation: Characters can be thought of as vectors in a  $|G|$ -dimensional inner product space. According to Equation (24.11), the characters of inequivalent irreducible representations are orthogonal. In particular, since there cannot be more orthogonal vectors than the dimension of a vector space, we conclude that the number of irreducible inequivalent representations of a group cannot be more than the cardinality of that group. Actually, we can do better. Restricting ourselves to unitary representations and collecting all elements belonging to the same conjugacy class together, we write

$$\sum_{i=1}^r c_i \chi_i^{(\alpha)} \chi_i^{(\beta)*} = |G| \delta_{\alpha\beta} \Rightarrow \langle \chi^{(\beta)} | \chi^{(\alpha)} \rangle = |G| \delta_{\alpha\beta}, \tag{24.12}$$

where  $i$  labels conjugacy classes,  $c_i$  is the number of elements in the  $i$ th class,  $r$  is the number of classes in  $G$ , and  $|\chi^{(\alpha)}\rangle \in \mathbb{C}^r$  is an  $r$ -dimensional vector with components  $\{c_i^{1/2} \chi_i^{(\alpha)}\}_{i=1}^r$ . Equation (24.12) shows that vectors belonging to different irreducible representations are orthogonal. Since there cannot be more orthogonal vectors than the dimension of a vector space, we conclude that *the number of inequivalent irreducible representations of a group cannot be more than the number of conjugacy classes of the group, i.e.,  $\rho \leq r$ .*

The characters of the adjoint representation are obtained from

$$\bar{\chi}(g) = \chi(g^{-1}) \Rightarrow \bar{\chi}_i = \chi_{i'},$$

where  $K_{i'}$  is the class consisting of all elements inverse to those of the class  $K_i$ . The equations involving characters of inverses of group elements can be written

in terms of the characters of the adjoint representation. For example, Equation (24.10) becomes

$$\sum_{g \in G} \chi^{(\alpha)}(g) \bar{\chi}^{(\beta)}(g) = |G| \delta_{\alpha\beta} \Rightarrow \sum_{i=1}^r c_i \chi_i^{(\alpha)} \bar{\chi}_i^{(\beta)} = |G| \delta_{\alpha\beta}. \tag{24.13}$$

Other relations can be obtained similarly.

### 24.3 Analysis of Representations

We can use the results obtained in the last section to gain insight into a given representation. Take the trace of both sides of Equation (24.4) and write the result as

$$\chi(g) = m_1 \chi^{(1)}(g) + \dots + m_\rho \chi^{(\rho)}(g) \equiv \sum_{\alpha=1}^{\rho} m_\alpha \chi^{(\alpha)}(g); \tag{24.14}$$

i.e., a compound character is a linear combination of simple characters with non-negative integer coefficients. Furthermore, the orthogonality of simple characters gives

$$m_\alpha = \frac{1}{|G|} \sum_{g \in G} \chi(g) \chi^{(\alpha)*}(g), \tag{24.15}$$

yielding the number of times the irreducible representation  $T^{(\alpha)}$  occurs in the representation  $T$ .

Another useful relation will be obtained if we multiply Equation (24.14) by its complex conjugate and sum over  $g$ ; the result is

$$\begin{aligned} \sum_{g \in G} |\chi(g)|^2 &= \sum_{g \in G} \chi(g) \chi^*(g) = \sum_{g \in G} \sum_{\alpha} m_\alpha \chi^{(\alpha)}(g) \sum_{\beta} m_\beta \chi^{(\beta)*}(g) \\ &= \sum_{\alpha, \beta} m_\alpha m_\beta \underbrace{\sum_{g \in G} \chi^{(\alpha)}(g) \chi^{(\beta)*}(g)}_{=|G| \delta_{\alpha\beta}} = |G| \sum_{\alpha} m_\alpha^2. \end{aligned} \tag{24.16}$$

In particular, if  $T$  is irreducible, all  $m_\alpha$  are zero except for one, which is unity. We therefore obtain the **criterion for irreducibility**:

$$\sum_{g \in G} |\chi(g)|^2 = \sum_{i=1}^r c_i |\chi_i|^2 = |G| \quad \text{if } T \text{ is irreducible.} \tag{24.17}$$

For groups of low order and representations of small dimensions, Equation (24.16) becomes a powerful tool for testing the irreducibility of the representation.

**24.3.1. Example.** Let  $G = S_3$  and consider the representation of Example 24.1.8. The characters of the first three elements of this representation are easily calculated:

$$\chi_1 = \text{tr } \mathbf{T}_1 = 3, \quad \chi_2 = \text{tr } \mathbf{T}_2 = 1, \quad \chi_3 = \text{tr } \mathbf{T}_3 = 1.$$

Similarly, one can obtain  $\chi_4 = 1, \chi_5 = 0,$  and  $\chi_6 = 0$ . Substituting this in Equation (24.16) yields

$$\sum_{g \in G} |\chi(g)|^2 = \sum_{j=1}^6 |\chi_j|^2 = 3^2 + 1^2 + 1^2 + 1^2 + 0^2 + 0^2 = 12.$$

Comparing this with the RHS of (24.16) with  $|G| = 6$  yields  $\sum_{\alpha} m_{\alpha}^2 = 2$ . This restricts the nonzero  $\alpha$ 's to two, say  $\alpha = 1$  and  $\alpha = 2$ . Moreover,  $m_1$  and  $m_2$  can be only 1. Thus, the representation of Example 24.1.8 is reducible, and there are precisely two inequivalent irreducible representations in it, each occurring once.

We can actually find the invariant subspaces corresponding to the two irreducible representations revealed above. The first is easy to guess. Just taking the sum of the three functions  $\psi_1, \psi_2,$  and  $\psi_3$  gives a one-dimensional invariant subspace; so, let  $\phi_1 \equiv \psi_1 + \psi_2 + \psi_3,$  and note that the space  $\mathcal{W}_1$  spanned by  $\phi_1$  is invariant. The second is harder to discover. However, if we assume that  $\psi_1, \psi_2,$  and  $\psi_3$  are orthonormal, then using the Gram-Schmidt process, we can find the other two functions orthogonal to  $\phi_1$  (but not orthogonal to each other!). These are

$$\phi_2 = -\psi_1 + 2\psi_2 - \psi_3 \quad \phi_3 = -\psi_1 - \psi_2 + 2\psi_3.$$

The reader is urged to convince himself/herself that the subspace  $\mathcal{W}^{(2)}$  spanned by  $\phi_2$  and  $\phi_3$  is the complement of  $\mathcal{W}^{(1)}$  [i.e.,  $\mathcal{V} = \mathcal{W}^{(1)} \oplus \mathcal{W}^{(2)}$ ] and that it is invariant under all  $\mathbf{T}_g$ 's. ■

A very useful representation can be constructed as follows. Let  $G = \{g_j\}_{j=1}^m,$  and recall that left multiplication of elements of  $G$  by a fixed element  $g_i$  is a permutation of  $(g_1, g_2, \dots, g_m).$  Denote this permutation by  $\pi_i.$  Now define a regular representation  $R : G \rightarrow GL(\mathbb{R}^m),$  called the **regular representation**, by

$$\mathbf{R}_{g_i}(x_1, x_2, \dots, x_m) = (x_{\pi_i(1)}, x_{\pi_i(2)}, \dots, x_{\pi_i(m)}).$$

That this is indeed a representation is left as a problem for the reader. One can obtain a matrix representation of  $R$  by choosing the standard basis  $\{\hat{e}_j\}_{j=1}^m$  of  $\mathbb{R}^m$  and noting that  $\mathbf{R}_{g_i} \hat{e}_j = \hat{e}_{\pi_i^{-1}(j)}.$  From such a matrix representation it follows that all characters  $\chi^R$  of the regular representations are zero except for the identity, whose character is  $\chi^R(e) = m$  [see Equation (24.9)]. Now use Equation (24.14) for  $g = e$  and for the regular representation to obtain  $m = \sum_{\alpha=1}^{\rho} m_{\alpha} n_{\alpha}$  where  $n_{\alpha}$  is the dimension of the  $\alpha$ -th irreducible representation. We can find  $m_{\alpha}$  by using Equation (24.15) and noting that only  $g = e$  contributes to the sum:

$$m_{\alpha} = \frac{1}{|G|} \sum_{g \in G} \chi^R(g) \chi^{(\alpha)*}(g) = \frac{1}{m} \chi^R(e) \underbrace{\chi^{(\alpha)*}(e)}_{n_{\alpha}} = n_{\alpha}.$$



In words,

**24.3.2. Box.** *The number of times an irreducible representation occurs in the regular representation is equal to the dimension of that irreducible representation.*

We therefore obtain the important relations

$$\chi_i^R = |G|\delta_{i1} = \sum_{\alpha=1}^{\rho} n_{\alpha} \chi_i^{(\alpha)} \quad \text{and} \quad |G| = \sum_{\alpha=1}^{\rho} n_{\alpha}^2, \quad (24.18)$$

where we have assumed that the first conjugacy class is that of the identity. For finite groups of small order, the second equation can be very useful in obtaining the dimensions of irreducible representations.

**24.3.3. Example.** A group of order 2 or 3 has only one-dimensional inequivalent irreducible representations, because the only way that Equation (24.18) can be satisfied for  $|G| = 2$  or  $3$  is for all  $n_{\alpha}$ 's to be 1. A group of order 4 can have either 4 one-dimensional or one 2-dimensional inequivalent irreducible representations. The symmetric group  $S_3$ , being of order 6, can have 6 one-dimensional, or 2 one-dimensional and one 2-dimensional inequivalent irreducible representations. We shall see later that if all inequivalent irreducible representations of a group are one-dimensional, then the group must be abelian. Thus, the first possibility for  $S_3$  must be excluded. ■

## 24.4 Group Algebra

Think of group elements as (linearly independent) vectors. In fact, given any set, one can generate a vector space by taking linear combinations of the elements of the set assumed to form a basis. In the case of groups one gets a bonus: The product already defined on the basis (group elements) can be extended by linearity to all elements of the vector space to turn it into an algebra called the **group algebra**. For  $G = \{g_j\}_{j=1}^m$ , a typical element of the group algebra is  $\mathbf{a} = \sum_{i=1}^m a_i g_i$ . One can add two vectors as usual. But the product of two vectors is also defined:

$$\mathbf{a}\mathbf{b} = \left( \sum_{i=1}^m a_i g_i \right) \left( \sum_{j=1}^m b_j g_j \right) = \sum_{i=1}^m \sum_{j=1}^m a_i b_j \underbrace{g_i g_j}_{g_k} \equiv \sum_{k=1}^m c_k g_k,$$

where  $c_k$  is a sum involving  $a_i$  and  $b_j$ . The best way to learn this is to see an example.

**24.4.1. Example.** Let  $G = S_3$  and consider  $\mathbf{a} = 2\pi_1 - 3\pi_3 + \pi_5$  and  $\mathbf{b} = \pi_2 - 2\pi_4 + 3\pi_6$ . Then, using Table 23.1, we obtain

$$\begin{aligned} \mathbf{ab} &= (2\pi_1 - 3\pi_3 + \pi_5)(\pi_2 - 2\pi_4 + 3\pi_6) \\ &= 2\pi_1\pi_2 - 4\pi_1\pi_4 + 6\pi_1\pi_6 - 3\pi_3\pi_2 + 6\pi_3\pi_4 \\ &\quad - 9\pi_3\pi_6 + \pi_5\pi_2 - 2\pi_5\pi_4 + 3\pi_5\pi_6 \\ &= 2\pi_2 - 4\pi_4 + 6\pi_6 - 3\pi_6 + 6\pi_5 - 9\pi_2 + \pi_4 - 2\pi_3 + 3\pi_1 \\ &= 3\pi_1 - 7\pi_2 - 2\pi_3 - 3\pi_4 + 6\pi_5 + 3\pi_6. \end{aligned}$$

Let  $\mathcal{A}$  be any algebra. As a finite-dimensional vector space, we can always find two proper subspaces  $\mathcal{L}_1$  and  $\mathcal{L}_2$  such that  $\mathcal{A}$  is a direct sum of  $\mathcal{L}_1$  and  $\mathcal{L}_2$ . We write this as<sup>2</sup>  $\mathcal{A} = \mathcal{L}_1 \dot{+} \mathcal{L}_2$ . If we demand that this sum be invariant under left (right) algebra multiplication, then it is clear that  $\mathcal{L}_1$  and  $\mathcal{L}_2$  must be left (right) ideals, in which case we write

$$\mathcal{A} = \mathcal{L}_1 \oplus \mathcal{L}_2.$$

Now assume that  $\mathcal{A}$  has an identity  $\mathbf{1}$ , which as a vector in  $\mathcal{A}$  can be decomposed uniquely as

$$\mathbf{1} = \mathbf{1}_1 + \mathbf{1}_2 \tag{24.19}$$

with  $\mathbf{1}_1 \in \mathcal{L}_1$  and  $\mathbf{1}_2 \in \mathcal{L}_2$ . Similarly,

$$\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2 \quad \forall \mathbf{a} \in \mathcal{A}.$$

Multiplying these two equations, we obtain

$$(\mathbf{a}_1 + \mathbf{a}_2)\mathbf{1} = \mathbf{a}_1 + \mathbf{a}_2 = \mathbf{a}(\mathbf{1}_1 + \mathbf{1}_2) = \mathbf{a}\mathbf{1}_1 + \mathbf{a}\mathbf{1}_2.$$

The uniqueness of the decomposition of  $\mathbf{a}$  now implies that

$$\mathbf{a}_1 = \mathbf{a}\mathbf{1}_1, \quad \mathbf{a}_2 = \mathbf{a}\mathbf{1}_2.$$

So, if  $\mathbf{a} \in \mathcal{L}_1$ , then  $\mathbf{a}_2 = 0$  and

$$\mathbf{a} = \mathbf{a}\mathbf{1}_1, \quad \mathbf{a}\mathbf{1}_2 = 0,$$

with a similar result for  $\mathbf{a} \in \mathcal{L}_2$ . Since  $\mathbf{1}_1 \in \mathcal{L}_1$  and  $\mathbf{1}_2 \in \mathcal{L}_2$ , we have

$$\mathbf{1}_1^2 = \mathbf{1}_1, \quad \mathbf{1}_1\mathbf{1}_2 = 0; \quad \mathbf{1}_2^2 = \mathbf{1}_2, \quad \mathbf{1}_2\mathbf{1}_1 = 0. \tag{24.20}$$

idempotent element  
of algebras

An element  $\mathbf{a} \in \mathcal{A}$  that satisfies  $\mathbf{a}^2 = \mathbf{a}$  is called an **idempotent**.<sup>3</sup> Thus,  $\mathbf{1}_1$  and  $\mathbf{1}_2$  are idempotents. Furthermore, they generate  $\mathcal{L}_1$  and  $\mathcal{L}_2$ , respectively; i.e.,  $\mathcal{L}_1 = \mathcal{A}\mathbf{1}_1$  and  $\mathcal{L}_2 = \mathcal{A}\mathbf{1}_2$  (see the last section of Chapter 1).

<sup>2</sup>The reason for changing to this new notation is to reserve  $\oplus$  for the following.

<sup>3</sup>In the algebra of operators, we called such an element a projection operator.

### 24.4.1 Group Algebra and Representations

Group algebra is very useful for the construction and analysis of representations of groups. In fact, we have already used a similar approach in the construction of the regular representation. Instead of  $\mathbb{R}^m$  used before, use the  $m$ -dimensional vector space  $\mathcal{A}$ , the group algebra. Then left-multiplication by a group element  $g$  can be identified with  $T_g^{(R)}$ , the operators of the regular representation, and the invariant subspaces of  $\mathcal{A}$  become the left ideals of  $\mathcal{A}$ , and we can write

$$\mathcal{A} = \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \cdots \oplus \mathcal{L}_r.$$

Moreover, since the identity element of the group is the identity element of the algebra as well, the argument at the end of the last subsection gives the resolution<sup>4</sup> of the identity

$$e = e_1 + \cdots + e_r \quad e_i^2 = e_i, \quad e_i e_j = 0 \quad \text{for } i \neq j. \quad (24.21)$$

essentially idempotent elements

It is clear that if  $\mathbf{a}^2 = \alpha \mathbf{a}$ , then  $\mathbf{a}/\alpha$  will be idempotent. So, we can essentially ignore the constant  $\alpha$ , which is why  $\mathbf{a}$  is called **essentially idempotent**. Now consider the element of the group algebra

$$P = \sum_{x \in G} x \quad (24.22)$$

and note that  $gP = \sum_{x \in G} gx = P$ . It follows that

$$P^2 = \sum_{g \in G} g \sum_{x \in G} x = \sum_{g \in G} \sum_{x \in G} gx = \sum_{g \in G} P = |G|P.$$

So,  $P$  is essentially idempotent. Furthermore, the reader may verify that the ideal generated by  $P$  is one-dimensional.

primitive idempotents

An idempotent that cannot be resolved into other idempotents satisfying Equation (24.21) is called a **primitive idempotent**. The reader may check that the following holds.

**24.4.2. Proposition.** *A left ideal is minimal if and only if it is generated by a primitive idempotent.*

Let us now apply the notion of the group algebra to derive further relations among characters. Denote the elements of the  $i$ th class  $K_i$  of  $G$  by  $\{x_l^{(i)}\}_{l=1}^{c_i}$  and construct the element of the group algebra  $\kappa_i \equiv \sum_{l=1}^{c_i} x_l^{(i)}$ . If in the product of two such quantities

$$\kappa_i \kappa_j = \sum_{l=1}^{c_i} \sum_{m=1}^{c_j} x_l^{(i)} x_m^{(j)}, \quad (24.23)$$

<sup>4</sup>In the algebra of operators, we called elements satisfying these relations orthogonal projection operators.

$x_l^{(i)} x_m^{(j)} \equiv y \in G$ , is in a certain conjugacy class, then the rest of that class can be obtained by taking all conjugates of  $y$ , i.e., elements of  $G$  that can be written as

$$gyg^{-1} = gx_l^{(i)} x_m^{(j)} g^{-1} = \underbrace{gx_l^{(i)} g^{-1}}_{\in K_i} \underbrace{gx_m^{(j)} g^{-1}}_{\in K_j}.$$

It follows that if one member of a class appears in the double sum of Equation (24.23), all members will appear there. The reader may check that if  $y$  occurs  $k$  times in the double sum, then all members of the class of  $y$  occur  $k$  times as well. Collecting all such members together, we can write

$$\kappa_i \kappa_j = \sum_{l=1}^r c_{ijl} \kappa_l, \quad (24.24)$$

where  $c_{ijl}$  are positive integers.

Now consider the  $\alpha$ th irreducible representation, and add all operators corresponding to a given class:

$$\mathbf{T}_i^{(\alpha)} \equiv \sum_{g \in K_i} \mathbf{T}_g^{(\alpha)} \Rightarrow \mathbf{T}_i^{(\alpha)} \mathbf{T}_j^{(\alpha)} = \sum_{l=1}^r c_{ijl} \mathbf{T}_l^{(\alpha)}, \quad (24.25)$$

where the second equation follows from the same sort of argument used above to establish Equation (24.24). One can show that  $\mathbf{T}_i^{(\alpha)}$  commutes with all  $\mathbf{T}_g^{(\alpha)}$ . Therefore, by Schur's lemma,  $\mathbf{T}_i^{(\alpha)} = \lambda_i^{(\alpha)} \mathbf{1}$ , and the second equation in (24.25) becomes

$$\lambda_i^{(\alpha)} \lambda_j^{(\alpha)} = \sum_{l=1}^r c_{ijl} \lambda_l^{(\alpha)}. \quad (24.26)$$

Taking the characters of both sides of  $\mathbf{T}_i^{(\alpha)} = \lambda_i^{(\alpha)} \mathbf{1}$  and using the first equation in (24.25), noting that all elements of a class have the same character, we get

$$c_i \chi_i^{(\alpha)} = \lambda_i^{(\alpha)} n_\alpha \Rightarrow \lambda_i^{(\alpha)} = \frac{c_i \chi_i^{(\alpha)}}{n_\alpha}.$$

Substituting this in Equation (24.26), we obtain

$$c_i c_j \chi_i^{(\alpha)} \chi_j^{(\alpha)} = n_\alpha \sum_{l=1}^r c_{ijl} c_l \chi_l^{(\alpha)}. \quad (24.27)$$

This is another equation that is useful for computing characters. Note that this equation connects the purely group properties ( $c_i$ 's and  $c_{ijl}$ 's) with the properties

of the representation  $(\chi_i^{(\alpha)}$ 's and  $n_\alpha$ ). Summing Equation (24.27) over  $\alpha$  and using the first equation in (24.18), we get

$$c_i c_j \sum_{\alpha=1}^{\rho} \chi_i^{(\alpha)} \chi_j^{(\alpha)} = \sum_{l=1}^r c_{ijl} c_l \underbrace{\sum_{\alpha=1}^{\rho} n_\alpha \chi_l^{(\alpha)}}_{=|G|\delta_{l1} \text{ by (24.18)}} = c_{ij1}|G|$$

because  $c_1 = 1$  (there is only one element in the class of the identity). Problem 24.12 shows that  $c_{ij1} = c_i \delta_{i'j}$  where  $K_{i'}$  is the class consisting of inverses of elements of  $K_i$ . It then follows that

$$\sum_{\alpha=1}^{\rho} \chi_i^{(\alpha)} \chi_j^{(\alpha)} = \frac{|G|}{c_j} \delta_{i'j}. \tag{24.28}$$

For a unitary representation,  $\chi_{i'}^{(\alpha)} = \chi_i^{(\alpha)*}$ , so Equation (24.28) becomes

$$\sum_{\alpha=1}^{\rho} \chi_i^{(\alpha)} \chi_j^{(\alpha)*} = \frac{|G|}{c_j} \delta_{ij} \Rightarrow \langle \chi_j | \chi_i \rangle = \frac{|G|}{c_j} \delta_{ij}, \tag{24.29}$$

where  $|\chi_i\rangle \in \mathbb{C}^\rho$  is a  $\rho$ -dimensional vector with components  $\{\chi_i^{(\alpha)}\}_{\alpha=1}^{\rho}$ . This equation can also be written in terms of group elements rather than classes. Since  $\chi_i^{(\alpha)} = \chi^{(\alpha)}(x)$  for any  $x \in K_i$ , we have

$$\sum_{\alpha=1}^{\rho} \chi^{(\alpha)}(x) \chi^{(\alpha)*}(y) = \frac{|G|}{|K_x^G|} \delta(K_x^G, K_y^G), \tag{24.30}$$

where  $K_x^G$  is the conjugacy class of  $G$  containing  $x$ ,  $|K_x^G|$  is the number of its elements, and

$$\delta(K_x^G, K_y^G) = \begin{cases} 1 & \text{if } K_x^G = K_y^G, \\ 0 & \text{otherwise.} \end{cases}$$

Equation (24.29) shows that the  $r$   $\rho$ -dimensional vectors  $\chi_i^{(\alpha)}$  are mutually orthogonal; therefore,  $r \leq \rho$ . The statement after Equation (24.12) was that  $r \geq \rho$ . We thus have the following:

**24.4.3. Theorem.** *The number of inequivalent irreducible representations of a finite group is equal to the number of conjugacy classes in the group.*

It is convenient to summarize our result in a square table with rows labeled by the irreducible representation and columns labeled by the conjugacy classes of  $G$ . Then on the  $\alpha$ th row and  $i$ th column we list  $\chi_i^{(\alpha)}$ , and we get Table 24.1, called the **character table** of  $G$ . Note that  $c_i$ , the order of  $K_i$ , is written as a left superscript.

character table of a  
finite group

	$c_1 K_1$	$c_2 K_2$	...	$c_i K_i$	...	$c_r K_r$
$T^{(1)}$	$\chi_1^{(1)}$	$\chi_2^{(1)}$	...	$\chi_i^{(1)}$	...	$\chi_r^{(1)}$
$T^{(2)}$	$\chi_1^{(2)}$	$\chi_2^{(2)}$	...	$\chi_i^{(2)}$	...	$\chi_r^{(2)}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$T^{(\alpha)}$	$\chi_1^{(\alpha)}$	$\chi_2^{(\alpha)}$	...	$\chi_i^{(\alpha)}$	...	$\chi_r^{(\alpha)}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$T^{(r)}$	$\chi_1^{(r)}$	$\chi_2^{(r)}$	...	$\chi_i^{(r)}$	...	$\chi_r^{(r)}$

**Table 24.1** A typical character table.

Character tables have the property that any two of their rows are orthogonal in the sense of Equation (24.12), and any two of their columns are orthogonal in the sense of Equation (24.29).

If all inequivalent irreducible representations of a group  $G$  have dimension one, then there will be  $|G|$  of them [by Equation (24.18)]. Hence, there will be  $|G|$  conjugacy classes; i.e., each class consists of a single element. By Problem 23.16, the group must be abelian. Combining this with Theorem 24.2.3, we have the following theorem.

**24.4.4. Theorem.** *A finite group is abelian if and only if all its inequivalent irreducible representations are one-dimensional.*

### 24.5 Relationship of Characters to Those of a Subgroup

Let  $H$  be a subgroup of  $G$ . Denote by  $K_h^H$  and  $K_g^G$  the  $H$ -class containing  $h \in H$  and the  $G$ -class containing  $g$ , respectively. Let  $d_j$  and  $c_i$  be the number of elements in the  $j$ th  $H$ -class and  $i$ th  $G$ -class, respectively. Any representation of  $G$  defines a representation of  $H$  by restriction. An irreducible representation of  $G$  may be reducible as a representation of  $H$ . This is because although the subspace  $\mathcal{W}^{(\alpha)}$  of the carrier space that is irreducible under  $G$  is the smallest such subspace containing a given vector, it is possible to generate a smaller subspace by applying a subset of the operators  $T_g$  corresponding to those  $g$ 's that belong to  $H$ . It follows that

$$T^{(\alpha)}(h) = \sum_{\sigma} m_{\alpha\sigma} t^{(\sigma)}(h), \quad h \in H, \tag{24.31}$$

where  $m_{\alpha\sigma}$  are nonnegative integers as in Equation (24.14) and  $t^{(\sigma)}$  are irreducible representations of  $H$ . If  $\chi^{(\alpha)}$  and  $\xi^{(\sigma)}$  denote the characters of irreducible representations of  $G$  and  $H$ , respectively, then the equivalent equation for the characters

is

$$\chi^{(\alpha)}(h) = \sum_{\sigma} m_{\alpha\sigma} \xi^{(\sigma)}(h), \quad h \in H. \quad (24.32)$$

Multiply both sides by  $\xi^{(\kappa)*}(h)$ , sum over  $h \in H$ , and take the complex conjugate at the end. Then by the orthogonality relation (24.11), applied to  $H$ , we obtain

$$m_{\alpha\kappa} = \frac{1}{|H|} \sum_{h \in H} \chi^{(\alpha)*}(h) \xi^{(\kappa)}(h). \quad (24.33)$$

Now multiply both sides of Equation (24.33) by  $\chi^{(\alpha)}(g)$ , sum over  $\alpha$ , and use Equation (24.30) to obtain

$$\sum_{\alpha} m_{\alpha\kappa} \chi^{(\alpha)}(g) = \frac{|G|}{|H| |K_g^G|} \sum_{h \in H} \delta(K_h^G, K_g^G) \xi^{(\kappa)}(h). \quad (24.34)$$

The sum on the right can be transformed into a sum over conjugacy classes of  $H$ . Then Equation (24.34) becomes

$$\sum_{\alpha} m_{\alpha\kappa} \chi_i^{(\alpha)} = \frac{|G|}{|H|c_i} \sum_j d_j \xi_j^{(\kappa)}, \quad i = 1, 2, \dots, r, \quad (24.35)$$

where the sum on the LHS is over irreducible representations of  $G$ , and on the RHS it is over those  $H$ -classes  $j$  that lie in the  $i$ th  $G$ -class. Note that the coefficients  $|G|d_j/(|H|c_i)$  are integers by Problem 23.17.

Equations (24.34) and (24.35) are useful for obtaining characters of  $G$  when those of a subgroup  $H$  are known. The general procedure is to note that the RHS of these equations are completely determined by the structure of the group  $G$  and the characters of  $H$ . Varying  $i$ , the RHS of (24.35) determines the  $r$  components of a (compound) character  $|\psi\rangle$ , which, by the LHS, can be written as a linear combination of characters of  $G$ :

$$|\psi\rangle \equiv \sum_{\alpha=1}^r m_{\alpha} |\chi^{(\alpha)}\rangle, \quad (24.36)$$

where we have suppressed the irrelevant subscript  $\kappa$ . If we know some of the  $|\chi^{(\alpha)}\rangle$ 's, we may be able to determine the rest by taking successive inner products to find the integers  $m_{\alpha}$ , and subtracting each irreducible factor of the sum from the LHS. We illustrate this procedure for  $S_n$  in the following example.

**24.5.1. Example.** Let  $K_1 = (1^2)$  and  $K_2 = (2)$  for  $S_2$  (see Section 23.4 for notation). Example 24.1.13 showed that we can construct two irreducible representations for any  $S_n$ , the symmetric and the antisymmetric representations. The reader may verify that these two representations are inequivalent. Since the number of inequivalent irreducible representations is equal to the number of classes in a group, we have all the information needed to

	${}^1K_1$	${}^1K_2$
$T^{(1)}$	1	1
$T^{(2)}$	1	-1

**Table 24.2** Character table for  $S_2$ .

	${}^1K_1$	${}^3K_2$	${}^2K_3$
$T^{(1)}$	1	1	1
$T^{(2)}$	1	-1	1
$T^{(3)}$	?	?	?

**Table 24.3** Partially filled character table for  $S_3$ .

construct the character table for  $S_2$ . Table 24.2 shows this character table. We want to use the  $S_2$  character table to construct the character table for  $S_3$ . With our knowledge of the symmetric and the antisymmetric representations, we can partially fill in the  $S_3$  character table. Let  $K_1 = (1^3)$ ,  $K_2 = (2, 1)$ , and  $K_3 = (3)$  and note that  $c_1 = 1$ ,  $c_2 = 3$ , and  $c_3 = 2$ . Then we obtain Table 24.3. To complete the table, we start with  $\kappa = 1$ , and write the RHS of Equation (24.35) as

$$\psi_i = \frac{6}{2c_i} \sum_j d_j \xi_j^{(1)} = \frac{3}{c_i} \sum_j \xi_j^{(1)}$$

because  $d_j = 1$  for the two classes of  $S_2$ . The sum on the RHS is over  $S_2$ -classes that are inside the  $i$ th  $S_3$ -class. For  $i = 1$ , only the the first  $S_2$ -class contributes. Noting that  $\xi_j^{(\kappa)}$  are the entries of Table 24.2, we get

$$\psi_1 = \frac{3}{c_1} \xi_1^{(1)} = \frac{3}{1} \cdot 1 = 3.$$

Similarly,

$$\psi_2 = \frac{3}{c_2} \xi_2^{(1)} = \frac{3}{3} \cdot 1 = 1 \quad \text{and} \quad \psi_3 = \frac{3}{c_3} \cdot 0 = 0.$$

The second equation follows from the fact that there are no classes of  $S_2$  inside the third class of  $S_3$ . Equation (24.36) now gives

$$|\psi\rangle = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} = \sum_{\alpha=1}^r m_\alpha |\chi^{(\alpha)}\rangle.$$

We can find the number of times  $|\chi^{(1)}\rangle$  occurs in this compound character by taking the inner product:

$$\langle \chi^{(1)} | \psi \rangle = \sum_{\alpha=1}^r m_\alpha \langle \chi^{(1)} | \chi^{(\alpha)} \rangle = m_1 |G| = 6m_1.$$



	${}^1K_1$	${}^3K_2$	${}^2K_3$
$T^{(1)}$	1	1	1
$T^{(2)}$	1	-1	1
$T^{(3)}$	2	0	-1

**Table 24.4** Complete character table for  $S_3$ .

But

$$\langle \chi^{(1)} | \psi \rangle = \sum_{i=1}^3 c_i \chi_i^{(1)} \psi_i = 1 \cdot 1 \cdot 3 + 3 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 0 = 6.$$

These two equations show that  $m_1 = 1$ . So,

$$\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + m_2 |\chi^{(2)}\rangle + m_3 |\chi^{(3)}\rangle.$$

Subtracting the column vectors, we get a new character:

$$|\psi'\rangle \equiv \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} = m_2 |\chi^{(2)}\rangle + m_3 |\chi^{(3)}\rangle.$$

Taking the inner product with  $|\chi^{(2)}\rangle$  yields  $m_2 = 0$ . It follows that  $|\psi'\rangle$  is a simple character. In fact,

$$\sum_i c_i |\psi'_i|^2 = 1 \cdot 2^2 + 3 \cdot 0^2 + 2 \cdot (-1)^2 = 6,$$

and the criterion of irreducibility, Equation (24.17), is satisfied.

We can now finish up Table 24.3 to obtain Table 24.4, which is the complete character table for  $S_3$ . ■

## 24.6 Irreducible Basis Functions

We have studied the operators  $T_g$  and their characters representing group elements in rather extensive detail. Let us now turn our attention to the carrier space itself. In particular, we want to concentrate on the basis functions of the irreducible representations. We choose “functions,” rather than vectors, because of their use in quantum mechanics as discussed at the beginning of this chapter.

Let  $\{|\psi_i^{(\alpha)}\rangle\}_{i=1}^{n_\alpha}$  be a set of basis functions for  $\mathcal{W}^{(\alpha)}$ , the  $\alpha$ th invariant irreducible subspace. Invariance of  $\mathcal{W}^{(\alpha)}$  implies that

$$T_g |\psi_i^{(\alpha)}\rangle = \sum_{j=1}^{n_\alpha} T_{ji}^{(\alpha)}(g) |\psi_j^{(\alpha)}\rangle,$$

functions belonging to the  $i$ th row of the  $\alpha$ th irreducible representation

where  $T_{ji}^{(\alpha)}(g)$  are elements of the matrix  $T_g^{(\alpha)}$  representing  $g \in G$ .

**24.6.1. Definition.** A function (or vector)  $|\phi_i^{(\alpha)}\rangle$  is said to **belong to the  $i$ th row of the  $\alpha$ th irreducible representation** (or to transform according to the  $i$ th row of the  $\alpha$ th irreducible representation) if there exists a basis  $\{|\psi_i^{(\alpha)}\rangle\}_{i=1}^{n_\alpha}$  of the  $\alpha$ th irreducible representation of  $G$  with matrices  $(T_{ji}^{(\alpha)}(g))$  and  $n_\alpha - 1$  other functions  $\{|\phi_j^{(\alpha)}\rangle\}$  such that

$$T_g |\phi_i^{(\alpha)}\rangle = \sum_{j=1}^{n_\alpha} T_{ji}^{(\alpha)}(g) |\phi_j^{(\alpha)}\rangle. \tag{24.37}$$

Functions that belong to rows of irreducible representations have some remarkable properties. Let  $|\psi_i^{(\alpha)}\rangle$  and  $|\phi_j^{(\beta)}\rangle$  transform according to the  $i$ th and  $j$ th rows of the  $\alpha$ th and  $\beta$ th irreducible representations, respectively. Choose an inner product for the carrier space such that all representations are unitary. Then we have

$$\begin{aligned} \langle \psi_i^{(\alpha)} | \phi_j^{(\beta)} \rangle &= \langle T_g \psi_i^{(\alpha)} | T_g \phi_j^{(\beta)} \rangle \\ &= \sum_{l=1}^{n_\alpha} \sum_{m=1}^{n_\beta} T_{li}^{(\alpha)*}(g) T_{mj}^{(\beta)}(g) \langle \psi_l^{(\alpha)} | \phi_m^{(\beta)} \rangle. \end{aligned}$$

Summing this equation over  $g$  yields  $|G| \langle \psi_i^{(\alpha)} | \phi_j^{(\beta)} \rangle$  for the LHS, while

$$\begin{aligned} RHS &= \sum_{l=1}^{n_\alpha} \sum_{m=1}^{n_\beta} \overbrace{\sum_{g \in G} T_{li}^{(\alpha)*}(g) T_{mj}^{(\beta)}(g)}^{(|G|/n_\alpha) \delta_{\alpha\beta} \delta_{lm} \delta_{ij}} \langle \psi_l^{(\alpha)} | \phi_m^{(\beta)} \rangle \\ &= \frac{|G|}{n_\alpha} \delta_{\alpha\beta} \delta_{ij} \sum_{l=1}^{n_\alpha} \langle \psi_l^{(\alpha)} | \phi_l^{(\beta)} \rangle, \end{aligned}$$

where we have made use of Equation (24.8). Therefore,

$$\langle \psi_i^{(\alpha)} | \phi_j^{(\beta)} \rangle = \frac{1}{n_\alpha} \delta_{\alpha\beta} \delta_{ij} \sum_{l=1}^{n_\alpha} \langle \psi_l^{(\alpha)} | \phi_l^{(\alpha)} \rangle. \tag{24.38}$$

This shows that functions belonging to different irreducible representations are orthogonal. We should expect this, because in our construction of invariant irreducible subspaces, we kept dividing the whole space into orthogonal complements. What is surprising is that *functions transforming according to different rows of an irreducible representation are orthogonal*. We had no control over this property! It is a consequence of Equation (24.37). Another surprise is the independence of the inner product from  $i$ : If we let  $i = j$  and  $\alpha = \beta$  on both sides of (24.38), we obtain

$$\langle \psi_i^{(\alpha)} | \phi_i^{(\alpha)} \rangle = \frac{1}{n_\alpha} \sum_{l=1}^{n_\alpha} \langle \psi_l^{(\alpha)} | \phi_l^{(\alpha)} \rangle, \tag{24.39}$$

symmetry and the quantum mechanical perturbation theory; lifting of degeneracy

which indicates that the inner product on the LHS is independent of  $i$ .

**24.6.2. Example.** The quantum-mechanical perturbation theory starts with a known Hamiltonian  $H_0$  with eigenvalues  $E_i$  and the corresponding eigenstates  $|E_i\rangle$ . Subsequently, a (small) perturbing “potential”  $V$  is added to the Hamiltonian, and the eigenvalues and eigenstates of the new Hamiltonian  $H = H_0 + V$  are sought. One can draw important conclusions about the eigenvalues and eigenstates of the total Hamiltonian by symmetry arguments.

Suppose the symmetry group of  $H_0$  is  $G$ , and that of  $H$  is  $H$ , which has to be a subgroup of  $G$ . In most cases, the eigenspaces of  $H_0$  are irreducible carrier spaces of  $G$ , i.e., their basis vectors transform according to the rows of irreducible representations of  $G$ . If  $H$  is a proper subgroup of  $G$ , then the eigenspaces of  $H_0$  will split according to Equation (24.31). We say that some of the *degeneracy is lifted* because of the perturbation  $V$ . The nature of the split, i.e., the number and the dimensionality of the vector spaces into which a given eigenspace splits, can be obtained by the characters of  $G$  and  $H$  and Equation (24.32). The original eigenspaces are represented on an *energy diagram* with a line corresponding to each eigenspace. The split of the eigenspace into  $k$  new subspaces is then indicated by the branching of the old line into  $k$  new lines.

To the lowest approximation—first-order perturbation theory—the magnitude of the split, i.e., the difference between the eigenvalues of  $H_0$  and those of  $H$ , is given by [see Equation (21.57)] the expectation value  $\langle \phi_i^{(\alpha)} | V | \phi_j^{(\alpha)} \rangle$ , where  $|\phi_i^{(\alpha)}\rangle$  belongs to the  $i$ th row of the  $\alpha$ th irreducible representation, and  $|\phi_j^{(\alpha)}\rangle$  to its  $j$ th row ( $i \neq j$ ). Only if this expectation value is nonzero will a split occur. This, in turn, depends on the symmetry of  $V$ : If  $V$  is *at least* as symmetric as  $H_0$  (corresponding to  $G = H$ ), then  $\langle \phi_i^{(\alpha)} | V | \phi_j^{(\alpha)} \rangle = 0$ , and no splitting occurs (Problem 24.17). If, on the other hand,  $V$  is less symmetric than  $H_0$  (corresponding to  $H \subset G$ ), then  $V |\phi_j^{(\alpha)}\rangle$  will not belong to the  $j$ th row of the  $\alpha$ th irreducible representation, and in general,  $\langle \phi_i^{(\alpha)} | V | \phi_j^{(\alpha)} \rangle \neq 0$ . ■

We had decomposed the carrier space  $\mathcal{V}$  of a representation into invariant irreducible subspaces  $\mathcal{W}^{(\alpha)}$ . The argument above shows that each  $\mathcal{W}^{(\alpha)}$  has a basis consisting of the “rows” of the irreducible representations. Corresponding to such a basis, there is a set of projection operators  $\mathbf{P}_i^{(\alpha)}$  with the property  $\sum_{\alpha,i} \mathbf{P}_i^{(\alpha)} = 1$  (Chapter 4). Our aim is to find an expression for these operators, which have the defining property  $\mathbf{P}_i^{(\alpha)} |\psi_i^{(\alpha)}\rangle = |\psi_i^{(\alpha)}\rangle$ . We start with Equation (24.37), multiply both sides of it by  $T_{lm}^{(\beta)*}(g)$ , sum over  $g \in G$ , and use Equation (24.8) to obtain

$$\begin{aligned} \sum_{g \in G} T_{lm}^{(\beta)*}(g) \mathbf{T}_g |\psi_i^{(\alpha)}\rangle &= \sum_{j=1}^{n_\alpha} |\psi_j^{(\alpha)}\rangle \sum_{g \in G} T_{lm}^{(\beta)*}(g) T_{ji}^{(\alpha)}(g) \\ &= \frac{|G|}{n_\alpha} \sum_{j=1}^{n_\alpha} |\psi_j^{(\alpha)}\rangle \delta_{lj} \delta_{mi} \delta_{\alpha\beta} = \frac{|G|}{n_\alpha} |\psi_i^{(\alpha)}\rangle \delta_{mi} \delta_{\alpha\beta}. \end{aligned}$$

Let  $\beta = \alpha$ ,  $m = l = i$ , and multiply both sides by  $n_\alpha/|G|$ . Then this equation becomes

$$\frac{n_\alpha}{|G|} \sum_{g \in G} T_{ii}^{(\alpha)*}(g) \mathbf{T}_g |\psi_i^{(\alpha)}\rangle = |\psi_i^{(\alpha)}\rangle,$$

projection operator onto the  $i$ th row of the  $\alpha$ th irreducible representation

which suggests the identification

$$P_i^{(\alpha)} = \frac{n_\alpha}{|G|} \sum_{g \in G} T_{ii}^{(\alpha)*}(g) T_g \tag{24.40}$$

with the properties

$$P_i^{(\alpha)} |\psi_j^{(\beta)}\rangle = |\psi_i^{(\alpha)}\rangle \delta_{ij} \delta_{\alpha\beta}, \quad P_i^{(\alpha)} |\phi\rangle = |\phi_i^{(\alpha)}\rangle, \tag{24.41}$$

where  $|\phi_i^{(\alpha)}\rangle$  is the projection of  $|\phi\rangle$  along the  $i$ th row of the  $\alpha$ th irreducible representation.

We are also interested in the projection operator that projects onto the irreducible subspace  $\mathcal{W}^{(\alpha)}$ . Such an operator is obtained by summing  $P_i^{(\alpha)}$  over  $i$ . We thus obtain

projection operator onto the  $\alpha$ th irreducible representation

$$P^{(\alpha)} = \frac{n_\alpha}{|G|} \sum_{g \in G} \underbrace{\sum_{i=1}^{n_\alpha} T_{ii}^{(\alpha)*}(g)}_{=\chi^{(\alpha)*}(g)} T_g = \frac{n_\alpha}{|G|} \sum_{g \in G} \chi^{(\alpha)*}(g) T_g \tag{24.42}$$

and

$$P^{(\alpha)} |\psi^{(\beta)}\rangle = |\psi^{(\alpha)}\rangle \delta_{\alpha\beta} \quad P^{(\alpha)} |\phi\rangle = |\phi^{(\alpha)}\rangle, \tag{24.43}$$

where  $|\phi^{(\alpha)}\rangle$  is the projection of  $|\phi\rangle$  onto the  $\alpha$ th irreducible invariant subspace. These formulas are extremely useful in identifying the irreducible subspaces of a given carrier space: Start with a basis  $\{|a_i\rangle\}$  of the carrier space, apply  $P^{(\alpha)}$  to all basis vectors, and collect all the linearly independent vectors of the form  $P^{(\alpha)} |a_i\rangle$ . These vectors form a basis of the  $\alpha$ th irreducible representation. The following example illustrates this point.

**24.6.3. Example.** Consider the representation of  $S_3$  given in Example 24.1.8, where the carrier space is the span of the three functions  $|\psi_1\rangle = xy$ ,  $|\psi_2\rangle = yz$ , and  $|\psi_3\rangle = xz$ .

We refer to the character table for  $S_3$  (Table 24.4) and use Equation (24.42) to obtain

$$\begin{aligned} P^{(1)} &= \frac{1}{6} (T_1 + T_2 + T_3 + T_4 + T_5 + T_6), \\ P^{(2)} &= \frac{1}{6} (T_1 - T_2 - T_3 - T_4 + T_5 + T_6), \\ P^{(3)} &= \frac{2}{6} (2T_1 - T_5 - T_6), \end{aligned}$$

where, as in Example 24.1.8, we have used the notation  $T_i$  for  $T_{\pi_i}$ , and the result  $n_1 = n_2 = 1$  and  $n_3 = 2$  obtained from Equation (24.18), Theorem 24.4.4, and the fact that  $S_3$  is nonabelian.

To get the first irreducible subspace of this representation, we apply  $P^{(1)}$  to  $|\psi_1\rangle$ . Since this subspace is one-dimensional, the procedure will give a basis for it if the vector so

obtained is nonzero:

$$\begin{aligned}\mathbf{P}^{(1)}|\psi_1\rangle &= \frac{1}{6}(\mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 + \mathbf{T}_4 + \mathbf{T}_5 + \mathbf{T}_6)|\psi_1\rangle \\ &= \frac{1}{6}(|\psi_1\rangle + |\psi_1\rangle + |\psi_2\rangle + |\psi_3\rangle + |\psi_3\rangle + |\psi_2\rangle) = \frac{1}{3}(|\psi_1\rangle + |\psi_2\rangle + |\psi_3\rangle).\end{aligned}$$

This is a basis for the carrier space of the irreducible identity representation.

For the second irreducible representation, we get

$$\mathbf{P}^{(2)}|\psi_1\rangle = \frac{1}{6}(|\psi_1\rangle - |\psi_1\rangle - |\psi_2\rangle - |\psi_3\rangle + |\psi_3\rangle + |\psi_2\rangle) = 0.$$

Similarly,  $\mathbf{P}^{(2)}|\psi_2\rangle = 0$  and  $\mathbf{P}^{(2)}|\psi_3\rangle = 0$ . This means that  $T^{(2)}$  is not included in the representation we are working with. We should have expected this, because if this one-dimensional irreducible representation were included, it would force the last irreducible representation to be one-dimensional as well [see Equation (24.18)], and, by Theorem 24.4.4, the group  $S_3$  to be abelian!

The last irreducible representation is obtained similarly. We have

$$\begin{aligned}\mathbf{P}^{(3)}|\psi_1\rangle &= \frac{1}{3}(2\mathbf{T}_1 - \mathbf{T}_5 - \mathbf{T}_6)|\psi_1\rangle = \frac{1}{3}(2|\psi_1\rangle - |\psi_3\rangle - |\psi_2\rangle), \\ \mathbf{P}^{(3)}|\psi_2\rangle &= \frac{1}{3}(2\mathbf{T}_1 - \mathbf{T}_5 - \mathbf{T}_6)|\psi_2\rangle = \frac{1}{3}(2|\psi_2\rangle - |\psi_1\rangle - |\psi_3\rangle).\end{aligned}$$

These two vectors are linearly independent. Therefore, they form a basis for the last irreducible representation. The reader may check that  $\mathbf{P}^{(3)}|\psi_3\rangle$  is a linear combination of  $\mathbf{P}^{(3)}|\psi_1\rangle$  and  $\mathbf{P}^{(3)}|\psi_2\rangle$ . ■

## 24.7 Tensor Product of Representations

A simple quantum mechanical system possessing a group of symmetry is described by vectors that transform irreducibly (or according to a row of an irreducible representation). For example, a rotationally invariant system can be described by an eigenstate of angular momentum, the generator of rotation.<sup>5</sup> These eigenstates transform as rows of irreducible representations of the rotation group. At a more fundamental level, the very concept of a particle or field is thought of as states that transform irreducibly under the fundamental group of spacetime, the Poincaré group.

Often these irreducible states are “combined” to form new states. For example, the state of two (noninteracting) particles is described by a two-particle state, labeled by the combined eigenvalues of the two sets of operators that describe each particle separately. In the case of angular momentum, the single-particle states may be labeled as  $|l_i, m_i\rangle$  for  $i = 1, 2$ . Then the combined state will be labeled as  $|l_1, m_1; l_2, m_2\rangle$ , and one can define an action of the rotation group on

<sup>5</sup>Chapter 27 will make explicit the connection between groups and their generators.

the vector space spanned by these combined states to construct a representation. We now describe the way in which this is done.

Suppose that  $T : G \rightarrow GL(\mathcal{V})$  and  $S : G \rightarrow GL(\mathcal{W})$  are two representations of a group  $G$ . Let  $\mathcal{V} \otimes \mathcal{W}$  be the tensor product of  $\mathcal{V}$  and  $\mathcal{W}$  (see Example 1.3.19). Now define an action of the group  $G$  on  $\mathcal{V} \otimes \mathcal{W}$  via the representation  $T \otimes S : G \rightarrow GL(\mathcal{V} \otimes \mathcal{W})$  given by

$$T \otimes S(g)(|v\rangle, |w\rangle) = (T(g)|v\rangle, S(g)|w\rangle).$$

We note that

$$\begin{aligned} T \otimes S(g_1 g_2)(|v\rangle, |w\rangle) &= (T(g_1 g_2)|v\rangle, S(g_1 g_2)|w\rangle) = (T(g_1)T(g_2)|v\rangle, S(g_1)S(g_2)|w\rangle) \\ &= T \otimes S(g_1)(T(g_2)|v\rangle, S(g_2)|w\rangle) = T \otimes S(g_1)T \otimes S(g_2)(|v\rangle, |w\rangle). \end{aligned}$$

Kronecker product representation

It follows that  $T \otimes S$  is indeed a representation, called the **tensor product** or **direct product** or **Kronecker product** representation. It is common, especially in the physics literature, to write  $|v, w\rangle$ , or simply  $|vw\rangle$  for  $(|v\rangle, |w\rangle)$ , and  $TS$  for  $T \otimes S$ . If we choose the orthonormal bases  $\{|v_i\rangle\}$  for  $\mathcal{V}$  and  $\{|w_a\rangle\}$  for  $\mathcal{W}$ , and define an inner product on  $\mathcal{V} \otimes \mathcal{W}$  by

$$\langle v, w | v', w' \rangle \equiv \langle v | v' \rangle \langle w | w' \rangle,$$

we obtain a matrix representation of the group with matrix elements given by

$$\begin{aligned} (T \otimes S)_{ia, jb}(g) &\equiv \langle v_i, w_a | T \otimes S(g) | v_j, w_b \rangle \\ &= \langle v_i | T(g) | v_j \rangle \langle w_a | S(g) | w_b \rangle \equiv T_{ij}(g) S_{ab}(g). \end{aligned}$$

Note that the rows and columns of this matrix are distinguished by double indices. If the matrix  $T$  is  $m \times m$  and  $S$  is  $n \times n$ , then the matrix  $T \otimes S$  is  $(mn) \times (mn)$ . The character of the tensor product representation is

character of a product representation is a product of characters

$$\begin{aligned} \chi^{T \otimes S}(g) &= \sum_{i,a} (T \otimes S)_{ia, ia}(g) = \sum_{i,a} T_{ii}(g) S_{aa}(g) = \sum_i T_{ii}(g) \sum_a S_{aa}(g) \\ &= \chi^T(g) \chi^S(g) \Rightarrow \chi_i^{T \otimes S} = \chi_i^T \cdot \chi_i^S. \end{aligned} \tag{24.44}$$

So the character of the tensor product is the product of the individual characters. An important special case is the tensor product of a representation with itself. For such a representation, the matrix elements satisfy the symmetry relation  $(T \otimes T)_{ia, jb}(g) = (T \otimes T)_{ai, bj}(g)$ . This symmetry can be used to decompose the tensor product space into two subspaces that are separately invariant under the action of the group. To do this, take the span of all the symmetric vectors of the form  $(|v_i w_j\rangle + |v_j w_i\rangle) \in \mathcal{V} \otimes \mathcal{V}$  and denote it by  $(\mathcal{V} \otimes \mathcal{V})_s$ . Similarly, take the span of all the antisymmetric vectors of the form  $(|v_i w_j\rangle - |v_j w_i\rangle) \in \mathcal{V} \otimes \mathcal{V}$  and denote it by  $(\mathcal{V} \otimes \mathcal{V})_a$ . Next note that

$$|v_i w_j\rangle = \frac{1}{2}(|v_i w_j\rangle + |v_j w_i\rangle) + \frac{1}{2}(|v_i w_j\rangle - |v_j w_i\rangle).$$

It follows that every vector of the product space can be written as the sum of a symmetric and an antisymmetric vector. Furthermore, the only vector that is both symmetric and antisymmetric is the zero vector. Therefore,

$$\mathcal{V} \otimes \mathcal{V} = (\mathcal{V} \otimes \mathcal{V})_s \oplus (\mathcal{V} \otimes \mathcal{V})_a.$$

Now consider the action of the group on each of these subspaces separately. From the relation

$$\begin{aligned} T \otimes T(g) |v_i w_j\rangle &\equiv T \otimes T(g) (|v_i\rangle, |w_j\rangle) \\ &= \left( \sum_k T_{ki}(g) |v_k\rangle, \sum_l T_{lj}(g) |w_l\rangle \right) \\ &= \sum_{k,l} T_{ki} T_{lj}(g) (|v_k\rangle, |w_l\rangle) = \sum_{k,l} (T \otimes T)_{kl,ij}(g) |v_k w_l\rangle \end{aligned}$$

we obtain

$$\begin{aligned} T \otimes T(g) (|v_i w_j\rangle \pm |v_j w_i\rangle) \\ = \sum_{k,l} [(T \otimes T)_{kl,ij}(g) \pm (T \otimes T)_{kl,ji}(g)] |v_k w_l\rangle. \end{aligned} \tag{24.45}$$

Kronecker product reduces to the symmetric and the antisymmetric representations

Problem 24.21 shows that the RHS can be written as a sum over the symmetric (for the plus sign) or antisymmetric (for the minus sign) vectors alone. It follows that

**24.7.1. Box.** *The Kronecker product of a representation with itself is always reducible into two representations, the symmetrized product and the antisymmetrized product representations.*

### 24.7.1 Clebsch–Gordan Decomposition

A common situation in quantum mechanics is to combine two simple systems into a composite system and see which properties of the original simple systems the composite system retains. For example, combining the angular momenta of two particles gives a new total angular momentum operator. The question of what single-particle angular momentum states are included in the states of the total angular momentum operator is the content of **selection rules** and is of great physical interest: A quark and an antiquark (two fermions) with spin  $\frac{1}{2}$  always combine to form a meson (a boson), because the resulting composite state has no projection onto the subspace spanned by half-integer-spin particles. In this section, we study the mathematical foundation of this situation. The tensor product of two irreducible representations  $T^{(\alpha)}$  and  $T^{(\beta)}$  of  $G$  is denoted by  $T^{(\alpha \times \beta)}$ , and it is,

in general, a reducible representation. The characters, generally compound, are denoted by  $\chi^{(\alpha \times \beta)}$ . Equation (24.14), combined with Equation (24.44), tells us what irreducible representations are present in the tensor product, and therefore onto which irreducible representations the product representation has nonzero projection:

$$\chi_i^{(\alpha \times \beta)} = \chi_i^{(\alpha)} \cdot \chi_i^{(\beta)} = \sum_{\sigma=1}^r m_{\sigma}^{\alpha\beta} \chi_i^{(\sigma)},$$

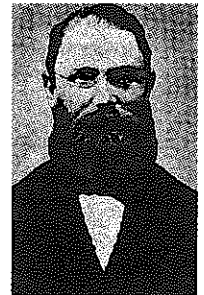
where  $m_{\sigma}^{\alpha\beta}$  are nonnegative integers. We rewrite this more conveniently in terms of vectors as

$$\begin{aligned} |\chi^{(\alpha \times \beta)}\rangle &= \sum_{\sigma=1}^r m_{\sigma}^{\alpha\beta} |\chi^{(\sigma)}\rangle, \\ m_{\sigma}^{\alpha\beta} &= \frac{1}{|G|} \langle \chi^{(\sigma)} | \chi^{(\alpha \times \beta)} \rangle = \frac{1}{|G|} \sum_{i=1}^r c_i \bar{\chi}_i^{(\sigma)} \chi_i^{(\alpha)} \chi_i^{(\beta)}. \end{aligned} \quad (24.46)$$

simply reducible  
group

A group for which  $m_{\sigma}^{\alpha\beta} = 0, 1$  is called **simply reducible**.

**Rudolph Friedrich Alfred Clebsch** (1833–1872) studied mathematics in the shadow of Jacobi at the University of Königsberg, two of his teachers having been students of Jacobi. After graduation he held a number of positions in Germany, including positions at the universities of Berlin, Giessen, and finally Göttingen, where he remained until his death. He and Carl Neumann, son of one of the aforementioned Jacobian teachers, founded the *Mathematische Annalen*.



Clebsch began his career in mathematical physics, producing a doctoral thesis on hydrodynamics and a book on elasticity in which he treated the elastic vibrations of rods and plates. These works were primarily mathematical, however, and he soon turned his attention more to pure mathematics. His links to Jacobi gave rise to his first work in that vein, concerning problems in variational calculus and partial differential equations, in which he surpassed the results of Jacobi's work.

Clebsch first achieved significant recognition for his work in projective invariants and algebraic geometry. He was intrigued by the interplay between algebra and geometry, and, since many results in the theory of invariants have geometric interpretations, the two fields seemed natural choices.

**24.7.2. Example.** Referring to Table 24.8 of Problem 24.15, and using Equation (24.44), we can construct the compound character  $|\chi^{(4 \times 5)}\rangle$  with components 9, -1, 1, 0, -1. Then,



we have

$$|\chi^{(4 \times 5)}\rangle = \begin{pmatrix} 9 \\ -1 \\ 1 \\ 0 \\ -1 \end{pmatrix} = \sum_{\sigma=1}^5 m_{\sigma}^{45} |\chi^{(\sigma)}\rangle, \quad m_{\sigma}^{45} = \frac{1}{24} \langle \chi^{(\sigma)} | \chi^{(4 \times 5)} \rangle.$$

For the first irreducible representation, we get

$$\begin{aligned} m_1^{45} &= \frac{1}{24} \langle \chi^{(1)} | \chi^{(4 \times 5)} \rangle = \frac{1}{24} \sum_{i=1}^5 c_i \chi_i^{(1)*} \chi_i^{(4 \times 5)} \\ &= \frac{1}{24} [1 \cdot 1 \cdot 9 + 6 \cdot 1 \cdot (-1) + 3 \cdot 1 \cdot 1 + 8 \cdot 1 \cdot 0 + 6 \cdot 1 \cdot (-1)] = 0. \end{aligned}$$

For the second irreducible representation, we get

$$\begin{aligned} m_2^{45} &= \frac{1}{24} \langle \chi^{(2)} | \chi^{(4 \times 5)} \rangle \\ &= \frac{1}{24} [1 \cdot 1 \cdot 9 + 6 \cdot (-1) \cdot (-1) + 3 \cdot 1 \cdot 1 + 8 \cdot (-1) \cdot 0 + 6 \cdot (-1) \cdot (-1)] = 1. \end{aligned}$$

Similarly,  $m_3^{45} = 1$ ,  $m_4^{45} = 1$ , and  $m_5^{45} = 1$ . We thus see that the identity representation is not included in the direct product of irreducible representations 4 and 5; all other irreducible representations of  $S_4$  occur once in  $T^{(4 \times 5)}$ . ■

Clebsch–Gordan series

In terms of representations themselves, we have the so-called **Clebsch–Gordan series**

$$T^{(\alpha \times \beta)}(g) = \sum_{\sigma=1}^r m_{\sigma}^{\alpha\beta} T^{(\sigma)}(g), \quad m_{\sigma}^{\alpha\beta} = \frac{1}{|G|} \sum_{i=1}^r c_i \bar{\chi}_i^{(\sigma)} \chi_i^{(\alpha)} \chi_i^{(\beta)}, \quad (24.47)$$

where we have used Equation (24.13)

How to obtain invariants from the product of representations

**24.7.3. Example.** The one-dimensional identity representation plays a special role in the application of group theory to physics because any vector (function) in its carrier space is invariant under the action of the group, and invariant vectors often describe special states of the quantum mechanical systems. For example, the ground state of an atomic system with rotational invariance has zero orbital angular momentum, corresponding to a spherically symmetric state.

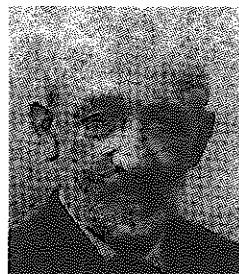
Another example comes from particle physics. Quarks are usually placed in the states of an irreducible representation of a group  $[SU(n)]$ , where  $n$  is the number of “flavors” such as up, down, charm, and antiquarks in its adjoint. A question of great importance is what combination of quarks and antiquarks leads to particles—called singlets—that are an invariant of the group. For the case of quark-antiquark combination, the answer comes in the analysis of the tensor product of one irreducible representation, say  $T^{(\alpha)}$ , and one adjoint representation, say  $\bar{T}^{(\beta)}$ . In fact, using Equation (24.47), we have

$$m_1^{\alpha\beta} = \frac{1}{|G|} \sum_{i=1}^r c_i \bar{\chi}_i^{(1)} \chi_i^{(\alpha)} \bar{\chi}_i^{(\beta)} = \frac{1}{|G|} \sum_{i=1}^r c_i \chi_i^{(\alpha)} \bar{\chi}_i^{(\beta)} = \delta_{\alpha\beta},$$

where we used Equation (24.13) and the fact that all characters of the identity representation are unity. Thus to construct an invariant state, we need to combine a representation with its adjoint, in which case we obtain the identity representation only once. ■

---

**Paul Albert Gordan** (1837–1912), the son of David Gordan, a merchant, attended gymnasium and business school, then worked for several years in banks. His early interest in mathematics was encouraged by the private tutoring he received from a professor at the *Friedrich Wilhelm Gymnasium*. He attended Ernst Kummer's lectures in number theory at the University of Berlin in 1855, then studied at the universities of Breslau, Königsberg, and Berlin. At Königsberg he came under the influence of Karl Jacobi's school, and at Berlin his interest in algebraic equations was aroused. His dissertation (1862), which concerned geodesics on spheroids, received a prize offered by the philosophy faculty of the University of Breslau. The techniques that Gordan employed in it were those of Lagrange and Jacobi.



Gordan's interest in function theory led him to visit G. F. B. Riemann in Göttingen in 1862, but Riemann was ailing, and their association was brief. The following year, Gordan was invited to Giessen by Clebsch, thus beginning the fruitful collaboration most physicists recognize. Together they produced work on the theory of Abelian functions, based on Riemann's fundamental paper on that topic, and several of Clebsch's papers are considered important steps toward establishing for Riemann's theories a firm foundation in terms of pure algebraic geometry. Of course, the Clebsch–Gordan collaboration also produced the famous coefficients that bear their names, so indispensable to the theory of angular momentum coupling found in almost every area of modern physics. In 1874 Gordan became a professor at Erlangen, where he remained until his retirement in 1910. He married Sophie Deuer, the daughter of a Giessen professor of Roman law, in 1869. In 1868 Clebsch introduced Gordan to the theory of invariants, which originated in an observation of George Boole's in 1841 and was further developed by Arthur Cayley in 1846. Following the work of these two Englishmen, a German branch of the theory was developed by S. H. Aronhold and Clebsch, the latter elaborating the former's symbolic methods of characterizing algebraic forms and their invariants. Invariant theory was Gordan's main interest for the rest of his mathematical career; he became known as the greatest expert in the field, developing many techniques for representing and generating forms and their invariants.

Gordan made important contributions to algebra and solutions of algebraic equations, and gave simplified proofs of the transcendence of  $e$  and  $\pi$ . The overall style of Gordan's mathematical work was algorithmic. He shied away from presenting his ideas in informal literary forms. He derived his results computationally, working directly toward the desired goal without offering explanations of the concepts that motivated his work.

Gordan's only doctoral student, Emmy Noether, was one of the first women to receive a doctorate in Germany. She carried on his work in invariant theory for a while, but under the stimulus of Hilbert's school at Göttingen her interests shifted and she became one of the primary contributors to modern algebra.

---

So far, we have concentrated on the reduction of the operators and carrier spaces into irreducible components. Let us now direct our attention to the vectors themselves. Given two irreducible representations  $T^{(\alpha)}$  and  $T^{(\beta)}$  with carrier spaces spanned by vectors  $\{|\phi_i^{(\alpha)}\rangle\}_{i=1}^{n_\alpha}$  and  $\{|\psi_j^{(\beta)}\rangle\}_{j=1}^{n_\beta}$ , we form the direct product representation  $T^{(\alpha \times \beta)}$  with the carrier space spanned by vectors  $\{|\phi_i^{(\alpha)}\psi_j^{(\beta)}\rangle\}$ . We know that  $T^{(\alpha \times \beta)}$  is reducible, and Equation (24.47) tells us how many times each irreducible factor occurs in  $T^{(\alpha \times \beta)}$ . This means that the span of  $\{|\phi_i^{(\alpha)}\psi_j^{(\beta)}\rangle\}$  can be decomposed into invariant irreducible subspaces; i.e., there must exist a basis of the carrier of the product space the vectors of which belong to irreducible representations of  $G$ . More specifically, we should be able to form the linear combinations

Clebsch–Gordan coefficients

$$|\Psi_k^{(\sigma),q}\rangle = \sum_{ij} C(\alpha\beta; \sigma, q|ij; k) |\phi_i^{(\alpha)}\psi_j^{(\beta)}\rangle, \tag{24.48}$$

which transform according to the rows of the  $\sigma$ th irreducible representation. Here the subscript  $k$  refers to the row of the  $\sigma$ th representation, and  $q$  distinguishes among functions that have the same  $\sigma$  and  $k$ , corresponding to the case where  $m_\sigma^{\alpha\beta} \geq 2$ . For simply reducible groups, the label  $q$  is unnecessary. The coefficients  $C(\alpha\beta; \sigma, q|ij; k)$  are called the **Clebsch–Gordan coefficients** for  $G$ . These coefficients are normalized such that

$$\sum_{ij} C^*(\alpha\beta; \sigma, q|ij; k) C(\alpha\beta; \sigma', q'|ij; k') = \delta_{\sigma\sigma'} \delta_{qq'} \delta_{kk'},$$

$$\sum_{\sigma q k} C^*(\alpha\beta; \sigma, q|ij; k) C(\alpha\beta; \sigma, q|i'j'; k) = \delta_{i'i} \delta_{j'j}.$$

This will guarantee that  $|\Psi_k^{(\sigma),q}\rangle$  are orthonormal if the product vectors form an orthonormal set. Using these relations, we can write the inverse of Equation (24.48) as

$$|\phi_i^{(\alpha)}\psi_j^{(\beta)}\rangle = \sum_{\sigma q k} C^*(\alpha\beta; \sigma, q|ij; k) |\Psi_k^{(\sigma),q}\rangle. \tag{24.49}$$

### 24.7.2 Irreducible Tensor Operators

An operator  $\mathbf{A}$  acting in the carrier space of the representation of a group  $G$  is transformed into another operator,  $\mathbf{A} \mapsto \mathbf{T}_g \mathbf{A} \mathbf{T}_g^{-1}$ , by the action of the group. Just as in the case of vector spaces, one can thus construct a set of operators that transform among themselves by such action and lump these operators in irreducible sets.

irreducible set of operators

**24.7.4. Definition.** An operator  $\mathbf{A}_i^{(\alpha)}$  is said to transform according to the  $i$ th row of the  $\alpha$ th irreducible representation if there exist  $n_\alpha - 1$  other operators

$\{\mathbf{A}_j^{(\alpha)}\}$  and a basis  $\{|\psi_i^{(\alpha)}\rangle\}$  such that

$$\mathbf{T}_g \mathbf{A}_i^{(\alpha)} \mathbf{T}_g^{-1} = \sum_{j=1}^{n_\alpha} T_{ji}^{(\alpha)}(g) \mathbf{A}_j^{(\alpha)}, \tag{24.50}$$

where  $(T_{ji}^{(\alpha)}(g))$  is the matrix representation of  $g$ . The set of such operators is called an **irreducible set of operators** (or **irreducible tensorial set**).

scalar operator

In particular, if  $T_{ij}^{(\alpha)}(g) = \delta_{ij}$ , i.e., if the representation is the identity representation, then  $\mathbf{A} = \mathbf{T}_g \mathbf{A} \mathbf{T}_g^{-1}$ , and  $\mathbf{A}$  is called a **scalar operator**. The term “scalar” refers to the fact that  $\mathbf{A}$  has only one “component,” in contrast to the other operators of Equation (24.50), which may possess several components.

Consider the set of vectors (functions) defined by  $|\psi_{ij}\rangle \equiv \mathbf{A}_i^{(\alpha)} |\phi_j^{(\beta)}\rangle$ , where  $|\phi_j^{(\beta)}\rangle$  transform according to the  $\beta$ th irreducible representation. These vectors transform according to

$$\begin{aligned} \mathbf{T}_g |\psi_{ij}\rangle &= \mathbf{T}_g \mathbf{A}_i^{(\alpha)} \mathbf{T}_g^{-1} \mathbf{T}_g |\phi_j^{(\beta)}\rangle = \sum_{k=1}^{n_\alpha} T_{ki}^{(\alpha)}(g) \mathbf{A}_k^{(\alpha)} \sum_{l=1}^{n_\beta} T_{lj}^{(\beta)}(g) |\phi_l^{(\beta)}\rangle \\ &= \sum_{k,l} T_{ki}^{(\alpha)}(g) T_{lj}^{(\beta)}(g) \mathbf{A}_k^{(\alpha)} |\phi_l^{(\beta)}\rangle = \sum_{k,l} T_{kl,ij}^{(\alpha \times \beta)}(g) |\psi_{kl}\rangle, \end{aligned} \tag{24.51}$$

i.e., according to the representation  $T^{(\alpha \times \beta)}$ . This means that the vectors  $|\psi_{ij}\rangle$  have the same transformation properties as the tensor product vectors  $|\phi_i^{(\alpha)} \psi_j^{(\beta)}\rangle$ . Therefore, using Equation (24.49), we can write

$$\mathbf{A}_i^{(\alpha)} |\phi_j^{(\beta)}\rangle = \sum_{\sigma q k} C^*(\alpha\beta; \sigma, q | ij; k) |\Psi_k^{(\sigma),q}\rangle,$$

and more importantly,

$$\begin{aligned} \langle \phi_m^{(\gamma)} | \mathbf{A}_i^{(\alpha)} | \phi_j^{(\beta)} \rangle &= \sum_{\sigma q k} C^*(\alpha\beta; \sigma, q | ij; k) \underbrace{\langle \phi_m^{(\gamma)} | \Psi_k^{(\sigma),q} \rangle}_{\text{use Eq. (24.38) here}} \\ &= \sum_q C^*(\alpha\beta; \gamma, q | ij; m) \langle \phi_m^{(\gamma)} | \Psi_m^{(\gamma),q} \rangle. \end{aligned} \tag{24.52}$$

It follows that the matrix element of the operator  $\mathbf{A}_i^{(\alpha)}$  will vanish unless the irreducible representation  $T^{(\gamma)}$  occurs in the reduction of the tensor product  $T^{(\alpha)} \otimes T^{(\beta)}$ , and this can be decided from the character tables and the Clebsch–Gordan series, Equation (24.47).

There is another remarkable property of Equation (24.52) that has significant physical consequences. Notice how the dependence on  $i$  and  $j$  is contained entirely in the Clebsch–Gordan coefficients. From Equation (24.39) it follows that

Wigner–Eckart theorem and reduced matrix elements

$\langle \phi_m^{(\gamma)} | \Psi_m^{(\gamma),q} \rangle$  is independent of  $m$ . Therefore, this dependence must *also* be contained entirely in Clebsch–Gordan coefficients. One therefore writes (24.52) as

$$\langle \phi_m^{(\gamma)} | \mathbf{A}_i^{(\alpha)} | \phi_j^{(\beta)} \rangle \equiv \sum_q C^*(\alpha\beta; \gamma, q | i j; m) \underbrace{\langle \phi^{(\gamma)} | \mathbf{A}^{(\alpha)} | \phi^{(\beta)} \rangle_q}_{\text{reduced matrix element}}. \quad (24.53)$$

This equation is known as the **Wigner–Eckart theorem**, and the numbers multiplying the Clebsch–Gordan coefficients are known as the **reduced matrix elements**.

From the point of view of physics, Equation (24.53) can be very useful in calculating matrix elements (expectation values and transition between states), once we know the transformation properties of the physical operator. For example, for a **scalar operator**  $\mathbf{S}$ , which, by definition, transforms according to the identity representation, (24.53) becomes

$$\langle \phi_m^{(\gamma)} | \mathbf{A} | \phi_j^{(\beta)} \rangle = \langle \phi^{(\gamma)} | \mathbf{A}^{(\alpha)} | \phi^{(\beta)} \rangle \delta_{\gamma\beta} \delta_{mj};$$

i.e., scalar operators have no matrix elements between different irreducible representations of a group, and within an irreducible representation, they are multiples of the identity matrix. This result is also a consequence of Schur’s lemma.

## 24.8 Representations of the Symmetric Group

The symmetric (permutation) group is an important prototype of finite groups. In fact, **Cayley’s theorem** (see [Rotm 84, p. 46] for a proof) states that any finite group of order  $n$  is isomorphic to a subgroup of  $S_n$ . Moreover, the representation of  $S_n$  leads directly to the representation of many of the Lie groups encountered in physical applications. It is, therefore worthwhile to devote some time to the analysis of the representations of  $S_n$ .

### 24.8.1 Analytic Construction

The starting point of the construction of representations of the symmetric group is Equation (24.35), which is valid for any finite group. There is one simple character that every group has, namely, the character of the one-dimensional symmetric representation in which all elements of the group are mapped to  $1 \in \mathbb{R}$ . Setting  $\xi_j^{(k)} = 1$  in (24.35), and noting that  $\sum_j d_j = d_i$ , we obtain

$$\psi_i^H \equiv \frac{|G|d_i}{|H|c_i}, \quad (24.54)$$

where  $\{\psi_i^H\}$  are the components of a compound character of  $G$ .

Frobenius has shown that by a clever choice of  $H$ , one can completely solve the problem of the construction of the irreducible representations of  $S_n$ . He proceeded

as follows. Consider a partition  $(\lambda) = (\lambda_1, \dots, \lambda_n)$  of  $n$ . The symmetric groups  $\{S_{\lambda_i}\}$  are subgroups of  $S_n$  and have no elements in common—therefore they all commute with one another. The direct product of these subgroups is a subgroup of  $S_n$ , which we denote by  $S_{(\lambda)}$ :

$$S_{(\lambda)} = S_{\lambda_1} \times \dots \times S_{\lambda_n}.$$

If we denote the compound character of Equation (24.54) by  $|\psi^{(\lambda)}\rangle$  in this case, calculate  $|H|$ ,  $c_i$ , and  $d_i$ , and substitute the results in (24.54)—all of which can be done in closed form—we obtain an explicit formula for the components of  $|\psi^{(\lambda)}\rangle$ . This formula is messy, and we shall not derive it here. The interested reader may refer to [Hame 89, pp. 189–192] for details.

We are really interested in the simple characters of  $S_n$ , and Frobenius came up with a powerful method of calculating them. Since there is a one-to-one correspondence between the irreducible representations and conjugacy classes, and another one between conjugacy classes of  $S_n$  and partitions of  $n$ , we shall label the simple characters of  $S_n$  by partitions of  $n$ . Thus, instead of our common notation  $\chi_i^{(\alpha)}$ , we use  $\chi^{(\lambda)}$ , where  $(\lambda)$  denotes a partition of  $n$ , and  $(l)$  a cycle structure of  $S_n$ .

Suppose we want to find the irreducible characters corresponding to the cycle structure  $(l) = (1^\alpha, 2^\beta, 3^\gamma, \dots)$ . These form a *column* under the class  $(l)$  in a character table. To calculate the irreducible characters, form two polynomials in  $(x_1, x_2, \dots, x_n)$  as follows. The first one, which is completely symmetric in all variables, is

$$s^{(l)} \equiv \left(\sum_{i=1}^n x_i\right)^\alpha \left(\sum_{i=1}^n x_i^2\right)^\beta \left(\sum_{i=1}^n x_i^3\right)^\gamma \dots \tag{24.55}$$

The second one is completely antisymmetric, and can be written as

$$\begin{aligned} D(x_1, \dots, x_n) &\equiv \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{pmatrix} = \prod_{i < j} (x_i - x_j) \\ &= \sum_{\pi} \epsilon_{\pi} x_{\pi(1)}^{n-1} x_{\pi(2)}^{n-2} \dots x_{\pi(n-1)} x_{\pi(n)}^0. \end{aligned} \tag{24.56}$$

It can be shown that the simple characters of  $S_n$  are coefficients of certain terms of the product of these polynomials. To be exact, we have

$$\begin{aligned} s^{(l)} D(x_1, \dots, x_n) &= \sum_{(\lambda)} \chi^{(\lambda)} \sum_{\pi} \epsilon_{\pi} x_{\pi(1)}^{\lambda_1+n-1} x_{\pi(2)}^{\lambda_2+n-2} \dots x_{\pi(n-1)}^{\lambda_{n-1}+1} x_{\pi(n)}^{\lambda_n} \end{aligned} \tag{24.57}$$

The outer sum goes over all partitions of  $n$ , the inner sum over all permutations of  $S_n$ . The procedure for finding the simple characters of  $S_n$  should now be clear from (24.57):

**24.8.1. Box.** To find the simple character  $\chi_{(\lambda_1 \dots \lambda_n)}^{(\alpha_1 \dots \alpha_n)}$ , construct the corresponding symmetric and antisymmetric polynomials, multiply them together, collect all terms of the form

$$x_{\pi(1)}^{\lambda_1+n-1} x_{\pi(2)}^{\lambda_2+n-2} \dots x_{\pi(n-1)}^{\lambda_{n-1}+1} x_{\pi(n)}^{\lambda_n}$$

for all permutations  $\pi \in S_n$ ; the coefficient of such a term will be the desired character.

Since the coefficients of  $x_{\pi(1)}^{\lambda_1+n-1} x_{\pi(2)}^{\lambda_2+n-2} \dots x_{\pi(n-1)}^{\lambda_{n-1}+1} x_{\pi(n)}^{\lambda_n}$  for various  $\pi$ 's differ by a sign, in practice, the coefficient of just  $x_1^{\lambda_1+n-1} \dots x_{n-1}^{\lambda_{n-1}+1} x_n^{\lambda_n}$  is precisely the simple character we are looking for.

**24.8.2. Example.** The best way to understand the procedure described above is to go through an example in detail. We calculate the characters of  $S_3$  using the above method. Label the rows of the character table with the partitions of 3. These are (3), (2, 1), and (1, 1, 1). Similarly, label the columns with the conjugacy classes, or cycle structures: (1<sup>3</sup>), (1, 2), and (3). The first cycle structure has  $\alpha = 3, \beta = 0 = \gamma$ . Therefore,

$$s_{(1^3)} = (x_1 + x_2 + x_3)^3 = x_1^3 + x_2^3 + x_3^3 + 3(x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 + x_2^2 x_3 + x_1 x_3^2 + x_2 x_3^2) + 6x_1 x_2 x_3 \tag{24.58}$$

and

$$D(x_1, x_2, x_3) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3),$$

so that

$$s_{(1^3)} D(x_1, x_2, x_3) = s_{(1^3)} (x_1^2 x_2 - x_1^2 x_3 - x_2^2 x_1 + x_2^2 x_3 - x_3^2 x_2 + x_3^2 x_1). \tag{24.59}$$

Now we note that for  $(\lambda) = (3), \lambda_1 = 3, \lambda_2 = 0, \text{ and } \lambda_3 = 0$ . Therefore, the coefficient of  $x_1^{\lambda_1+n-1} x_2^{\lambda_2+n-2} \dots x_n^{\lambda_n} = x_1^5 x_2$  gives  $\chi_{(1^3)}^{(3)}$ . Similarly, for  $(\lambda) = (2, 1, 0), \lambda_1 = 2, \lambda_2 = 1, \text{ and } \lambda_3 = 0$ , and the coefficient of  $x_1^{\lambda_1+n-1} x_2^{\lambda_2+n-2} \dots x_n^{\lambda_n} = x_1^4 x_2^2$  gives  $\chi_{(1^3)}^{(2,1)}$ . Finally, for  $(\lambda) = (1, 1, 1), \lambda_1 = \lambda_2 = \lambda_3 = 1$ , and the coefficient of  $x_1^{\lambda_1+n-1} x_2^{\lambda_2+n-2} \dots x_n^{\lambda_n} = x_1^3 x_2^2 x_3$  gives  $\chi_{(1^3)}^{(1,1,1)}$ . These coefficients can be read off by scanning through Equation (24.58) while multiplying its terms by those of Equation (24.59) and keeping track of the coefficients of the products of the relevant powers of  $x_1, x_2,$  and  $x_3$ . The reader may verify that there is only one term of the form  $x_1^3 x_2^2 x_3$ , whose coefficient is

	(1 <sup>3</sup> )	(1, 2)	(3)
$T^{(3)}$	1	1	1
$T^{(2,1)}$	2	0	-1
$T^{(1,1,1)}$	1	-1	1

**Table 24.5** The character table for  $S_3$ . Each column corresponds to a conjugacy class, each row to a partition of 3. The last two rows have been switched compared to Table 24.4.

1, giving  $\chi_{(1^3)}^{(3)} = 1$ ; there are two terms of the form  $x_1^4 x_2^2$ , whose coefficients are  $-1$  and  $3$ , giving  $\chi_{(1^3)}^{(2,1)} = 2$ ; and there are four terms of the form  $x_1^3 x_2^2 x_3$ , whose coefficients are  $+1, -3, -3$ , and  $+6$ , giving  $\chi_{(1^3)}^{(1,1,1)} = 1$ . Therefore, the first column of the character table of  $S_3$  is  $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ . To obtain the second column, we consider the second conjugacy class,  $(1, 2)$ , with  $\alpha = 1 = \beta$  and  $\gamma = 0$ . The corresponding symmetric polynomial is

$$\begin{aligned} s_{(1,2)} &= (x_1 + x_2 + x_3)(x_1^2 + x_2^2 + x_3^2) \\ &= x_1^3 + x_2^3 + x_3^3 + x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 + x_2 x_3^2 + x_1 x_3^2 + x_2 x_3^2. \end{aligned} \tag{24.60}$$

$D(x_1, x_2, x_3)$  is the same as before. Multiplying and keeping track of the coefficients of  $x_1^5 x_2$ ,  $x_1^4 x_2^2$ , and  $x_1^3 x_2^2 x_3$ , we obtain  $\chi_{(1,2)}^{(3)} = 1$ ,  $\chi_{(1,2)}^{(2,1)} = 0$ , and  $\chi_{(1,2)}^{(1,1,1)} = -1$ . It follows that the second column of the character table of  $S_3$  is  $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ .

The last column is obtained similarly. We note that  $\alpha = 0 = \beta$ , and  $\gamma = 1$ . Therefore, the symmetric polynomial is

$$s_{(3)} = x_1^3 + x_2^3 + x_3^3,$$

and the antisymmetric polynomial is the same as before. Multiplying these two polynomials and extracting the coefficients as before, we get  $\chi_{(3)}^{(3)} = 1$ ,  $\chi_{(3)}^{(2,1)} = -1$ , and  $\chi_{(3)}^{(1,1,1)} = 1$ .

It follows that the third column of the character table of  $S_3$  is  $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ .

Collecting all the data obtained above, we can reconstruct the character table of  $S_3$ . This is shown in Table 24.5. The irreducible representations are labeled by the three possible partitions of 3, and the conjugacy classes by the three cycle structures. ■

### 24.8.2 Graphical Construction

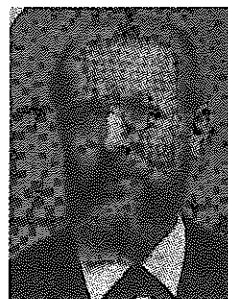
The analytic construction of the previous subsection can be handled using graphical techniques that are considerably simpler. To begin with, let us find the character of the identity element of  $S_n$ . The cycle structure is  $(1^n)$ , i.e., all cycles consist of a single element. Thus,  $\alpha = n$ , and  $\beta, \gamma$ , etc. are all zero. It follows that the LHS of Equation (24.57) is  $(\sum x_i)^n D(x_j)$ . We calculate this product one power of  $\sum x_i$



at a time. For the same reason as in the example above,  $\chi_{(1^n)}^{(\lambda)}$  will be the coefficient of

$$x_1^{\lambda_1+n-1} x_2^{\lambda_2+n-2} \dots x_{n-1}^{\lambda_{n-1}+1} x_n^{\lambda_n}.$$

**Ferdinand Georg Frobenius** (1849–1917), the son of a parson, was born in Berlin and began his mathematical studies at Göttingen in 1867. He received his doctorate in Berlin three years later. Four years later, on the basis of his mathematical research, he was appointed assistant professor at the University of Berlin. He achieved the rank of full professor at the *Eidgenössische Polytechnikum Zürich* before returning to Berlin as a professor of mathematics in 1892. During the early years of Frobenius's career, modern group theory was in its infancy. He combined its three main branches of study—the theory of solutions to algebraic equations (permutation groups and the work of Galois), geometry (transformation and Lie groups), and number theory—to produce the concept of the abstract group. He collaborated with Issai Schur in representation and character theory of groups.



His paper *Über die Gruppencharacterere* is of fundamental importance. It was presented to the Berlin Academy on 16 July 1896 and it contains work that Frobenius had done in the preceding few months. In a series of letters to Dedekind, the first on 12 April 1896, his ideas on group characters quickly develop, and Frobenius is able to construct a complete set of representations by complex numbers. In a letter to Dedekind on 26 April 1896 Frobenius finds the irreducible characters for the alternating group, and the symmetric groups.

In 1897 Frobenius reformulated the work of Molien—the Latvian student of Klein, who, in his thesis, classified the semisimple algebras using the method of group rings—in terms of matrices and then showed that his characters are the traces of the irreducible representations. Frobenius's character theory found important applications in quantum mechanics and was used with great effect by Burnside, who wrote it up in the 1911 edition of his *Theory of Groups of Finite Order*.

Frobenius is also remembered as the originator of a series method for solving ordinary differential equations. Despite the clearly greater importance of his work in group theory, this *method of Frobenius* serves admirably to perpetuate his name.

If we multiply  $D(x_j)$  by  $\sum x_i$  one  $x$  at a time, we increase the power of one of the  $x_i$ 's by one. If at any stage, two of the exponents become equal, the term must vanish, due to the antisymmetry of  $(\sum x_i)D(x_j)$ . Therefore, as we raise the degree of the polynomial by one at each stage, the power of  $x_1$  must be raised at least as fast as  $x_2$ , and the power of  $x_2$  must be raised at least as fast as  $x_3$ , etc. Our goal is to raise the power of  $x_1$  by  $\lambda_1$ , that of  $x_2$  by  $\lambda_2$ , and, in general, the power of  $x_i$  by  $\lambda_i$ , making sure that at each stage, the number of multiplications by  $x_1$  is greater than or equal to the number of multiplications by  $x_2$ , etc. The total number

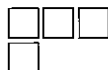
of ways by which we can reach this goal will be  $\chi_{(1^n)}^{(\lambda)}$ , which is also the dimension of the irreducible representation  $(\lambda)$  by Equation (24.9).

To see the argument more clearly, suppose that we are interested in the dimension of the irreducible representation of  $S_4$  corresponding to  $(3, 1)$ . Then we must raise the power of  $x_1$  by 3 and the power of  $x_2$  by 1;  $x_3$  and  $x_4$  will remain intact, and therefore will not enter in the following discussion. It follows that  $D(x_j)$  is to be multiplied by  $x_1^3 x_2$ , one  $x$ -factor at a time, the number of  $x_1$ -factors always exceeding the number of  $x_2$ -factors. The possible ways of doing this are

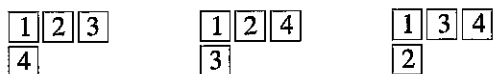
$$x_1^3 x_2, \quad x_1^2 x_2 x_1, \quad x_1 x_2 x_1^2. \tag{24.61}$$

Note that as we count the factors from left to right, the number of  $x_1$ 's is always greater than or equal to the number of  $x_2$ 's. Thus  $x_2 x_1^3$  is absent because  $x_2$  occurs without  $x_1$  occurring first. It follows that the dimension of the irreducible representation  $(3, 1)$  is 3.

A graphical way to arrive at the same result is to draw  $\lambda_1 = 3$  boxes on top and  $\lambda_2 = 1$  box below it:



The next step is to fill in the boxes with numbers corresponding to the position of  $x_1$  (filling up the first row) and  $x_2$  factors (filling up the second row) in Equation (24.61). Since in the first term of (24.61), the  $x_1$ 's occupy the first, second, and third positions, we enter 1, 2, and 3 in the first row, and 4 in the second row corresponding to the last position occupied by  $x_2$ . Similarly, in the second term of (24.61), the  $x_1$ 's occupy the first, second, and fourth positions; therefore, we enter 1, 2, and 4 in the first row, and 3 in the second row corresponding to the position occupied by  $x_2$ . Finally, in the last term of (24.61), the  $x_1$ 's occupy the first, third, and fourth positions; therefore, we enter 1, 3, and 4 in the first row, and 2 in the second row corresponding to the position occupied by  $x_2$ . The result is the graph shown below:



Young frame defined

**24.8.3. Definition.** Let  $(\lambda) = (\lambda_1, \lambda_2, \dots, \lambda_n)$  be a partition of  $n$ . The **Young frame** (or the **Young pattern**) associated with  $(\lambda)$  is a collection of rows of boxes (squares) aligned at the left such that the first row has  $\lambda_1$  boxes, the second row  $\lambda_2$  boxes, etc. Since  $\lambda_i \geq \lambda_{i+1}$ , the length of the rows decreases as one goes to the bottom of the frame.

The Young frame associated with  $(\lambda)$  represents  $x_1^{\lambda_1} \dots x_n^{\lambda_n}$ , which multiplies the antisymmetric polynomial  $D(x_j)$ . To find the dimension of the irreducible representation  $T^{(\lambda)}$ , we have to count the number of ways in which the  $x$ -factors can be permuted among themselves such that as we scan the product, the number of  $x_i$ 's is never less than number of  $x_j$ 's if  $j > i$ . This leads to

standard Young tableaux defined

**24.8.4. Definition.** A *standard Young tableau* (or *diagram*, or *graph*) is a Young frame filled with numbers 1 through  $n$  such that

1. the numbers are placed consecutively left to right on the rows starting with 1 in the far-left box of the first row;
2. no box of any row is to be filled unless all boxes to its left are already filled;
3. at each stage, the number of boxes filled in any row is never less than the number of rows below it.

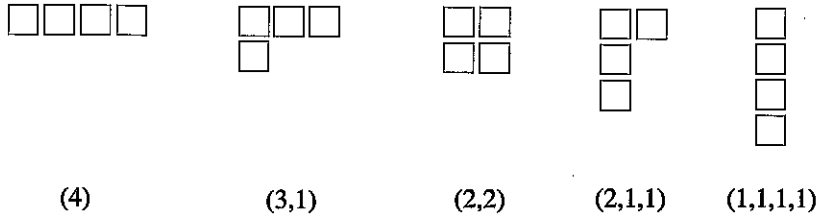
regular graphs

Tableaux satisfying the last condition are called **regular graphs**.

It follows that in a Young tableau, the number 1 is always in the upper left-hand box, and that going down in a column, the numbers must increase.

**24.8.5. Theorem.** Let  $(\lambda)$  be a partition of  $n$ . Then the dimension of the irreducible representation  $T^{(\lambda)}$  is equal to the number of standard Young tableaux associated with  $(\lambda)$ .

**24.8.6. Example.** We wish to calculate the dimension of each irreducible representation of  $S_4$ . The partitions are  $(4)$ ,  $(3, 1)$ ,  $(2, 2)$ ,  $(2, 1, 1)$ , and  $(1, 1, 1, 1)$  whose associated Young frames are shown below:



The number of standard Young tableaux associated with  $(\lambda) = (4)$  is 1, because there is only one way to place the numbers 1 through 4 in the four boxes. Thus, the dimension of  $T^{(4)}$  is 1. For  $(\lambda) = (3, 1)$ , we can place 2 either to the right of 1 or below it. The first choice gives rise to two possibilities for the placement of 3: Either to the right of 2 or below 1. The second choice gives rise to only one possibility for 3, namely to the right of 1. With 1, 2, and 3 in place, the position of 4 is predetermined. Thus, we have 3 possibilities for  $(\lambda) = (3, 1)$ , and the dimension of  $T^{(3,1)}$  is 3. For  $(\lambda) = (2, 2)$ , we can place 2 either to the right of 1 or below it. Both choices give rise to only one possibility for 3: In the first case, 3 can only go under 1; in the second case to its right. With 1, 2, and 3 in place, the position of 4 is again predetermined. Thus, we have 2 possibilities for  $(\lambda) = (2, 2)$ , and the dimension of  $T^{(2,2)}$  is 2. The reader may check that the dimension of  $T^{(2,1,1)}$  is 3, and that of  $T^{(1,1,1,1)}$  is 1. Figure 24.1 summarizes these findings. We note that the dimensions satisfy  $1^2 + 3^2 + 2^2 + 3^2 + 1^2 = 24$ , the second equation of (24.18). ■

(4)	$\begin{array}{ c c c c } \hline 1 & 2 & 3 & 4 \\ \hline \end{array}$	$n_{(4)} = 1$		
(3,1)	$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 4 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline 1 & 3 & 4 \\ \hline 2 \\ \hline \end{array}$	$\begin{array}{ c c c } \hline 1 & 2 & 4 \\ \hline 3 \\ \hline \end{array}$	$n_{(3,1)} = 3$
(2,2)	$\begin{array}{ c c } \hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array}$	$\begin{array}{ c c } \hline 1 & 3 \\ \hline 2 & 4 \\ \hline \end{array}$	$n_{(2,2)} = 2$	
(2,1,1)	$\begin{array}{ c c } \hline 1 & 2 \\ \hline 3 \\ \hline 4 \\ \hline \end{array}$	$\begin{array}{ c c } \hline 1 & 4 \\ \hline 2 \\ \hline 3 \\ \hline \end{array}$	$\begin{array}{ c c } \hline 1 & 3 \\ \hline 2 \\ \hline 4 \\ \hline \end{array}$	$n_{(2,1,1)} = 3$
(1,1,1,1)	$\begin{array}{ c } \hline 1 \\ \hline 2 \\ \hline 3 \\ \hline 4 \\ \hline \end{array}$	$n_{(1,1,1,1)} = 1$		

**Figure 24.1** The standard Young tableaux, and the dimensions of irreducible representations of  $S_4$ .

### 24.8.3 Graphical Construction of Characters

The product of the symmetric polynomial  $s_{(l)}$  and the antisymmetric polynomial  $D(x_j)$  contains all the information regarding the representations of  $S_n$ . We can extract the simple characters by looking at the coefficients of appropriate products of the  $x$ -factors. This can also be done graphically. Without going into the combinatorics of the derivation of the results, we state the rules for calculating the simple characters, and examine one particular case in detail to elucidate the procedure, whose statement can be very confusing.

As before, we label the irreducible representations with the partitions of  $n$ . However, we separate out the common factors in a cyclic structure, labeling the cycles by  $l_1, l_2$ , etc. For example,  $(2, 1^2)$  has  $l_1 = 2, l_2 = 1$ , and  $l_3 = 1$ . So, we write  $(l)$  as  $(l_1, l_2, \dots, l_m)$ .

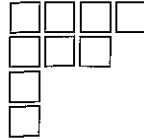
regular application **24.8.7. Definition.** A *regular application* of  $r$  identical symbols to a Young frame is the placement of those symbols in the boxes of the frame as follows. Add the symbols to any given row, starting with the first (farthest to the left) unoccupied cell, until the symbols are all used<sup>6</sup> or the number exceeds that of the preceding

<sup>6</sup>It is understood that if you cannot place all symbols on the first row, then you should start at the second row.

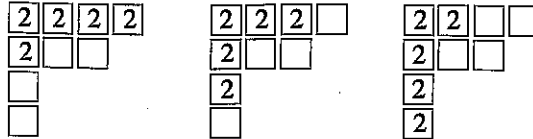
positive and negative applications

line by one. In the latter case, go to the preceding line and repeat the procedure, making sure that the final result of adding all  $r$  symbols will be a regular graph. If in this process the  $r$  symbols are divided among an even number of rows, we speak of **negative application**. If the number of rows is odd, we have a **positive application**.

As an illustration, consider the regular application of five 2's to the blank Young frame shown below.



We cannot start on the first row because it does not have enough boxes for the five 2's. We can start on the second row and put one 2 in the first box. This brings the number of 2's in the second row to one more than in the first row; therefore, we should now go to the first row and put the rest of the symbols there. We could start at the third row, put one 2 in the first box, put a second 2 in the first box of the second row, and the rest in the first row. Altogether we will have 3 regular applications of the five 2's. These are shown in the diagram below.



Of these the first and the last tableaux are negative applications, and the middle one is positive.

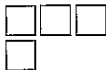
**24.8.8. Theorem.** *The character of the irreducible representation  $T^{(\lambda)}$  of the class  $(l) = (l_1, l_2, \dots, l_m)$  is obtained by successive regular applications of  $l_1$  identical symbols (usually taken to be 1's), then  $l_2$  identical symbols of a different kind (usually taken to be 2's), etc. The character  $\chi_{(l)}^{(\lambda)}$  is then equal to the number of ways of building positive applications minus the number of ways of building negative applications.*

The order in which the  $l_i$ 's are applied is irrelevant. However, it is usually convenient to start with the largest cycle.

The best way to understand the procedure is to construct a character table. Let us do this for  $S_4$ . As usual, the rows are labeled by the various partitions  $(\lambda)$  of 4. We choose the order  $(4), (3, 1), (2, 2) = (2^2), (2, 1, 1) = (2, 1^2), (1, 1, 1, 1) = (1^4)$ . The columns are labeled by classes  $(l)$  in the following order:  $(1^4), (2, 1^2), (2^2), (3, 1), (4)$ , where, for example,  $(2, 1^2)$  means that  $l_1 = 2, l_2 = 1,$  and  $l_3 = 1$ . Example 24.8.6 gives us the first column of the character table. Similarly, the first row has 1 in all places, because it is the trivial representation. Our task is therefore

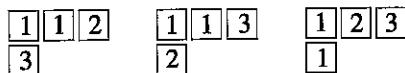
Detailed analysis of the construction of the character table for  $S_4$

to fill in the rest of the table one row at a time. The second row, with  $(\lambda) = (3, 1)$ , has a Young frame that looks like



and for each class (column) labeled  $(l_1, \dots, l_m)$ , we need to fill this in with  $l_1$  identical symbols (1's),  $l_2$  identical symbols of a different kind (2's), etc.

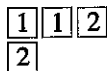
The second column has  $l_1 = 2, l_2 = 1 = l_3$ . So we have two 1's, one 2, and one 3. If we start with the first row, the two 1's can be placed in its first two boxes. If we start with the second row, the two 1's must be placed vertically on top of each other. In the first case, we have two choices for the 2: Either on the first row next to the two 1's, or on the second line. In the second case, we have only one choice for the 2: in the first row next to 1. With 1's and 2 in place, the position of 3 is determined. The three possibilities are shown below:



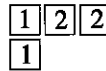
The first two are positive applications, the third is negative because the 1's occupy an even number of rows. We therefore have

$$\chi_{(2,1^2)}^{(3,1)} = +1 + 1 - 1 = +1.$$

The third column has  $l_1 = 2 = l_2$ . So we have two 1's and two 2's. We place the 1's as before. When the two 1's are placed vertically, we can put the 2's on the first row and we are done. When the 1's are initially placed in the first row, we have no way of placing the 2's by regular application. We cannot start on the first row because there is only one spot available (remember, we cannot go down once we start at a row). We cannot start on the second row because once we place the first 2, we are blocked, and the number of symbols in the second row does not exceed that of the first row by one. So, there is only one possibility:



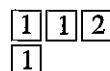
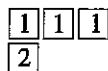
Not allowed



Allowed

The only allowed diagram is obtained by a negative application of 1's. Therefore,  $\chi_{(2^2)}^{(3,1)} = -1$ .

The fourth column has  $l_1 = 3$  and  $l_2 = 1$ . So we have three 1's and one 2. There are two ways to place the 1's: all on the first row, or starting on the second row and working our way up until all boxes are filled except the last box of the first row. The placement of 2 will be then predetermined. The result is the two diagrams shown below:



The first diagram is obtained by a positive application of 1's, the second by a negative application. Therefore,

$$\chi_{(3,1)}^{(3,1)} = +1 - 1 = 0.$$

Finally, for the last column,  $l_1 = 4$ . There is only one way to put all the 1's in the frame, and that is a negative application. Thus,  $\chi_{(4)}^{(3,1)} = -1$ .

Rather than going through the rest of the table in the same gory detail, we shall point out some of the trickier calculations, and leave the rest of the table for the reader to fill in. One confusion may arise in the calculation of  $\chi_{(2^2)}^{(2^2)}$ . The frame looks like this,



and we need to fill this with two 1's and two 2's. The 1's can go into the first row or the first column. The 2's then can be placed in the second row or the second column. The result is



The first diagram has no negative application. The second has *two* negative applications, one for the 1's, and one for the 2's. Therefore, the overall sign for the second diagram is positive. It follows that  $\chi_{(2^2)}^{(2^2)} = +1 + 1 = +2$ .

The calculation of  $\chi_{(4)}^{(2^2)}$  may also be confusing. We need to place four 1's in the frame. If we start on the first row, we are stuck, because there is room for only two 1's. If we start in the second row, then we can only go up: Putting the first 1 in the second row causes that row to have one extra 1 in comparison with the preceding row. However, once we go up, we have room for only two 1's (we cannot go back down). So, there is no way we can place the four 1's in the  $(2^2)$  frame, and  $\chi_{(4)}^{(2^2)} = 0$ .

The character table for  $S_4$  is shown in Table 24.6 (see Problem 24.15 as well). The reader is urged to verify all entries not calculated above. The character table for  $S_5$  can also be calculated with only minor tedium. We quote the result here in Table 24.7 and let the reader check the entries of the table.

### 24.8.4 Young Operators

The group algebra techniques of Section 24.4—which we used in our discussion of representation theory in a very limited way—provide a powerful and elegant tool

	$(1^4)$	$(2, 1^2)$	$(2^2)$	$(3, 1)$	$(4)$
$T^{(4)}$	1	1	1	1	1
$T^{(3,1)}$	3	1	-1	0	-1
$T^{(2^2)}$	2	0	2	-1	0
$T^{(2,1^2)}$	3	-1	-1	0	1
$T^{(1^4)}$	1	-1	1	1	-1

**Table 24.6** Character table for  $S_4$ . The rows and columns are labeled by the partitions of 4 and cycle structures, respectively.

for unraveling the representations of finite groups. These techniques have been particularly useful in the analysis of the representations of the symmetric group. Our emphasis on the symmetric group is not merely due to the importance of  $S_n$  as a paradigm of all finite groups. It has also to do with the unexpected usefulness of the representations of  $S_n$  in studying the representations of  $GL(V)$ , the paradigm of all (classical) *continuous* groups. We shall come back to this observation later when we discuss representations of Lie groups in Chapter 27.

To begin with, consider the element of the  $S_n$  group algebra as defined in Equation (24.22). Since multiplying  $P$  (on the left) by a group element does not change  $P$ , the ideal generated by  $P$  is not only one-dimensional, but all elements of  $S_n$  are represented by the number 1. Therefore, the ideal  $AP$  corresponds to the (irreducible) identity representation.

For  $S_n$ , there is another group algebra element that has similar properties. This is

$$Q = \sum_{i=1}^{n!} \epsilon_{\pi_i} \pi_i, \quad \pi_i \in S_n. \quad (24.62)$$

The reader may check that

$$\pi_j Q = \epsilon_{\pi_j} Q \quad \text{and} \quad Q^2 = n! Q.$$

As in the case of  $P$ ,  $Q$  generates a one-dimensional ideal, but a left multiplication may introduce a minus sign (when the permutation is odd). Thus, the ideal generated by  $Q$  must correspond to the antisymmetric (or alternating) representation.

All the irreducible representations, including the special one-dimensional cases above, can be obtained using this group-algebraic method. We shall not give the proofs here, and we refer the reader to the classic book [Boer 63, pp. 102–125]. The starting point is the Young frame corresponding to the partition  $(\lambda) = (\lambda_1, \dots, \lambda_m)$ . One puts the numbers 1 through  $n$  in the frame in *any* order, consistent with tableau construction, so that the end product is a Young tableau. Let  $p$  be any permutation of a Young tableau that does not change the row



	(1 <sup>5</sup> )	(2, 1 <sup>3</sup> )	(2 <sup>2</sup> , 1)	(3, 2)	(3, 1 <sup>2</sup> )	(4, 1)	(5)
$T^{(5)}$	1	1	1	1	1	1	1
$T^{(4,1)}$	4	2	0	-1	1	0	-1
$T^{(3,2)}$	5	1	1	1	-1	-1	0
$T^{(3,1^2)}$	6	0	-2	0	0	0	1
$T^{(2^2,1)}$	5	-1	1	-1	-1	1	0
$T^{(2,1^3)}$	4	-2	0	1	1	0	-1
$T^{(1^5)}$	1	-1	1	-1	1	-1	1

**Table 24.7** Character table for  $S_5$ . The rows and columns are labeled by the partitions of 5 and cycle structures, respectively.

horizontal and  
vertical permutations

of the numbers it acts on. Such a  $p$  is called a **horizontal permutation**. Similarly, let  $q$  be a **vertical permutation** of the Young tableau, i.e., a permutation that does not change the column of the numbers it acts on.

Young operators  
defined

**24.8.9. Definition.** Consider the  $k$ th Young tableau corresponding to the partition  $(\lambda)$ . Let the **Young symmetrizer**  $P_k^{(\lambda)}$  and **antisymmetrizer**  $Q_k^{(\lambda)}$  be the elements of the group algebra of  $S_n$  defined as

$$P_k^{(\lambda)} = \sum_p p, \quad Q_k^{(\lambda)} = \sum_q \epsilon_q q.$$

Then, the **Young operator**  $Y_k^{(\lambda)}$  of this tableau, another element of the group algebra, is given by  $Y_k^{(\lambda)} = Q_k^{(\lambda)} P_k^{(\lambda)}$ .

It can be shown that the following holds.

**24.8.10. Theorem.** The Young operator  $Y_k^{(\lambda)}$  is essentially idempotent, and generates a minimal left ideal, hence an irreducible representation for  $S_n$ . Representations thus obtained from different frames are inequivalent. Different tableaux with the same frame give equivalent irreducible representations.

In practice, one usually chooses the standard Young tableaux and applies the foregoing procedure to them to obtain the entire collection of inequivalent irreducible representations of  $S_n$ . We have already seen how to calculate characters of  $S_n$  employing both analytical and graphical methods. Theorem 24.8.10 gives yet another approach to analyzing representations of  $S_n$ . For low values of  $n$  this technique may actually be used to determine the characters, but as  $n$  grows, it becomes unwieldy, and the graphical method becomes more manageable.

**24.8.11. Example.** Let us apply this method to  $S_3$ . The partitions are (3), (2, 1), and (1<sup>3</sup>). There is only one standard Young tableau associated with (3) and (1<sup>3</sup>). Thus,

$$Y^{(3)} = P^{(3)} = \frac{1}{3!} \sum_{j=1}^6 \pi_j = \frac{1}{6} (e + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6),$$

$$Y^{(1^3)} = Q^{(1^3)} = \frac{1}{3!} \sum_{j=1}^6 \epsilon_{\pi_j} \pi_j = \frac{1}{6} (e - \pi_2 - \pi_3 - \pi_4 + \pi_5 + \pi_6),$$

where we have divided these Young operators by 6 to make them idempotent; we have also used the notation of Example 23.4.1. One can show directly that  $Y^{(3)}Y^{(1^3)} = 0$ . In fact, one can prove this for general  $S_n$  (see Problem 24.30).

For the partition (2, 1), there are two Young tableaux. The first one has the numbers 1 and 2 in the first row and 3 in the second. In the second tableau the numbers 2 and 3 are switched. Therefore, using the multiplication table for  $S_3$  as given in Example 23.4.1, we have

$$Y_1^{(2,1)} = Q_1^{(2,1)} P_1^{(2,1)} = (e - \pi_3)(e + \pi_2) = e + \pi_2 - \pi_3 - \pi_6,$$

$$Y_2^{(2,1)} = Q_2^{(2,1)} P_2^{(2,1)} = (e - \pi_2)(e + \pi_3) = e - \pi_2 + \pi_3 - \pi_5.$$

The reader may verify that the product of any two Young operators corresponding to different Young tableaux is zero and that

$$Y_1^{(2,1)} Y_1^{(2,1)} = 3Y_1^{(2,1)}, \quad Y_2^{(2,1)} Y_2^{(2,1)} = 3Y_2^{(2,1)}.$$

Let us calculate the left ideal generated by these four Young operators. We already know from our discussion at the beginning of this subsection that  $\mathcal{L}^{(3)}$  and  $\mathcal{L}^{(1^3)}$ , the ideals generated by  $Y^{(3)}$  and  $Y^{(1^3)}$ , are one-dimensional. Let us find  $\mathcal{L}_1^{(2,1)}$ , the ideal generated by  $Y_1^{(2,1)}$ . This is the span of all vectors obtained by multiplying  $Y_1^{(2,1)}$  on the left by elements of the group algebra. It is sufficient to multiply  $Y_1^{(2,1)}$  by the basis of the algebra, namely the group elements:

$$eY_1^{(2,1)} = Y_1^{(2,1)},$$

$$\pi_2 Y_1^{(2,1)} = \pi_2 + e - \pi_5 - \pi_4 \equiv X_1^{(2,1)},$$

$$\pi_3 Y_1^{(2,1)} = \pi_3 + \pi_6 - e - \pi_2 = -Y_1^{(2,1)},$$

$$\pi_4 Y_1^{(2,1)} = \pi_4 + \pi_5 - \pi_6 - \pi_3 = -X_1^{(2,1)} + Y_1^{(2,1)},$$

$$\pi_5 Y_1^{(2,1)} = \pi_5 + \pi_4 - \pi_2 - e = -X_1^{(2,1)},$$

$$\pi_6 Y_1^{(2,1)} = \pi_6 + \pi_3 - \pi_4 - \pi_5 = X_1^{(2,1)} - Y_1^{(2,1)}.$$

It follows from the above calculation that  $\mathcal{L}_1^{(2,1)}$ , as a vector space, is spanned by  $\{Y_1^{(2,1)}, X_1^{(2,1)}\}$ , and since these two vectors are linearly independent,  $\mathcal{L}_1^{(2,1)}$  is a two-dimensional minimal ideal corresponding to a two-dimensional irreducible representation of  $S_3$ . One can use this basis to find representation matrices and the simple characters of  $S_3$ .

The other two-dimensional irreducible representation of  $S_3$ , *equivalent* to the one above, is obtained by constructing the ideal  $\mathcal{L}_2^{(2,1)}$  generated by  $Y_2^{(2,1)}$ . This construction is left for the reader, who is also asked to verify its dimensionality.

The resolution of the identity is easily verified:

$$e = \underbrace{\frac{1}{6}Y^{(3)}}_{\equiv e_1} + \underbrace{\frac{1}{6}Y^{(1^3)}}_{\equiv e_2} + \underbrace{\frac{1}{3}Y_1^{(2,1)}}_{\equiv e_3} + \underbrace{\frac{1}{3}Y_2^{(2,1)}}_{\equiv e_4}.$$

The  $e_i$ 's are idempotents that satisfy  $e_i e_j = 0$  for  $i \neq j$ . ■

### 24.8.5 Products of Representations of $S_n$

In the quantum theory of systems of many *identical* particles, the wave function must have a particular symmetry under exchange of the particles: If the particles are all fermions (bosons), the overall wave function must be completely antisymmetric (symmetric). Since the space of functions of several variables can provide a carrier space for the representation of any group, we can, in the case of  $S_n$ , think of the antisymmetric (symmetric) functions as basis functions for the one-dimensional irreducible identity (alternating) representation. To obtain these basis functions, we apply the Young operator  $Y^{(1^n)}$  (or  $Y^{(n)}$ ) to the arguments of any given function of  $n$  variables to obtain the completely antisymmetric (or symmetric) wave function.<sup>7</sup>

Under certain conditions, we may require mixed symmetries. For instance, in the presence of spin, the product of the total spin wave function and the total space wave function must be completely antisymmetric for Fermions. Thus, the space part (or the spin part) of the wave functions will, in general, have mixed symmetry. Such a mixed symmetry corresponds to some other Young operator, and the wave function is obtained by applying that Young operator to the arguments of the wave function.

Now suppose that we have two separate systems consisting of  $n_1$  and  $n_2$  particles, respectively, which are all assumed to be identical. As long as the two systems are not interacting, each will consist of states that are classified according to the irreducible representations of its symmetric group. When the two systems interact, we should classify the states of the total system according to the irreducible representations of all  $n_1 + n_2$  particles. We have already encountered the mathematical procedure for such classification: It is the Clebsch–Gordan decomposition of the direct product of the states of the two systems. Since the initial states correspond to Young tableaux, and since we are interested in the inequivalent irreducible representations, we need to examine the decomposition of the direct product of Young frames into a sum of Young frames. We first state (without proof) the procedure for such a decomposition, and then give an example to illustrate it.

**24.8.12. Theorem.** *To find the components of Young frames in the product of two Young frames, draw one of the frames. In the other frame, assign the same symbol,*

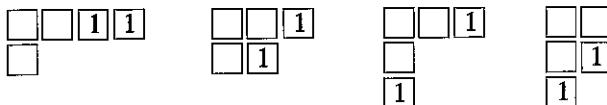
<sup>7</sup>We must make the additional assumption that the permuted functions are all independent.

say 1, to all boxes in the first row, the same symbol 2 to all boxes in the second row, etc. Now attach the first row to the first frame, and enlarge in all possible ways subject to the restriction that no two 1's appear in the same column, and that the resultant graph be regular. Repeat with the 2's, etc., making sure in each step that as we read from right to left and top to bottom, no symbol is counted fewer times than the symbol that came after it. The product is the sum of all diagrams so obtained.

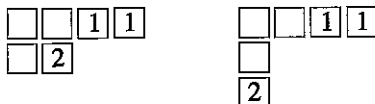
To illustrate the procedure, consider the product



We have put two 1's in the first row and one 2 in the second row of the frame to the right. Now apply the first row to the frame on the left. The result is



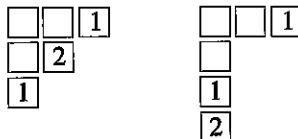
Now we apply the 2 to each of these graphs separately. We cannot put a 2 to the right of the 1's, because in that case, as we count from right to left, we would start with a 2 without having counted any 1's. The allowed graphs obtained from the first diagram are



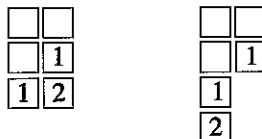
Applying the 2 to the second graph, we obtain



and to the third graph gives



Finally the last graph yields



The entire process described above is written in terms of frames as

$$\begin{array}{c}
 \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} \otimes \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} = \begin{array}{|c|c|c|c|} \hline \square & \square & \square & \square \\ \hline \square & \square & & \\ \hline \end{array} + \begin{array}{|c|c|c|c|} \hline \square & \square & \square & \square \\ \hline \square & & & \\ \hline \square & & & \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array} \\
 \\
 + 2 \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \\ \hline \square & & \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & & \\ \hline \square & & \\ \hline \end{array} + \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} + \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \\ \hline \end{array}
 \end{array}$$

Some simple products, some of which will be used later, are given in Figure 24.2.

### 24.9 Problems

**24.1.** Show that the action of a group  $G$  on the space of functions  $\psi$  given by  $T_g \psi(\mathbf{x}) = \psi(g^{-1} \cdot \mathbf{x})$  is a representation of  $G$ .

**24.2.** Complete Example 24.1.6.

**24.3.** Let the vector space carrying the representation of  $S_3$  be the space of functions. Choose  $\psi_1(x, y, z) \equiv xy$  and find the matrix representation of  $S_3$  in the minimal invariant subspace containing  $\psi_1$ . Hint: See Example 24.1.8.

**24.4.** Let the vector space carrying the representation of  $S_3$  be the space of functions. Choose (a)  $\psi_1(x, y, z) \equiv x$  and (b)  $\psi_1(x, y, z) \equiv x^2$ , and in each case, find the matrix representation of  $S_3$  in the minimal invariant subspace containing  $\psi_1$ .

**24.5.** Show that the representations  $T, \bar{T},$  and  $T^*$  are either all reducible or all irreducible.

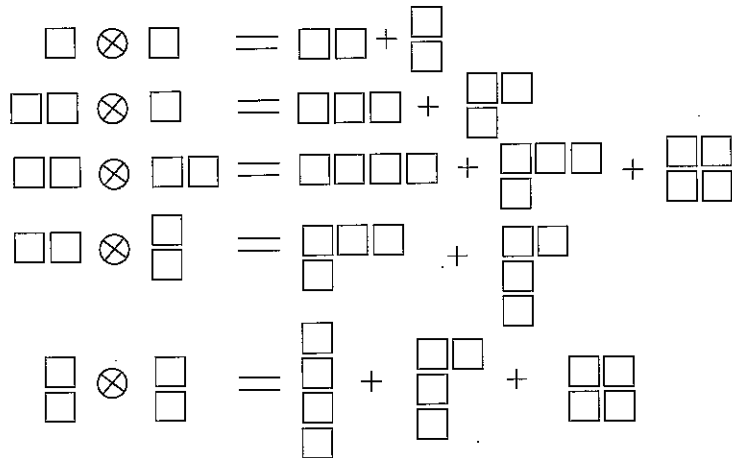
**24.6.** Show that the hermitian conjugate of Equation (24.5) gives  $T_g A^\dagger = A^\dagger T_g'$ . Using this relation show that  $S \equiv A^\dagger A$  commutes with all  $T_g$ 's. This result is used to prove Schur's lemmas in infinite dimensions.

**24.7.** Show that elements of a group belonging to the same conjugacy class have the same characters.

**24.8.** Show that the regular representation is indeed a representation, i.e., that  $R : G \rightarrow GL(m, \mathbb{R})$  is a homomorphism.

**24.9.** Show that a left ideal is minimal if and only if it is generated by a primitive idempotent.

**24.10.** Let  $G$  be a finite group. Define the element  $P = \sum_{x \in G} x$  of the group algebra and show that the left ideal generated by  $P$  is one-dimensional.



**Figure 24.2** Some products of Young frames for small values of  $n$ .

**24.11.** Show that  $T_i^{(\alpha)}$  defined in Equation (24.25) commutes with all operators  $T_g^{(\alpha)}$ . Hint: Consider  $T_g^{(\alpha)} T_i^{(\alpha)} (T_g^{(\alpha)})^{-1}$ .

- 24.12.** Let  $K_{i'}$  denote the set of inverses of a conjugacy class  $K_i$  with  $c_i$  elements.
- (a) Show that  $K_{i'}$  is also a class with  $c_i$  elements.
  - (b) Show that identity occurs exactly  $c_i$  times in the product  $\kappa_i \kappa_{i'}$ , and none in the product  $\kappa_i \kappa_j$  if  $j \neq i'$  [see Equation (24.23)].
  - (c) Conclude that

$$c_{ij1} = \begin{cases} 0 & \text{if } j \neq i', \\ c_i & \text{if } j = i'. \end{cases}$$

**24.13.** Show that the coefficients  $|G|d_j/|H|c_i$  of Equation (24.35) are integers.

**24.14.** Show that the symmetric and the antisymmetric representations of  $S_n$  are inequivalent.

**24.15.** Construct the character table for  $S_4$  from that of  $S_3$  (given as Table 24.4), and verify that it is given by Table 24.8.

**24.16.** Show that all functions transforming according to a given row of an irreducible representation have the same norm.

**24.17.** Show that if the group of symmetry of  $\mathbf{V}$  contains that of  $H_0$  and  $|\phi_j^{(\alpha)}\rangle$  belongs to the  $j$ th column of the  $\alpha$ th irreducible representation, then so does  $\mathbf{V}|\phi_j^{(\alpha)}\rangle$ . Conclude that  $\langle \phi_i^{(\alpha)} | \mathbf{V} | \phi_j^{(\alpha)} \rangle = 0$  for  $i \neq j$ .

**24.18.** Find the irreducible components of the representation of Example 24.1.6.

	${}^1K_1$	${}^6K_2$	${}^3K_3$	${}^8K_4$	${}^6K_5$
$T^{(1)}$	1	1	1	1	1
$T^{(2)}$	1	-1	1	1	-1
$T^{(3)}$	2	0	2	-1	0
$T^{(4)}$	3	1	-1	0	-1
$T^{(5)}$	3	-1	-1	0	1

**Table 24.8** Character table for  $S_4$ .

**24.19.** Show that  $\mathbf{P}^{(3)}|\psi_3\rangle$  of Example 24.6.3 is a linear combination of  $\mathbf{P}^{(3)}|\psi_1\rangle$  and  $\mathbf{P}^{(3)}|\psi_2\rangle$ .

**24.20.** Show that the tensor product of two unitary representations is unitary.

**24.21.** Switch the dummy indices of the double sum in (24.45), add (subtract) the two sums, and use  $(T \otimes T)_{ia,jb}(g) = (T \otimes T)_{ai,bj}(g)$  to show that the double sum can be written as a sum over the *symmetric (antisymmetric)* vectors alone.

**24.22.** Show that the characters  $\chi^S(g)$  and  $\chi^A(g)$  of the symmetrized and anti-symmetrized product representations are given, respectively, by

$$\chi^S(g) = \frac{1}{2}[(\chi(g))^2 + \chi(g^2)] \quad \text{and} \quad \chi^A(g) = \frac{1}{2}[(\chi(g))^2 - \chi(g^2)].$$

**24.23.** Suppose that  $\mathbf{A}_i^{(\alpha)}$  transforms according to  $T^{(\alpha)}$ , and  $\mathbf{A}_j^{(\beta)}$  according to  $T^{(\beta)}$ . Show that  $\mathbf{A}_i^{(\alpha)}\mathbf{A}_j^{(\beta)}$  transforms according to  $T^{(\alpha \times \beta)}$ .

**24.24.** Show that

$$\frac{1}{m_{\alpha\beta}} \sum_{ij} |\langle \phi_m^{(\gamma)} | \mathbf{A}_i^{(\alpha)} | \phi_j^{(\beta)} \rangle|^2 = |\langle \phi^{(\gamma)} | \mathbf{A}^{(\alpha)} | \phi^{(\beta)} \rangle_q|^2.$$

One can interpret this as the statement that the square of the reduced matrix element is proportional to the average (over  $i$  and  $j$ ) of the square of the full matrix elements.

**24.25.** Construct the character table of  $S_4$  using the analytical method and Equation (24.57).

**24.26.** Find all the standard Young tableaux for  $S_5$ . Thus, determine the dimension of each irreducible representations of  $S_5$ . Check that the dimensions satisfy the second equation of (24.18).

**24.27.** Verify the remaining entries of Table 24.6.

**24.28.** Construct the character table of  $S_5$ .

24.29. Suppose that  $Q$ , an element of the group algebra of  $S_n$ , is given by

$$Q = \sum_{i=1}^{n!} \epsilon_{\pi_i} \pi_i, \quad \pi_i \in S_n.$$

Show that

$$\pi_j Q = \epsilon_{\pi_j} Q \quad \text{and} \quad Q^2 = n!Q.$$

24.30. Show that  $Y^{(n)}Y^{(1^n)} = 0$ . Hint: There are as many even permutations in  $S_n$  as there are odd permutations.

24.31. Show that the product of any two Young operators of  $S_3$  corresponding to different Young tableaux is zero and that

$$Y_1^{(2,1)}Y_1^{(2,1)} = 3Y_1^{(2,1)}, \quad Y_2^{(2,1)}Y_2^{(2,1)} = 3Y_2^{(2,1)}.$$

24.32. Construct the ideal  $\mathcal{L}_2^{(2,1)}$  generated by  $Y_2^{(2,1)}$  and verify that it is two dimensional.

24.33. Using the ideal  $\mathcal{L}_1^{(2,1)}$  generated by  $Y_1^{(2,1)}$ , find the matrices of the irreducible representation  $T^{(2,1)}$ . From these matrices calculate the simple characters of  $S_3$  and compare your result with Table 24.4. Show that the ideal  $\mathcal{L}_2^{(2,1)}$  generated by  $Y_2^{(2,1)}$  gives the same set of characters.

24.34. Find all the Young operators for  $S_4$  corresponding to the first entry of each row of Figure 24.1. Find the ideals  $\mathcal{L}_1^{(3,1)}$  and  $\mathcal{L}_1^{(2^2)}$  generated by the Young operators  $Y_1^{(3,1)}$  and  $Y_1^{(2^2)}$  corresponding to the second and third rows of the table. Show that  $\mathcal{L}_1^{(3,1)}$  and  $\mathcal{L}_1^{(2^2)}$  have 3 and 2 dimensions, respectively.

24.35. Verify the products of the Young frames of Figure 24.2.

## Additional Reading

1. Barut, A. and Raczka, R. *Theory of Group Representations and Applications*, World Scientific, 1986. A comprehensive treatise on group theory and its application to physics written in a formal, but readable, language.
2. Boerner, H. *Representation of Groups*, North-Holland, 1963. A detailed introduction to group representation using group algebra techniques.
3. Elliott, J. and Dawber, P. *Symmetry in Physics* (2 volumes), Oxford University Press, 1979. Another informal book written for physicists.



4. Fulton, W. and Harris, J. *Representation Theory*, Springer-Verlag, 1991. A formal introduction to representation theory designed for graduate students in mathematics, but also useful for physicists with a solid math background.
5. Hamermesh, M. *Group Theory and its Application to Physical Problems*, Dover, 1989. A detailed and very readable account of the representation of symmetric groups written for physicists.

# Algebra of Tensors

Up until a mere two decades ago, tensors were almost completely synonymous with (general) relativity—except for a minor use in hydrodynamics. Students of physics did not need to study tensors until they took a course in the general theory of relativity. Then they would read the introductory chapter on tensor algebra and analysis, solve a few problems to condition themselves for index “gymnastics,” read through the book, learn some basic facts about relativity, and finally abandon it (unless they became relativists).

Today, with the advent of gauge theories of fundamental particles, the realization that gauge fields are to be thought of as geometrical objects, and the widespread belief that all fundamental interactions (including gravity) are different manifestations of the same superforce, the picture has changed drastically.

Two important developments have taken place as a consequence: Tensors have crept into other interactions besides gravity (such as the weak and strong nuclear interactions), and the geometrical (coordinate-independent) aspects of tensors have become more and more significant in the study of all interactions. The coordinate-independent study of tensors is the focus of the fascinating field of differential geometry and Lie groups, the subject of the remainder of the book. This chapter covers tensor algebra, while the next is devoted to tensor analysis.

As is customary, we will consider only real vector spaces and abandon the Dirac bra and ket notation, whose implementation is most advantageous in unitary (complex) spaces. From here on, the basis vectors<sup>1</sup> of a vector space  $\mathcal{V}$  will be distinguished by a subscript and those of its dual space by a superscript. If  $\{\mathbf{e}_i\}_{i=1}^N$  is a basis in  $\mathcal{V}$ , then  $\{\epsilon^j\}_{j=1}^N$  is a basis in  $\mathcal{V}^*$ . Also, **Einstein's summation convention** will be used:

Einstein's summation convention

---

<sup>1</sup>We denote vectors by roman boldface, and tensors of higher rank by sans serif bold letters.

**25.0.1. Box.** Repeated indices, of which one is an upper and the other a lower index, are assumed to be summed over:  $a_i^k b_j^i$  means  $\sum_{i=1}^N a_i^k b_j^i$ .

As a result of this convention, it is more natural to label the elements of a matrix representation of an operator  $\mathbf{A}$  by  $\alpha_i^j$  (rather than  $\alpha_{ji}$ ), because then  $\mathbf{A}e_i = \alpha_i^j e_j$ .

## 25.1 Multilinear Mappings

Since tensors are special kinds of linear operators on vector spaces, let us reconsider  $\mathcal{L}(\mathcal{V}, \mathcal{W})$ , the space of all linear mappings from the real vector space  $\mathcal{V}$  to the real vector space  $\mathcal{W}$ . We noted in Chapter 3 that  $\mathcal{L}(\mathcal{V}, \mathcal{W})$  is isomorphic to a space with dimension  $\dim \mathcal{V} \cdot \dim \mathcal{W}$ . The following proposition—whose proof we leave to the reader—shows this directly.

**25.1.1. Proposition.** Let  $\{e_i\}_{i=1}^{N_1}$  be a basis for  $\mathcal{V}$  and  $\{e'_\alpha\}_{\alpha=1}^{N_2}$  a basis for  $\mathcal{W}$ . Then

1. the linear transformations  $\mathbf{T}_\beta^j : \mathcal{V} \rightarrow \mathcal{W}$  in the vector space  $\mathcal{L}(\mathcal{V}, \mathcal{W})$ , defined by (note the new way of writing the Kronecker delta)

$$\mathbf{T}_\beta^j e_i = \delta_i^j e'_\beta, \quad j = 1, \dots, N_1; \beta = 1, \dots, N_2,$$

form a basis in  $\mathcal{L}(\mathcal{V}, \mathcal{W})$ . In particular,  $\dim \mathcal{L}(\mathcal{V}, \mathcal{W}) = N_1 N_2$ .

2. If  $\tau_j^\alpha$  are the matrix elements of a linear transformation  $\mathbf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ , then

$$\mathbf{T} = \tau_j^\alpha \mathbf{T}_\alpha^j.$$

The dual space  $\mathcal{V}^*$  is simply the space  $\mathcal{L}(\mathcal{V}, \mathbb{R})$ . Proposition 25.1.1 (with  $N_2 = 1$ ) then implies that  $\dim \mathcal{V}^* = \dim \mathcal{V}$ , which was shown in Chapter 1. The dual space is important in the discussion of tensors, so we consider some of its properties below.

The basis  $\{\mathbf{T}_\beta^j\}$  of Proposition 25.1.1 reduces to  $\{\mathbf{T}_1^j\}$  when  $\mathcal{W} = \mathbb{R}$  and is denoted by  $\{\epsilon^j\}_{j=1}^N$ , with  $N = \dim \mathcal{V}^* = \dim \mathcal{V}$ . The  $\epsilon^j$ 's have the property that

$$\epsilon^j(e_i) = \delta_i^j. \quad (25.1)$$

This relation was also established in Chapter 1. The basis  $B^* = \{\epsilon^j\}_{j=1}^N$  is simply the dual of the basis  $B = \{e_i\}_{i=1}^N$ . Note the “natural” position of the indices for  $B$  and  $B^*$ .

Now suppose that  $\{f_i\}_{i=1}^N = B'$  is another basis of  $\mathcal{V}$  and  $\mathbf{R}$  is the (invertible) matrix carrying  $B$  onto  $B'$ . Let  $B'^* = \{\varphi^j\}_{j=1}^N$  be the dual of  $B'$ . We want to find the matrix that carries  $B^*$  onto  $B'^*$ . If we denote this matrix by  $\mathbf{A}$  and its elements by  $a_i^j$ , we have

$$\delta_i^k = \varphi^k f_i = (a_i^k \epsilon^l)(r_l^j e_j) = a_i^k r_l^j \delta_j^l = a_i^k r_l^j = (\mathbf{A}\mathbf{R})_i^k,$$

where the first equality follows from the duality of  $B'$  and  $B'^*$ . In matrix form, this equation can be written as  $AR = 1$ , or  $A = R^{-1}$ . Thus,

**25.1.2. Box.** *The matrix that transforms bases of  $\mathcal{V}^*$  is the inverse of the matrix that transforms the corresponding bases of  $\mathcal{V}$ .*

In the above equations the upper index in matrix elements labels *rows*, and the lower index labels *columns*. This can be remembered by noting that the column vectors  $\mathbf{e}_i$  can be thought of as columns of a matrix, and the lower index  $i$  then labels those columns. Similarly,  $\epsilon^j$  can be thought of as rows of a matrix. We now generalize the concept of linear functionals.

multilinear mappings  
defined

**25.1.3. Definition.** *A map  $\mathbf{T} : \mathcal{V}_1 \times \mathcal{V}_2 \times \cdots \times \mathcal{V}_r \rightarrow \mathcal{W}$  is called  $r$ -linear if it is linear in all its variables:*

$$\begin{aligned} \mathbf{T}(\mathbf{v}_1, \dots, \alpha \mathbf{v}_i + \alpha' \mathbf{v}'_i, \dots, \mathbf{v}_r) \\ = \alpha \mathbf{T}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_r) + \alpha' \mathbf{T}(\mathbf{v}_1, \dots, \mathbf{v}'_i, \dots, \mathbf{v}_r) \end{aligned}$$

for all  $i$ .

We can easily construct a bilinear map. Let  $\tau_1 \in \mathcal{V}_1^*$  and  $\tau_2 \in \mathcal{V}_2^*$ . We define the map  $\tau_1 \otimes \tau_2 : \mathcal{V}_1 \times \mathcal{V}_2 \rightarrow \mathbb{R}$  by

$$\tau_1 \otimes \tau_2(\mathbf{v}_1, \mathbf{v}_2) = \tau_1(\mathbf{v}_1)\tau_2(\mathbf{v}_2). \quad (25.2)$$

The expression  $\tau_1 \otimes \tau_2$  is called the **tensor product** of  $\tau_1$  and  $\tau_2$ . Clearly, since  $\tau_1$  and  $\tau_2$  are separately linear, so is  $\tau_1 \otimes \tau_2$ .

An  $r$ -linear map can be multiplied by a scalar, and two  $r$ -linear maps can be added; in each case the result is an  $r$ -linear map. Thus, the set of  $r$ -linear maps from  $\mathcal{V}_1 \times \cdots \times \mathcal{V}_r$  into  $\mathcal{W}$  forms a vector space that is denoted by  $\mathcal{L}(\mathcal{V}_1, \dots, \mathcal{V}_r; \mathcal{W})$ .

We can also construct multilinear maps on the dual space. First, we note that we can define a natural linear functional on  $\mathcal{V}^*$  as follows. We let  $\tau \in \mathcal{V}^*$  and  $\mathbf{v} \in \mathcal{V}$ ; then  $\tau(\mathbf{v}) \in \mathbb{R}$ . Now we twist this around and *define* a mapping  $\mathbf{v} : \mathcal{V}^* \rightarrow \mathbb{R}$  given by  $\mathbf{v}(\tau) \equiv \tau(\mathbf{v})$ . It is easily shown that this mapping is linear. Thus, we have naturally constructed a linear functional on  $\mathcal{V}^*$  by identifying  $(\mathcal{V}^*)^*$  with  $\mathcal{V}$ .

Construction of multilinear maps on  $\mathcal{V}^*$  is now trivial. For example, let  $\mathbf{v}_1 \in \mathcal{V}_1$  and  $\mathbf{v}_2 \in \mathcal{V}_2$  and define the tensor product  $\mathbf{v}_1 \otimes \mathbf{v}_2 : \mathcal{V}_1^* \times \mathcal{V}_2^* \rightarrow \mathbb{R}$  by

$$\mathbf{v}_1 \otimes \mathbf{v}_2(\tau_1, \tau_2) = \mathbf{v}_1(\tau_1)\mathbf{v}_2(\tau_2) = \tau_1(\mathbf{v}_1)\tau_2(\mathbf{v}_2). \quad (25.3)$$

We can also construct mixed multilinear maps such as  $\mathbf{v} \otimes \tau : \mathcal{V}^* \times \mathcal{V} \rightarrow \mathbb{R}$  given by

$$\mathbf{v} \otimes \tau(\theta, \mathbf{u}) = \mathbf{v}(\theta)\tau(\mathbf{u}) = \theta(\mathbf{v})\tau(\mathbf{u}). \quad (25.4)$$

natural pairing of vectors and their duals

There is a bilinear map  $h : \mathcal{V}^* \times \mathcal{V} \rightarrow \mathbb{R}$  that naturally pairs  $\mathcal{V}$  and  $\mathcal{V}^*$ ; it is given by  $h(\theta, \mathbf{v}) \equiv \theta(\mathbf{v})$ . This mapping is called the **natural pairing** of  $\mathcal{V}$  and  $\mathcal{V}^*$  into  $\mathbb{R}$  and is denoted by using angle brackets:

$$h(\theta, \mathbf{v}) \equiv \langle \theta, \mathbf{v} \rangle \equiv \theta(\mathbf{v}).$$

tensors; covariant and contravariant degrees

**25.1.4. Definition.** Let  $\mathcal{V}$  be a vector space with dual space  $\mathcal{V}^*$ . Then a **tensor of type  $(r, s)$**  is a multilinear mapping<sup>2</sup>

$$T_{rs}^r : \underbrace{\mathcal{V}^* \times \mathcal{V}^* \times \cdots \times \mathcal{V}^*}_{r \text{ times}} \times \underbrace{\mathcal{V} \times \mathcal{V} \times \cdots \times \mathcal{V}}_{s \text{ times}} \rightarrow \mathbb{R}.$$

The set of all such mappings for fixed  $r$  and  $s$  forms a vector space denoted by  $T_s^r(\mathcal{V})$ . The number  $r$  is called the **contravariant degree** of the tensor, and  $s$  is called the **covariant degree** of the tensor.

As an example, let  $\mathbf{v}_1, \dots, \mathbf{v}_r \in \mathcal{V}$  and  $\tau^1, \dots, \tau^s \in \mathcal{V}^*$ , and define the tensor product  $T_s^r \equiv \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r \otimes \tau^1 \otimes \cdots \otimes \tau^s$  by

$$\begin{aligned} & \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r \otimes \tau^1 \otimes \cdots \otimes \tau^s(\theta^1, \dots, \theta^r, \mathbf{u}_1, \dots, \mathbf{u}_s) \\ &= \mathbf{v}_1(\theta^1) \dots \mathbf{v}_r(\theta^r) \tau^1(\mathbf{u}_1) \dots \tau^s(\mathbf{u}_s) = \prod_{i=1}^r \prod_{j=1}^s \theta^i(\mathbf{v}_i) \tau^j(\mathbf{u}_j). \end{aligned}$$

Each  $\mathbf{v}$  in the tensor product requires an element of  $\mathcal{V}^*$ ; that is why the number of factors of  $\mathcal{V}^*$  in the Cartesian product equals the number of  $\mathbf{v}$ 's in the tensor product. As explained in Chapter 0, the Cartesian product with  $s$  factors of  $\mathcal{V}$  is denoted by  $\mathcal{V}^s$  (similarly for  $\mathcal{V}^*$ ).

contravariant and covariant vectors and tensors

A tensor of type  $(0, 0)$  is defined to be a scalar, so  $T_0^0(\mathcal{V}) = \mathbb{R}$ . A tensor of type  $(1, 0)$ , an ordinary vector, is called a **contravariant vector**, and one of type  $(0, 1)$ , a dual vector (or a linear functional), is called a **covariant vector**. A tensor of type  $(r, 0)$  is called a **contravariant tensor** of rank  $r$ , and one of type  $(0, s)$  is called a **covariant tensor** of rank  $s$ . The union of  $T_s^r(\mathcal{V})$  for all possible  $r$  and  $s$  can be made into an (infinite-dimensional) algebra, called the **algebra of tensors**, by defining the following product on it:

tensors form an algebra

multiplication of the algebra of tensors

**25.1.5. Definition.** The **tensor product** of a tensor  $T$  of type  $(r, s)$  and a tensor  $U$  of type  $(k, l)$  is a tensor  $T \otimes U$  of type  $(r + k, s + l)$ , defined, as an operator on  $(\mathcal{V}^*)^{r+k} \times \mathcal{V}^{s+l}$ , by

$$\begin{aligned} & T \otimes U(\theta^1, \dots, \theta^{r+k}, \mathbf{u}_1, \dots, \mathbf{u}_{s+l}) \\ &= T(\theta^1, \dots, \theta^r, \mathbf{u}_1, \dots, \mathbf{u}_s) U(\theta^{r+1}, \dots, \theta^{r+k}, \mathbf{u}_{s+1}, \dots, \mathbf{u}_{s+l}). \end{aligned}$$

This product turns the (infinite-dimensional) vector space of all tensors into an associative algebra called a **tensor algebra**.

<sup>2</sup>Just as the space  $\mathcal{V}^*$  of linear functionals of a vector space  $\mathcal{V}$  is isomorphic to  $\mathcal{V}$ , so is the space of tensors of type  $(r, s)$  to the tensor product space  $\mathcal{V}_{r,s}$  (see Example 1.3.19). In fact,  $T_s^r(\mathcal{V}) = \mathcal{V}_{r,s}^*$ , as shown in [Warner, 83] on p. 58.

This definition is a generalization of Equations (25.2), (25.3), and (25.4). It is easily verified that the tensor product is associative and distributive (over tensor addition), *but not commutative*.

Making computations with tensors requires choosing a basis for  $\mathcal{V}$  and one for  $\mathcal{V}^*$  and representing the tensors in terms of numbers (components). This process is not, of course, new. Linear operators are represented by arrays of numbers, i.e., matrices. The case of tensors is merely a generalization of that of linear operators and can be stated as follows:

**25.1.6. Theorem.** *Let  $\{\mathbf{e}_i\}_{i=1}^N$  be a basis in  $\mathcal{V}$ , and  $\{\epsilon^j\}_{j=1}^N$  a basis in  $\mathcal{V}^*$ , usually taken to be the dual of  $\{\mathbf{e}_i\}_{i=1}^N$ . Then the set of all tensor products  $\mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s}$  forms a basis for  $\mathcal{T}_s^r(\mathcal{V})$ . Furthermore, the components of any tensor  $\mathbf{A} \in \mathcal{T}_s^r(\mathcal{V})$  are*

$$A_{i_1 \dots i_s}^{j_1 \dots j_r} = \mathbf{A}(\epsilon^{j_1}, \dots, \epsilon^{j_r}, \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_s}).$$

*Proof.* The proof is a simple exercise in employing the definition of tensor products and keeping track of the multilinear property of tensors. Details are left for the reader. □

A useful result of the theorem is the relation

$$\mathbf{A} = A_{j_1 \dots j_s}^{i_1 \dots i_r} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s}. \tag{25.5}$$

Note that for every factor in the basis vectors of  $\mathcal{T}_s^r(\mathcal{V})$  there are  $N$  possibilities, each one giving a new linearly independent vector of the basis. Thus, the number of possible tensor products is  $N^{r+s}$ , and we have  $\dim \mathcal{T}_s^r(\mathcal{V}) = (\dim \mathcal{V})^{r+s}$ .

**25.1.7. Example.** Let us consider the special case of  $\mathcal{T}_1^1(\mathcal{V})$  as an illustration. We can write  $\mathbf{A} \in \mathcal{T}_1^1(\mathcal{V})$  as  $\mathbf{A} = A_j^i \mathbf{e}_i \otimes \epsilon^j$ . Given any  $\mathbf{v} \in \mathcal{V}$ , we obtain<sup>2</sup>

$$\mathbf{A}(\mathbf{v}) = (A_j^i \mathbf{e}_i \otimes \epsilon^j)(\mathbf{v}) = A_j^i \mathbf{e}_i \underbrace{[\epsilon^j(\mathbf{v})]}_{\in \mathbb{R}}.$$

This shows that  $\mathbf{A}(\mathbf{v}) \in \mathcal{V}$  and  $\mathbf{A}$  can be interpreted as a linear operator on  $\mathcal{V}$ , i.e.,  $\mathbf{A} \in \mathcal{L}(\mathcal{V})$ . Similarly, for  $\tau \in \mathcal{V}^*$  we get

$$\mathbf{A}(\tau) = (A_j^i \mathbf{e}_i \otimes \epsilon^j)(\tau) = A_j^i \underbrace{[\mathbf{e}_i(\tau)]}_{\in \mathbb{R}} \epsilon^j.$$

$\mathcal{T}_1^1(\mathcal{V})$ ,  $\mathcal{L}(\mathcal{V})$ , and  $\mathcal{L}(\mathcal{V}^*)$  are all the same

Thus,  $\mathbf{A} \in \mathcal{L}(\mathcal{V}^*)$ . We have shown that given  $\mathbf{A} \in \mathcal{T}_1^1(\mathcal{V})$ , there corresponds a linear operator belonging to  $\mathcal{L}(\mathcal{V})$  [or  $\mathcal{L}(\mathcal{V}^*)$ ] having a natural relation to  $\mathbf{A}$ . Similarly, given any  $\mathbf{A} \in \mathcal{L}(\mathcal{V})$  [or  $\mathcal{L}(\mathcal{V}^*)$ ] with a matrix representation in the basis  $\{\mathbf{e}_i\}_{i=1}^N$  of  $\mathcal{V}$  (or  $\{\epsilon^j\}_{j=1}^N$  of  $\mathcal{V}^*$ ) given by  $A_j^i$ , then corresponding to it in a natural way is a tensor in  $\mathcal{T}_1^1(\mathcal{V})$ , namely  $A_j^i \mathbf{e}_i \otimes \epsilon^j$ . Problem 25.5 shows that the tensor defined in this way is basis-independent. Therefore,

natural isomorphism there is a *natural* one-to-one correspondence among  $\mathcal{T}_1^1(\mathcal{V})$ ,  $\mathcal{L}(\mathcal{V})$ , and  $\mathcal{L}(\mathcal{V}^*)$ . This natural correspondence is called a **natural isomorphism**. Whenever there is a natural isomorphism between two vector spaces, those vector spaces can be treated as being the same. ■

We have defined tensors as multilinear machines that take in a vector from a *specific* Cartesian product space of  $\mathcal{V}$ 's and  $\mathcal{V}^*$ 's and manufacture a real number. Given the representation in Equation (25.5), however, we can generalize the interpretation of a tensor as a linear machine so that it takes a vector belonging to a Cartesian product space and manufactures a tensor. This corresponds to a situation in which not all factors of (25.5) find “partners.” An illustration of this situation was presented in the preceding example. To clarify this let us consider  $\mathbf{A} \in \mathcal{T}_2^1(\mathcal{V})$ , given by

$$\mathbf{A} = A_{jk}^i \mathbf{e}_i \otimes \epsilon^j \otimes \epsilon^k.$$

This machine needs a Cartesian-product vector of the form  $(\tau, \mathbf{v}_1, \mathbf{v}_2)$ , with  $\tau \in \mathcal{V}^*$  and  $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ , to give a real number. However, if it is not fed enough, it will not complete its job. For instance, if we feed it only a dual vector  $\tau \in \mathcal{V}^*$ , it will give a tensor belonging to  $\mathcal{T}_2^0(\mathcal{V})$ :

$$\mathbf{A}(\tau) = (A_{jk}^i \mathbf{e}_i \otimes \epsilon^j \otimes \epsilon^k)(\tau) = A_{jk}^i [\mathbf{e}_i(\tau)] \epsilon^j \otimes \epsilon^k.$$

If we feed it a double vector  $(\mathbf{v}_1, \mathbf{v}_2)$ , it will manufacture a vector in  $\mathcal{V}$ :

$$\mathbf{A}(\mathbf{v}_1, \mathbf{v}_2) = (A_{jk}^i \mathbf{e}_i \otimes \epsilon^j \otimes \epsilon^k)(\mathbf{v}_1, \mathbf{v}_2) = A_{jk}^i \mathbf{e}_i [\epsilon^j(\mathbf{v}_1)] [\epsilon^k(\mathbf{v}_2)] \in \mathcal{V}$$

positioning of vectors and their duals to match the tensor

What if we feed it a single vector  $\mathbf{v}$ ? It will blow its whistles and buzz its buzzers, because it does not know whether to give  $\mathbf{v}$  to  $\epsilon^j$  or  $\epsilon^k$  (it is smart enough to know that it cannot give  $\mathbf{v}$  to  $\mathbf{e}_i$ ). That is why we have to inform the machine with which factor of  $\epsilon$  to match  $\mathbf{v}$ . This is done by properly positioning  $\mathbf{v}$  inside a pair of parentheses: If we write  $(\cdot, \mathbf{v}, \cdot)$ , the machine will know that  $\mathbf{v}$  belongs to  $\epsilon^j$ , and  $(\cdot, \cdot, \mathbf{v})$  tells the machine to pair  $\mathbf{v}$  with  $\epsilon^k$ . If we write  $(\mathbf{v}, \cdot, \cdot)$ , the machine will give us an “error message” because it cannot pair  $\mathbf{v}$  with  $\mathbf{e}_i$ !

The components of a tensor  $\mathbf{A}$ , as given in Equation (25.5), depend on the basis in which they are described. If the basis is changed, the components change. The relation between components of a tensor in different bases is called the **transformation law** for that particular tensor. Let us investigate this concept.

We use overbars to distinguish among various bases. For instance,  $B = \{\mathbf{e}_i\}_{i=1}^N$ ,  $\overline{B} = \{\overline{\mathbf{e}}_j\}_{j=1}^N$ , and  $\overline{\overline{B}} = \{\overline{\overline{\mathbf{e}}}_k\}_{k=1}^N$  are three different bases of  $\mathcal{V}$ . Similarly,  $B^* = \{\epsilon^i\}_{i=1}^N$ ,  $\overline{B}^* = \{\overline{\epsilon}^j\}_{j=1}^N$ , and  $\overline{\overline{B}}^* = \{\overline{\overline{\epsilon}}^k\}_{k=1}^N$  are bases of  $\mathcal{V}^*$ . The components are also distinguished with overbars. Recall that if  $\mathbf{R}$  is the matrix connecting  $B$  and

<sup>2</sup>Here, we are assuming that  $\mathbf{A}$  acts on an object (such as  $\mathbf{v}$ ) by “pairing it up” with an appropriate factor of which  $\mathbf{A}$  is composed (such as  $\epsilon^j$ ).

$\bar{B}$ , then  $S = R^{-1}$  connects  $B^*$  and  $\bar{B}^*$ . For a tensor  $\mathbf{A}$  of type  $(1, 2)$ , Theorem 25.1.6 gives

$$\begin{aligned}\bar{A}^i_{jk} &= \mathbf{A}(\bar{\epsilon}^i, \bar{\epsilon}_j, \bar{\epsilon}_k) = \mathbf{A}(s_m^i \epsilon^m, r_j^n \epsilon_n, r_k^p \epsilon_p) \\ &= s_m^i r_j^n r_k^p \mathbf{A}(\epsilon^m, \epsilon_n, \epsilon_p) = s_m^i r_j^n r_k^p A_{np}^m.\end{aligned}\quad (25.6)$$

This is the law that transforms the components of  $\mathbf{A}$  from one basis to another. In the classical coordinate-dependent treatment of tensors, Equation (25.6) was the *defining* relation for a tensor of type  $(1, 2)$ . In other words, a tensor of type  $(1, 2)$  was defined to be a collection of numbers,  $A_{np}^m$  that transformed to another collection of numbers  $\bar{A}^i_{jk}$  according to the rule in (25.6) when the basis was changed. In the modern treatment of tensors it is not necessary to introduce any basis to define tensors. Only when the components of a tensor are needed must bases be introduced. The advantage of the modern treatment is obvious, since a  $(1, 2)$ -type tensor has 27 components in three dimensions and 64 components in four dimensions, and all of these are represented by the single symbol  $\mathbf{A}$ . However, the role of components should not be downplayed. After all, when it comes to actual calculations, we are forced to choose a basis and manipulate components.

Since  $\mathcal{J}_s^r(\mathcal{V})$  are vector spaces, it is desirable to investigate mappings from one such space to another. We will be particularly interested in linear mappings. For example,  $\mathbf{f} : \mathcal{J}_0^1(\mathcal{V}) \rightarrow \mathcal{J}_0^0(\mathcal{V}) = \mathbb{R}$  is what was called a linear functional before. Similarly,  $\mathbf{t} : \mathcal{J}_0^1(\mathcal{V}) \rightarrow \mathcal{J}_0^1(\mathcal{V})$  is a linear transformation on  $\mathcal{V}$ . A special linear transformation is  $\text{tr} : \mathcal{J}_1^1(\mathcal{V}) \rightarrow \mathcal{J}_0^0(\mathcal{V}) = \mathbb{R}$ , given by

$$\text{tr } \mathbf{A} = \text{tr}(A_j^i \epsilon_i \otimes \epsilon^j) \equiv A_i^i \equiv \sum_{i=1}^N A_i^i.$$

This is the same trace encountered in the study of linear transformations in Chapter 3.

Although the above definition makes explicit use of components with respect to a basis, it can be easily shown (see Problem 25.7) that it is in fact basis-independent. Functions of tensors that do not depend on bases are called *invariants*. Another example of an invariant is a linear functional (see Problem 25.6).

**25.1.8. Example.** Consider the tensor  $\mathbf{A} \in \mathcal{J}_0^2(\mathcal{V})$  given by  $\mathbf{A} = \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_1$ . We calculate the analogue of the trace for  $\mathbf{A}$ :  $\sum_{i=1}^2 A_{ii} = 1 + 0 = 1$ . Now we change to a new basis,  $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2\}$ , given by  $\mathbf{e}_1 = \bar{\mathbf{e}}_1 + 2\bar{\mathbf{e}}_2$  and  $\mathbf{e}_2 = -\bar{\mathbf{e}}_1 + \bar{\mathbf{e}}_2$ . In terms of the new basis vectors,  $\mathbf{A}$  is given by

$$\begin{aligned}\mathbf{A} &= (\bar{\mathbf{e}}_1 + 2\bar{\mathbf{e}}_2) \otimes (\bar{\mathbf{e}}_1 + 2\bar{\mathbf{e}}_2) + (-\bar{\mathbf{e}}_1 + \bar{\mathbf{e}}_2) \otimes (\bar{\mathbf{e}}_1 + 2\bar{\mathbf{e}}_2) \\ &= 3\bar{\mathbf{e}}_2 \otimes \bar{\mathbf{e}}_1 + 6\bar{\mathbf{e}}_2 \otimes \bar{\mathbf{e}}_2\end{aligned}$$

with  $\sum_{i=1}^2 \bar{A}_{ii} = 0 + 6 = 6 \neq \sum_{i=1}^2 A_{ii}$ . This kind of “trace” is not invariant. ■



Whenever a quantity is defined without reference to a basis, it is clearly invariant. The trace is an example of an *invariant quantity* whose definition *does* require a basis. Another example is the determinant of a linear operator [or a tensor of type (1, 1)], in terms of a matrix representation of that operator, which is clearly basis-dependent. A basis-independent definition for the determinant of a tensor will be given later in this chapter.

Besides mappings of the form  $h : \mathcal{T}_s^r(\mathcal{V}) \rightarrow \mathcal{T}_l^k(\mathcal{V})$  that depend on a single variable, we can define mappings that depend on several variables, in other words, that take several elements of  $\mathcal{T}_s^r(\mathcal{V})$  and give an element of  $\mathcal{T}_l^k(\mathcal{V})$ . We then write  $h : (\mathcal{T}_s^r(\mathcal{V}))^m \rightarrow \mathcal{T}_l^k(\mathcal{V})$ . It is understood that  $h(\mathbf{t}_1, \dots, \mathbf{t}_m)$ , in which all  $\mathbf{t}_i$  are in  $\mathcal{T}_s^r(\mathcal{V})$ , is a tensor of type  $(k, l)$ . If  $h$  is linear in all of its variables, it is called a **tensor-valued multilinear map**. Furthermore, if  $h(\mathbf{t}_1, \dots, \mathbf{t}_m)$  does not depend on the choice of a basis of  $\mathcal{T}_s^r(\mathcal{V})$ , it is called a *multilinear invariant*. In most cases  $k = 0 = l$ , and we speak of *scalar-valued invariants*, or simply invariants. An example of a multilinear invariant is the determinant considered as a function of the rows of a matrix.

The following defines an important class of multilinear invariants:

**25.1.9. Definition.** A *contraction* of a tensor  $\mathbf{A} \in \mathcal{T}_s^r(\mathcal{V})$  with respect to a contravariant index  $p$  and a covariant index  $q$  is a linear mapping  $\mathbf{C} : \mathcal{T}_s^r(\mathcal{V}) \rightarrow \mathcal{T}_{s-1}^{r-1}(\mathcal{V})$  given in components by

$$[\mathbf{C}(\mathbf{A})]_{j_1 \dots j_{s-1}}^{i_1 \dots i_{r-1}} = A_{j_1 \dots j_{q-1} k j_{q+1} \dots j_s}^{i_1 \dots i_{p-1} k i_{p+1} \dots i_r} \equiv \sum_k A_{j_1 \dots j_{q-1} k j_{q+1} \dots j_s}^{i_1 \dots i_{p-1} k i_{p+1} \dots i_r}$$

It can be readily shown that contractions are invariants. The proof is exactly the same as that for the invariance of the trace. In fact, the trace is a special case of a contraction, in which  $r = s = 1$ .

The linearity inherent in the construction of tensor algebras carries along some of the properties and structures of the underlying vector spaces. One such property is isomorphism. Suppose that  $\mathbf{F} : \mathcal{V} \rightarrow \mathcal{U}$  is a vector space isomorphism. Then  $\mathbf{F}^* : \mathcal{U}^* \rightarrow \mathcal{V}^*$ , the pullback of  $\mathbf{F}$ , is also an isomorphism (Proposition 1.3.18). Associated to  $\mathbf{F}$  is a linear map—which we denote by the same symbol—from  $\mathcal{T}_s^r(\mathcal{V})$  to  $\mathcal{T}_s^r(\mathcal{U})$  defined by

$$\underbrace{[\mathbf{F}(\mathbf{T})]}_{\in \mathcal{T}_s^r(\mathcal{U})}(\theta^1, \dots, \theta^r, \mathbf{u}_1, \dots, \mathbf{u}_s) \equiv \mathbf{T}(\mathbf{F}^*\theta^1, \dots, \mathbf{F}^*\theta^r, \mathbf{F}^{-1}\mathbf{u}_1, \dots, \mathbf{F}^{-1}\mathbf{u}_s), \tag{25.7}$$

where  $\mathbf{T} \in \mathcal{T}_s^r(\mathcal{V})$ ,  $\theta^i \in \mathcal{U}^*$ , and  $\mathbf{u}_j \in \mathcal{U}$ . The reader may check that this map is an algebra isomorphism (see Section 1.4). We shall use this isomorphism to define derivatives for tensors in the next chapter.

## 25.2 Symmetries of Tensors

Many applications demand tensors that have some kind of symmetry property. We have already encountered a symmetric tensor—the metric “tensor” of an inner product: If  $\mathcal{V}$  is a vector space and  $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ , then  $g(\mathbf{v}_1, \mathbf{v}_2) = g(\mathbf{v}_2, \mathbf{v}_1)$ . The following generalizes this property.

symmetric tensor  
defined

**25.2.1. Definition.** A tensor  $\mathbf{A}$  is *symmetric* in the  $i$ th and  $j$ th variables if its value as a multilinear function is unchanged when these variables are interchanged. Clearly, the two variables must be of the same kind.

An immediate consequence of this definition is that in any basis, the components of a symmetric tensor do not change when the  $i$ th and  $j$ th indices are interchanged.

contravariant-  
symmetric;  
covariant-symmetric;  
symmetric

**25.2.2. Definition.** A tensor is *contravariant-symmetric* if it is symmetric in every pair of its contravariant indices and *covariant-symmetric* if it is symmetric in every pair of its covariant indices. A tensor is *symmetric* if it is both contravariant-symmetric and covariant-symmetric.

The set of all symmetric tensors  $\mathcal{S}^r(\mathcal{V})$  of type  $(r, 0)$  forms a subspace of the vector space<sup>3</sup>  $\mathcal{T}_0^r$ . Similarly, the set of symmetric tensors of type  $(0, s)$  forms a subspace  $\mathcal{S}_s$  of  $\mathcal{T}_s^0$ . The (independent) components of a symmetric tensor  $\mathbf{A} \in \mathcal{S}^r$  are  $A_{i_1 i_2 \dots i_r}$ , where  $i_1 \leq i_2 \leq \dots \leq i_r$ ; the other components are given by symmetry.

Although a set of symmetric tensors forms a vector space, it does not form an algebra under the usual multiplication of tensors. In fact, even if  $\mathbf{A} = A^{ij} \mathbf{e}_i \otimes \mathbf{e}_j$  and  $\mathbf{B} = B^{kl} \mathbf{e}_k \otimes \mathbf{e}_l$  are symmetric tensors of type  $(2, 0)$ , the tensor product  $\mathbf{A} \otimes \mathbf{B} = A^{ij} B^{kl} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l$  need not be a type  $(4, 0)$  symmetric tensor. For instance,  $A^{ik} B^{jl}$  may not equal  $A^{ij} B^{kl}$ . However, we can modify the definition of the tensor product (for symmetric tensors) to give a symmetric product out of symmetric factors.

symmetrizer

**25.2.3. Definition.** A *symmetrizer* is an operator  $\mathbb{S} : \mathcal{T}_0^r \rightarrow \mathcal{S}^r$  given by

$$[\mathbb{S}(\mathbf{A})](\tau^1, \dots, \tau^r) = \frac{1}{r!} \sum_{\pi} \mathbf{A}(\tau^{\pi(1)}, \dots, \tau^{\pi(r)}), \tag{25.8}$$

where the sum is taken over the  $r!$  permutations of the integers  $1, 2, \dots, r$ , and  $\tau^1, \dots, \tau^r$  are all in  $\mathcal{V}^*$ .  $\mathbb{S}(\mathbf{A})$  is often denoted by  $\mathbf{A}_s$ . Clearly,  $\mathbf{A}_s$  is a symmetric tensor.

A similar definition gives the symmetrizer  $\mathbb{S} : \mathcal{T}_s^0 \rightarrow \mathcal{S}_s$ . Instead of  $\tau^1, \dots, \tau^r$  in (25.8), we would have  $\mathbf{v}_1, \dots, \mathbf{v}_s$ .

<sup>3</sup>When there is no risk of confusion, we shall delete  $\mathcal{V}$  from  $\mathcal{T}_s^r(\mathcal{V})$ , it being understood that all tensors are defined on some given underlying vector space.

**25.2.4. Example.** For  $r = 2$ , we have only two permutations, and

$$\mathbf{A}_s(\tau^1, \tau^2) = \frac{1}{2}[\mathbf{A}(\tau^1, \tau^2) + \mathbf{A}(\tau^2, \tau^1)]$$

For  $r = 3$ , we have six permutations of 1, 2, 3, and (25.8) gives

$$\begin{aligned} \mathbf{A}_s(\tau^1, \tau^2, \tau^3) &= \frac{1}{6}[\mathbf{A}(\tau^1, \tau^2, \tau^3) + \mathbf{A}(\tau^2, \tau^1, \tau^3) + \mathbf{A}(\tau^1, \tau^3, \tau^2) \\ &\quad + \mathbf{A}(\tau^3, \tau^1, \tau^2) + \mathbf{A}(\tau^3, \tau^2, \tau^1) + \mathbf{A}(\tau^2, \tau^3, \tau^1)] \end{aligned}$$

It is clear that interchanging any pair of  $\tau$ 's on the RHS of the above two equations does not change the sum. Thus,  $\mathbf{A}_s$  is indeed a symmetric tensor. ■

It can be shown that

$$\dim S^r(\mathcal{V}) = \binom{N+r-1}{r} \equiv \frac{(N+r-1)!}{r!(N-1)!}.$$

The proof involves counting the number of different integers  $i_1, \dots, i_r$  for which  $1 \leq i_m \leq i_{m+1} \leq N$  for each  $m$ .

We are now ready to define a product on the collection of symmetric tensors and make it an algebra, called the **symmetric algebra**.

symmetric tensors  
form an algebra  
under this product

**25.2.5. Definition.** The symmetric product of symmetric tensors  $\mathbf{A} \in S^r(\mathcal{V})$  and  $\mathbf{B} \in S^s(\mathcal{V})$  is denoted by  $\mathbf{AB}$  and defined as

$$\begin{aligned} \mathbf{AB}(\tau^1, \dots, \tau^{r+s}) &\equiv \frac{(r+s)!}{r!s!} \mathbf{S}(\mathbf{A} \otimes \mathbf{B})(\tau^1, \dots, \tau^{r+s}) \\ &= \frac{1}{r!s!} \sum_{\pi} \mathbf{A}(\tau^{\pi(1)}, \dots, \tau^{\pi(r)}) \mathbf{B}(\tau^{\pi(r+1)}, \dots, \tau^{\pi(r+s)}), \end{aligned}$$

where the sum is over all permutations of  $1, 2, \dots, r+s$ . The symmetric product of  $\mathbf{A} \in S_r(\mathcal{V})$  and  $\mathbf{B} \in S_s(\mathcal{V})$  is defined similarly.

**25.2.6. Example.** Let us construct the symmetric tensor products of vectors. First we find the symmetric product of  $\mathbf{v}_1$  and  $\mathbf{v}_2$  both belonging to  $\mathcal{V} = \mathcal{J}_0^1(\mathcal{V})$ :

$$\begin{aligned} (\mathbf{v}_1 \mathbf{v}_2)(\tau^1, \tau^2) &\equiv \mathbf{v}_1(\tau^1) \mathbf{v}_2(\tau^2) + \mathbf{v}_1(\tau^2) \mathbf{v}_2(\tau^1) \\ &= \mathbf{v}_1(\tau^1) \mathbf{v}_2(\tau^2) + \mathbf{v}_2(\tau^1) \mathbf{v}_1(\tau^2) \\ &= (\mathbf{v}_1 \otimes \mathbf{v}_2 + \mathbf{v}_2 \otimes \mathbf{v}_1)(\tau^1, \tau^2). \end{aligned}$$

Since this is true for any pair  $\tau^1$  and  $\tau^2$ , we have

$$\mathbf{v}_1 \mathbf{v}_2 = \mathbf{v}_1 \otimes \mathbf{v}_2 + \mathbf{v}_2 \otimes \mathbf{v}_1.$$

In general,  $\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_r = \sum_{\pi} \mathbf{v}_{\pi(1)} \otimes \mathbf{v}_{\pi(2)} \otimes \cdots \otimes \mathbf{v}_{\pi(r)}$ . ■

It is clear from the definition that symmetric multiplication is commutative, associative, and distributive. If we choose a basis  $\{\mathbf{e}_i\}_{i=1}^N$  for  $\mathcal{V}$  and express all symmetric tensors in terms of symmetric products of  $\mathbf{e}_i$  using the above properties, then any symmetric tensor can be expressed as a linear combination of terms of the form  $(\mathbf{e}_1)^{n_1} \cdots (\mathbf{e}_N)^{n_N}$ .

Skew-symmetry is the same as symmetry except that in the interchange of variables the tensor changes sign. The following theorem, whose simple proof we leave as a problem, further characterizes skew-symmetric tensors.

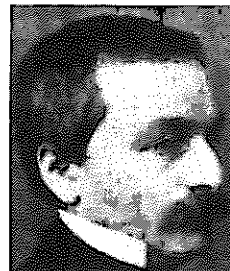
skew-symmetric  
tensors

**25.2.7. Theorem.** *A tensor  $\mathbf{A}$  is skew-symmetric in contravariant indices  $i$  and  $j$  if and only if for all  $\tau \in \mathcal{V}^*$ , substitution of  $\tau \in \mathcal{V}^*$  for both the  $i$ th and  $j$ th variables of  $\mathbf{A}$  yields zero, regardless of the values of the remaining variables.*

covariant and  
contravariant  
skew-symmetric  
tensors

**25.2.8. Definition.** *A covariant (contravariant) skew-symmetric tensor is one that is skew-symmetric in all pairs of covariant (contravariant) variables. A tensor is skew-symmetric if it is both covariant and contravariant skew-symmetric.*

**Leopold Kronecker** (1823–1891) was the son of Isidor Kronecker, a businessman, and Johanna Prausnitzer. They were wealthy and provided private tutoring at home for their son until he entered the Liegnitz Gymnasium. At the gymnasium, Kronecker's mathematics teacher was E. E. Kummer, who early recognized the boy's ability and encouraged him to do independent research. He also received Evangelical religious instruction, although he was Jewish; he formally converted to Christianity in the last year of his life.



Kronecker matriculated at the University of Berlin in 1841, where he attended lectures in mathematics given by Dirichlet. Like Gauss and Jacobi, he was interested in classical philology. He also attended Schelling's philosophy lectures; he was later to make a thorough study of the works of Descartes, Spinoza, Leibniz, Kant, and Hegel, as well as those of Schopenhauer, whose ideas he rejected.

Kronecker spent the summer semester of 1843 at the University of Bonn, and the fall semester at Breslau (now Wroclaw, Poland) because Kummer had been appointed professor there. He remained there for two semesters, returning to Berlin in the winter of 1844–1845 to take the doctorate. Kronecker took his oral examination consisting of questions not only in mathematics, but also in Greek history of legal philosophy. He was awarded the doctorate on 10 September 1845.

Dirichlet, his professor and examiner, was to remain one of Kronecker's closest friends, as was Kummer, his first mathematics teacher. In the meantime, in Berlin, Kronecker was also becoming better acquainted with Eisenstein and with Jacobi. During the same period Dirichlet introduced him to Alexander von Humboldt and to the composer Felix Mendelssohn, who was both Dirichlet's brother-in-law and the cousin of Kummer's wife.

Family business then called Kronecker from Berlin. In its interest he was required to spend a few years managing an estate near Liegnitz, as well as to dissolve the banking business of an uncle. In 1848 he married the latter's daughter, his cousin Fanny Prausnitzer; they

had six children. Having temporarily renounced an academic career, Kronecker continued to do mathematics as a recreation. He both carried on independent research and engaged in a lively mathematical correspondence with Kummer; he was not ambitious for fame, and was able to enjoy mathematics as a true amateur. By 1855, however, Kronecker's circumstances had changed enough to allow him to return to the academic life in Berlin as a financially independent private scholar.

In 1860 Kummer, seconded by Borchardt and Weierstrass, nominated Kronecker to the Berlin Academy, of which he became full member on 23 January 1861. Kronecker was increasingly active and influential in the affairs of the Academy, particularly in recruiting the most important German and foreign mathematicians for it. His influence outside Germany also increased. He was a member of many learned societies, among them the Paris Academy and the Royal Society of London. He established other contacts with foreign scientists in his numerous travels abroad and in extending to them the hospitality of his Berlin home. For this reason his advice was often solicited in regard to filling mathematical professorships both in Germany and elsewhere; his recommendations were probably as significant as those of his erstwhile friend Weierstrass.

The cause of the growing estrangement between Kronecker and Weierstrass was partly due to the very different temperaments of the two, and their professional and scientific differences. Since they had long maintained the same circle of friends, their friends, too, became involved on both levels. A characteristic incident occurred at the new year of 1884–1885, when H. A. Schwarz, who was both Weierstrass's student and Kummer's son-in-law, sent Kronecker a greeting that included the phrase: "He who does not honor the Smaller [Kronecker], is not worthy of the Greater [Weierstrass]." Kronecker read this allusion to physical size—he was a small man, and increasingly self-conscious with age—as a slur on his intellectual powers and broke with Schwarz completely.

Kronecker's mathematics lacked a systematic theoretical basis. Nevertheless, he was preeminent in uniting the separate mathematical disciplines. Moreover, in certain ways—his refusal to recognize an actual infinity, his insistence that a mathematical concept must be defined in a finite number of steps, and his opposition to the work of Cantor and Dedekind—his approach may be compared to that of intuitionists in the twentieth century. Kronecker's mathematics thus remains influential.

## 25.3 Exterior Algebra

The following discussion will concentrate on tensors of type  $(r, 0)$ . However, interchanging the roles of  $\mathcal{V}$  and  $\mathcal{V}^*$  makes all definitions, theorems, propositions, and conclusions valid for tensors of type  $(0, s)$  as well.

The set of all skew-symmetric tensors of type  $(p, 0)$  forms a subspace of  $\mathcal{T}^p(\mathcal{V})$ . This subspace is denoted by  $\Lambda^p(\mathcal{V})$ . It is not, however, an algebra unless we define a skew-symmetric product analogous to that for the symmetric case. First, we need the following:

**25.3.1. Definition.** An *antisymmetrizer* is a linear operator  $\mathbb{A} : \mathcal{T}_0^p(\mathcal{V}) \rightarrow \Lambda^p(\mathcal{V})$ ,

antisymmetrizer

given by

$$[\mathbb{A}(\mathbf{T})](\tau^1, \dots, \tau^p) = \frac{1}{p!} \sum_{\pi} \epsilon_{\pi} \mathbf{A}(\tau^{\pi(1)}, \dots, \tau^{\pi(p)}). \quad (25.9)$$

The sum is over all permutations of  $(1, 2, \dots, p)$ , and  $\epsilon_{\pi} = \epsilon_{\pi(1)\dots\pi(p)}$  is  $+1$  if the permutation is even and  $-1$  if the permutation is odd.  $\mathbb{A}(\mathbf{T})$  is sometimes denoted by  $\mathbf{T}_a$ .

**25.3.2. Example.** Let us write out Equation (25.9) for  $p = 3$ . The procedure is entirely analogous to Example 25.2.4:

$$\begin{aligned} \mathbf{T}_a(\tau^1, \tau^2, \tau^3) &= \frac{1}{6} [\epsilon_{123} \mathbf{A}(\tau^1, \tau^2, \tau^3) + \epsilon_{213} \mathbf{A}(\tau^2, \tau^1, \tau^3) + \epsilon_{132} \mathbf{A}(\tau^1, \tau^3, \tau^2) \\ &\quad + \epsilon_{312} \mathbf{A}(\tau^3, \tau^1, \tau^2) + \epsilon_{321} \mathbf{A}(\tau^3, \tau^2, \tau^1) + \epsilon_{231} \mathbf{A}(\tau^2, \tau^3, \tau^1)] \\ &= \frac{1}{6} [\mathbf{A}(\tau^1, \tau^2, \tau^3) - \mathbf{A}(\tau^2, \tau^1, \tau^3) - \mathbf{A}(\tau^1, \tau^3, \tau^2) \\ &\quad + \mathbf{A}(\tau^3, \tau^1, \tau^2) - \mathbf{A}(\tau^3, \tau^2, \tau^1) + \mathbf{A}(\tau^2, \tau^3, \tau^1)]. \end{aligned}$$

The reader may easily verify that all terms with a plus sign are obtained from (123) by an even number of interchanges of symbols, and those with a minus sign by an odd number. ■

We can now define the important product that turns the collection of all anti-symmetric tensors into an algebra.

exterior product  
defined

**25.3.3. Definition.** The exterior product (also called the wedge, Grassmann, alternating, or veck product) of two skew-symmetric tensors  $\mathbf{A} \in \Lambda^p(\mathcal{V})$  and  $\mathbf{B} \in \Lambda^q(\mathcal{V})$  is a skew-symmetric tensor belonging to  $\Lambda^{p+q}(\mathcal{V})$  and given by<sup>4</sup>

$$\mathbf{A} \wedge \mathbf{B} \equiv \frac{(r+s)!}{r!s!} \mathbb{A}(\mathbf{A} \otimes \mathbf{B}) = \frac{(r+s)!}{r!s!} (\mathbf{A} \otimes \mathbf{B})_a.$$

More explicitly,

$$\mathbf{A} \wedge \mathbf{B}(\tau^1, \dots, \tau^{r+s}) = \frac{1}{r!s!} \sum_{\pi} \epsilon_{\pi} \mathbf{A}(\tau^{\pi(1)}, \dots, \tau^{\pi(r)}) \mathbf{B}(\tau^{\pi(r+1)}, \dots, \tau^{\pi(r+s)}).$$

**25.3.4. Example.** Let us find the exterior product of vectors. First we find the exterior product of  $\mathbf{v}_1$  and  $\mathbf{v}_2$  both belonging to  $\mathcal{V} = \mathbb{T}_0^1(\mathcal{V})$ :

$$\begin{aligned} \mathbf{v}_1 \wedge \mathbf{v}_2(\tau^1, \tau^2) &= \sum_{\pi} \epsilon_{\pi(1)\pi(2)} (\mathbf{v}_1 \otimes \mathbf{v}_2)(\tau^{\pi(1)}, \tau^{\pi(2)}) \\ &= \epsilon_{12} (\mathbf{v}_1 \otimes \mathbf{v}_2)(\tau^1, \tau^2) + \epsilon_{21} (\mathbf{v}_1 \otimes \mathbf{v}_2)(\tau^2, \tau^1) \\ &= \mathbf{v}_1(\tau^1) \mathbf{v}_2(\tau^2) - \mathbf{v}_1(\tau^2) \mathbf{v}_2(\tau^1) \\ &= (\mathbf{v}_1 \otimes \mathbf{v}_2 - \mathbf{v}_2 \otimes \mathbf{v}_1)(\tau^1, \tau^2). \end{aligned}$$

<sup>4</sup>The reader should be warned that different authors may use different numerical factors in the definition of the exterior product.

Since this is true for arbitrary  $\tau^1$  and  $\tau^2$ , we have

$$v_1 \wedge v_2 = v_1 \otimes v_2 - v_2 \otimes v_1$$

In general,

$$v_1 \wedge v_2 \wedge \cdots \wedge v_r = \sum_{\pi} \epsilon_{\pi(1)\dots\pi(r)} v_{\pi(1)} \otimes v_{\pi(2)} \otimes \cdots \otimes v_{\pi(r)}.$$

In particular, this shows that the exterior product (of vectors) is associative. ■

If  $\{e_j\}_{j=1}^N$  is a basis with dual  $\{\epsilon^i\}_{i=1}^N$ , then the last result of Example 25.3.4 yields

$$\epsilon^{i_1} \wedge \epsilon^{i_2} \wedge \cdots \wedge \epsilon^{i_N}(e_{j_1}, e_{j_2}, \dots, e_{j_N}) = \sum_{\pi} \epsilon_{\pi} \delta_{\pi(j_1)}^{i_1} \delta_{\pi(j_2)}^{i_2} \cdots \delta_{\pi(j_N)}^{i_N}. \tag{25.10}$$

The following theorem contains the properties of the exterior product (for a proof, see [Abra 88, p. 394]):

**25.3.5. Theorem.** *The exterior product is associative and distributive with respect to the addition of tensors. Furthermore, it satisfies the following anticommutativity property:*

$$A \wedge B = (-1)^{pq} B \wedge A$$

whenever  $A \in \Lambda^p(\mathcal{V})$  and  $B \in \Lambda^q(\mathcal{V})$ . In particular,  $v_1 \wedge v_2 = -v_2 \wedge v_1$  for  $v_1, v_2 \in \mathcal{V}$ .

The wedge product of linear functionals of a vector space is particularly important in the analysis of tensors, as we shall see in the next chapter.

*p*-forms defined **25.3.6. Definition.** *The elements of  $\Lambda^p(\mathcal{V}^*)$  are called *p*-forms.*

A linear transformation  $T : \mathcal{V} \rightarrow \mathcal{W}$  induces a transformation<sup>6</sup>  $T^* : \Lambda^p(\mathcal{W}^*) \rightarrow \Lambda^p(\mathcal{V}^*)$  defined by

$$(T^*\rho)(v_1, \dots, v_p) \equiv \rho(Tv_1, \dots, Tv_p), \quad \rho \in \Lambda^p(\mathcal{W}^*), v_i \in \mathcal{V}. \tag{25.11}$$

$T^*\rho$  is called the **pullback** of  $\rho$  by  $T$ . The most important properties of pullback maps are given in the following:

**25.3.7. Proposition.** *Let  $T : \mathcal{V} \rightarrow \mathcal{U}$  and  $S : \mathcal{U} \rightarrow \mathcal{W}$ . Then*

1.  $T^* : \Lambda^p(\mathcal{U}^*) \rightarrow \Lambda^p(\mathcal{V}^*)$  is linear. (So is  $S^*$ , of course.)
2.  $(S \circ T)^* = T^* \circ S^*$ .

---

<sup>6</sup>Note that  $T^*$  is the extension of the pullback operator introduced at the end of Chapter 1.

3. If  $\mathbf{T}$  is the identity map, so is  $\mathbf{T}^*$ .
4. If  $\mathbf{T}$  is an isomorphism, so is  $\mathbf{T}^*$  and  $(\mathbf{T}^*)^{-1} = (\mathbf{T}^{-1})^*$ .
5. If  $\rho \in \Lambda^p(\mathcal{W}^*)$  and  $\sigma \in \Lambda^q(\mathcal{W}^*)$ , then  $\mathbf{T}^*(\rho \wedge \sigma) = \mathbf{T}^*\rho \wedge \mathbf{T}^*\sigma$ .

*Proof.* The proof follows directly from definitions and is left as an exercise for the reader.  $\square$

The components of  $\mathbf{A} \in \Lambda^p(\mathcal{V})$  are given by  $A^{i_1 \dots i_p}$ , where  $i_1 < i_2 < \dots < i_p$ . All other components are related to these by skew-symmetry. The number of independent components, which is the dimension of  $\Lambda^p(\mathcal{V})$ , is equal to the number of ways  $p$  numbers can be chosen from among  $N$  distinct numbers in such a way that no two of them are equal. This is simply the combination of  $N$  objects taken  $p$  at a time. Thus, we have

$$\dim \Lambda^p(\mathcal{V}) = \binom{N}{p} = \frac{N!}{p!(N-p)!}. \quad (25.12)$$

In particular,  $\dim \Lambda^N(\mathcal{V}) = 1$ . This should come as no surprise, because from a basis  $\{\mathbf{e}_i\}_{i=1}^N$  of  $\mathcal{V}$ , we can form a basis for  $\Lambda^p(\mathcal{V})$  by constructing all products of the form  $\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \dots \wedge \mathbf{e}_{i_p}$  [of which there are  $\binom{N}{p}$ ]. However, when  $p = N$ , there is, within a multiplicative constant, only one such product and that is  $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_N$ .

An elegant way of determining the linear independence of vectors using the formalism developed so far is given in the following proposition.

**25.3.8. Proposition.** *A set of vectors,  $\mathbf{v}_1, \dots, \mathbf{v}_p$ , is linearly independent if and only if  $\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_p \neq 0$ .*

*Proof.* If  $\{\mathbf{v}_i\}_{i=1}^p$  are independent, then they span a  $p$ -dimensional subspace  $\mathcal{M}$  of  $\mathcal{V}$ . Considering  $\mathcal{M}$  as a vector space in its own right, we have  $\dim \Lambda^p(\mathcal{M}) = 1$ . A basis for  $\Lambda^p(\mathcal{M})$  is simply  $\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_p$ , which cannot be zero.

Conversely, suppose that  $\alpha_1 \mathbf{v}_1 + \dots + \alpha_p \mathbf{v}_p = 0$ . Then taking the exterior product of the LHS with  $\mathbf{v}_2 \wedge \mathbf{v}_3 \wedge \dots \wedge \mathbf{v}_p$  makes all terms vanish (because each will have two factors of a vector in the wedge product) except the first one. Thus, we have  $\alpha_1 \mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_p = 0$ . The fact that the wedge product is not zero forces  $\alpha_1$  to be zero. Similarly, multiplying by  $\mathbf{v}_1 \wedge \mathbf{v}_3 \wedge \dots \wedge \mathbf{v}_p$  shows that  $\alpha_2 = 0$ , and so on.  $\square$

**25.3.9. Example.** Let  $\{\mathbf{e}_i\}_{i=1}^N$  be a basis for  $\mathcal{V}$ . Let  $\mathbf{v}_1 = \mathbf{e}_1 + 2\mathbf{e}_2 - \mathbf{e}_3$ ,  $\mathbf{v}_2 = 3\mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3$ ,  $\mathbf{v}_3 = -\mathbf{e}_1 - 3\mathbf{e}_2 + 2\mathbf{e}_3$ .

Take the wedge product of the first two  $\mathbf{v}$ 's:

$$\mathbf{v}_1 \wedge \mathbf{v}_2 = (\mathbf{e}_1 + 2\mathbf{e}_2 - \mathbf{e}_3) \wedge (3\mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3) = -5\mathbf{e}_1 \wedge \mathbf{e}_2 + 5\mathbf{e}_1 \wedge \mathbf{e}_3 + 5\mathbf{e}_2 \wedge \mathbf{e}_3.$$

All the wedge products that have repeated factors vanish. Now we multiply by  $\mathbf{v}_3$ :

$$\begin{aligned} \mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \mathbf{v}_3 &= -5\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge (-\mathbf{e}_1 - 3\mathbf{e}_2 + 2\mathbf{e}_3) + 5\mathbf{e}_1 \wedge \mathbf{e}_3 \wedge (-\mathbf{e}_1 - 3\mathbf{e}_2 + 2\mathbf{e}_3) \\ &\quad + 5\mathbf{e}_2 \wedge \mathbf{e}_3 \wedge (-\mathbf{e}_1 - 3\mathbf{e}_2 + 2\mathbf{e}_3) \\ &= -10\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3 - 15\mathbf{e}_1 \wedge \mathbf{e}_3 \wedge \mathbf{e}_2 - 5\mathbf{e}_2 \wedge \mathbf{e}_3 \wedge \mathbf{e}_1 = 0. \end{aligned}$$



We conclude that the three vectors are linearly dependent.  $\blacksquare$

As an application of Proposition 25.3.8, let us prove the following.

**Cartan's lemma** **25.3.10. Theorem.** (Cartan's lemma) *Suppose that  $\{\mathbf{e}_i\}_{i=1}^p$ ,  $p \leq \dim \mathcal{V}$ , form a linearly independent set of vectors in  $\mathcal{V}$  and that  $\{\mathbf{v}_i\}_{i=1}^p$  are also vectors in  $\mathcal{V}$  such that  $\sum_{i=1}^p \mathbf{e}_i \wedge \mathbf{v}_i = 0$ . Then all  $\mathbf{v}_i$  are linear combinations of only the set  $\{\mathbf{e}_i\}_{i=1}^p$ . Furthermore, if  $\mathbf{v}_i = \sum_{j=1}^p A_{ij} \mathbf{e}_j$ , then  $A_{ij} = A_{ji}$ .*

*Proof.* Multiplying both sides of  $\sum_{i=1}^p \mathbf{e}_i \wedge \mathbf{v}_i = 0$  by  $\mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_p$  gives

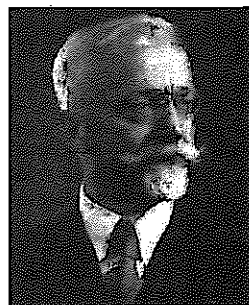
$$-\mathbf{v}_1 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_p = 0.$$

By Proposition 25.3.8,  $\mathbf{v}_1$  and the  $\mathbf{e}_i$  are linearly dependent. Similarly, by multiplying  $\sum_{i=1}^p \mathbf{e}_i \wedge \mathbf{v}_i = 0$  by the wedge product with  $\mathbf{e}_k$  missing, we show that  $\mathbf{v}_k$  and the  $\mathbf{e}_i$  are linearly dependent. Thus,  $\mathbf{v}_k = \sum_{i=1}^p A_{ki} \mathbf{e}_i$ , for all  $k$ . Furthermore, we have

$$0 = \sum_{k=1}^p \mathbf{e}_k \wedge \mathbf{v}_k = \sum_{k=1}^p \sum_{i=1}^p \mathbf{e}_k \wedge (A_{ki} \mathbf{e}_i) = \sum_{k < i} (A_{ki} - A_{ik}) \mathbf{e}_k \wedge \mathbf{e}_i,$$

where the last sum is over both  $k$  and  $i$  with  $k < i$ . Clearly,  $\{\mathbf{e}_k \wedge \mathbf{e}_i\}$  with  $k < i$  are linearly independent (show this!). Therefore, their coefficients must vanish.  $\square$

**Elie Joseph Cartan** (1869–1951), born in Dolomieu (near Chambéry), Savoie, Rhône-Alpes, France, became a student at the Ecole Normale in 1888 and obtained his doctorate in 1894. He lectured at Montpellier (1894–1896), Lyon (1896–1903), Nancy (1903–1909), and Paris (1909–1940). He had four children, one of whom, Henri Cartan, was to produce brilliant work in mathematics. Two others died tragically. Jean, a composer, died at the age of 25, while Louis, a physicist, was arrested by the Germans in 1942 and executed after 15 months in captivity.



Cartan added greatly to the theory of continuous groups, which had been initiated by Lie. His thesis (1894) contains a major contribution to Lie algebras wherein he completed the classification of the semisimple algebras that Killing had essentially found. He then turned to the theory of associative algebras and investigated the structure for these algebras over the real and complex fields. Wedderburn would complete Cartan's work in this area.

He then turned to representations of semisimple Lie groups. His work is a striking synthesis of Lie theory, classical geometry, differential geometry, and topology, which was to be found in all Cartan's work. He also applied Grassmann algebra to the theory of exterior differential forms.

By 1904 Cartan was turning to papers on differential equations, and from 1916 on he published mainly on differential geometry. Klein's *Erlanger Program* was seen to be

inadequate as a general description of geometry by Weyl and Veblen, and Cartan was to play a major role. He examined a space acted on by an arbitrary Lie group of transformations, developing a theory of moving frames that generalizes the kinematical theory of Darboux.

Cartan further contributed to geometry with his theory of symmetric spaces, which have their origins in papers he wrote in 1926. It develops ideas first studied by Clifford and Cayley and used topological methods developed by Weyl in 1925. This work was completed by 1932.

Cartan then went on to examine problems on a topic first studied by Poincaré. By this stage his son, Henri Cartan, was making major contributions to mathematics, and Elie Cartan was able to build on theorems proved by his son.

Cartan also published work on relativity and the theory of spinors. He is certainly one of the most important mathematicians of the first half of the twentieth century.

### 25.3.1 The Determinant

One of the most beautiful applications of exterior algebra is in the theory of determinants. We have already considered determinants in detail in Chapter 3, where we noted how messy it was to prove some of the theorems. With the machinery of exterior algebra at our disposal, we will see how elegant this theory becomes and how trivial some of the proofs will turn out to be.

First, let us recall that  $\dim \Lambda^N(\mathcal{V}) = 1$  when  $N = \dim \mathcal{V}$ . This means that if  $\{\mathbf{e}_i\}_{i=1}^N$  is a basis of  $\mathcal{V}$ , then  $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N$  is the only vector in the corresponding basis of  $\Lambda^N(\mathcal{V})$ . On the other hand, if  $\{\mathbf{v}_i\}_{i=1}^N$  is any set of  $N$  vectors, the product  $\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \cdots \wedge \mathbf{v}_N$  is either zero (if the  $\mathbf{v}_i$ 's are linearly dependent) or a nonzero product belonging to  $\Lambda^N(\mathcal{V})$ . Since  $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N$  is a basis of  $\Lambda^N(\mathcal{V})$ , we conclude that for any set of  $N$  vectors,  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N \in \mathcal{V}$ , the product  $\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \cdots \wedge \mathbf{v}_N$  is a multiple (possibly zero) of  $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N$ .

Now let  $\mathbf{A}$  be a linear operator on  $\mathcal{V}$ . The set of vectors  $\mathbf{A}\mathbf{e}_1, \mathbf{A}\mathbf{e}_2, \dots, \mathbf{A}\mathbf{e}_N$  all belong to  $\mathcal{V}$ . By the above remarks  $(\mathbf{A}\mathbf{e}_1) \wedge \cdots \wedge (\mathbf{A}\mathbf{e}_N)$  is proportional to  $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N$ . We now show that the proportionality constant is simply  $\det \mathbf{A}$ .

determinant in terms  
of wedge products

**25.3.11. Theorem.** *Let  $\mathbf{A} : \mathcal{V} \rightarrow \mathcal{V}$  be linear. Let  $\{\mathbf{e}_i\}_{i=1}^N$  be a basis for  $\mathcal{V}$ . Then*

$$(\mathbf{A}\mathbf{e}_1) \wedge \cdots \wedge (\mathbf{A}\mathbf{e}_N) = (\det \mathbf{A}) \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_N. \quad (25.13)$$

*Proof.* Let  $\mathbf{A}\mathbf{e}_r = A_r^i \mathbf{e}_i$ , for  $r = 1, 2, \dots, N$ . Then

$$\begin{aligned} (\mathbf{A}\mathbf{e}_1) \wedge \cdots \wedge (\mathbf{A}\mathbf{e}_N) &= (A_1^{i_1} \mathbf{e}_{i_1}) \wedge (A_2^{i_2} \mathbf{e}_{i_2}) \wedge \cdots \wedge (A_N^{i_N} \mathbf{e}_{i_N}) \\ &= A_1^{i_1} A_2^{i_2} \cdots A_N^{i_N} \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_N}. \end{aligned}$$

It is straightforward to show that

$$\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_N} = \epsilon_{i_1 i_2 \dots i_N} \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N, \quad (25.14)$$

so that  $(\mathbf{Ae}_1) \wedge \cdots \wedge (\mathbf{Ae}_N) = (A_1^{i_1} A_2^{i_2} \cdots A_N^{i_N} \epsilon_{i_1 i_2 \dots i_N}) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N$ . The expression in parentheses on the RHS is simply the determinant as defined in Chapter 3 (recall the summation convention).  $\square$

Levi-Civita tensor and determinants

**25.3.12. Example.** The symbol  $\epsilon_{i_1 i_2 \dots i_N}$ , called the **Levi-Civita tensor**, can be defined by Equation (25.14). Consider the linear operator  $\mathbf{E}$  whose action on a basis  $\{\mathbf{e}_i\}_{i=1}^N$  is to permute the vectors so that  $\mathbf{Ee}_k = \mathbf{e}_{i_k}$ . Then on the one hand,

$$(\mathbf{Ee}_1) \wedge \cdots \wedge (\mathbf{Ee}_N) = (\det \mathbf{E}) \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_N,$$

and on the other hand, by the definition of the action of  $\mathbf{E}$ ,

$$(\mathbf{Ee}_1) \wedge \cdots \wedge (\mathbf{Ee}_N) = \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_N}.$$

Comparison of these two equations with (25.14) gives  $\det \mathbf{E} = \epsilon_{i_1 i_2 \dots i_N}$ .  $\blacksquare$

Since the determinant is basis-independent, the result of the previous example can be summarized as follows:

**25.3.13. Box.** *The Levi-Civita tensor  $\epsilon_{i_1 i_2 \dots i_N}$  takes the same value in all coordinate systems.*

Let us now look at the determinant of the product of two operators  $\mathbf{A}$  and  $\mathbf{B}$ . By definition,  $(\mathbf{A}\mathbf{B}\mathbf{e}_1) \wedge \cdots \wedge (\mathbf{A}\mathbf{B}\mathbf{e}_N) = [\det(\mathbf{A}\mathbf{B})] \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_N$ . However, we also have

$$\begin{aligned} (\mathbf{A}\mathbf{B}\mathbf{e}_1) \wedge \cdots \wedge (\mathbf{A}\mathbf{B}\mathbf{e}_N) &= [\mathbf{A}(\mathbf{B}\mathbf{e}_1)] \wedge \cdots \wedge [\mathbf{A}(\mathbf{B}\mathbf{e}_N)] \\ &= (\det \mathbf{A})(\mathbf{B}\mathbf{e}_1) \wedge \cdots \wedge (\mathbf{B}\mathbf{e}_N) \\ &= (\det \mathbf{A})(\det \mathbf{B}) \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_N. \end{aligned}$$

It follows that  $\det \mathbf{A}\mathbf{B} = (\det \mathbf{A})(\det \mathbf{B})$ . Here the power and elegance of exterior algebra can be truly appreciated—to prove the same result in Chapter 3, we had to go through a maze of index shuffling and reshuffling.

### 25.3.2 Orientation

The reader is no doubt familiar with the right-handed and left-handed coordinate systems in  $\mathbb{R}^3$ . In this section, we generalize the idea to arbitrary vector spaces.

oriented basis defined

**25.3.14. Definition.** An **oriented basis** of an  $N$ -dimensional vector space is an ordered collection of  $N$  linearly independent vectors.

If  $\{\mathbf{v}_i\}_{i=1}^N$  is one oriented basis and  $\{\mathbf{u}_i\}_{i=1}^N$  is a second one, then

$$\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \cdots \wedge \mathbf{u}_N = (\det \mathbf{R}) \mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \cdots \wedge \mathbf{v}_N,$$

where  $\mathbf{R}$  is the transformation matrix and  $\det \mathbf{R}$  is a nonzero number ( $\mathbf{R}$  is invertible), which can be positive or negative. Accordingly, we have the following definition.

oriented vector spaces defined

**25.3.15. Definition.** An *orientation* is the collection of all oriented bases related by a transformation matrix having a positive determinant. A vector space for which an orientation is specified is called an **oriented vector space**.

Clearly, there are only two orientations in any vector space. Each oriented basis is positively related to any oriented basis belonging to the same orientation and negatively related to any oriented basis belonging to the other orientation. For example, in  $\mathbb{R}^3$ , the bases  $\{e_x, e_y, e_z\}$  and  $\{e_y, e_x, e_z\}$  belong to different orientations because

$$e_x \wedge e_y \wedge e_z = -e_y \wedge e_x \wedge e_z.$$

The first basis is (by convention) called a **right-handed** coordinate system, and the second is called a **left-handed** coordinate system. Any other basis is either right-handed or left-handed. There is no third alternative!

volume element of a vector space

**25.3.16. Definition.** Let  $\mathcal{V}$  be a vector space. Let  $\mathcal{V}^*$  have the oriented basis  $\{\epsilon^i\}_{i=1}^N$ . The oriented **volume element**  $\mu \in \Lambda^N(\mathcal{V}^*)$  of  $\mathcal{V}$  is defined as

$$\mu \equiv \epsilon^1 \wedge \epsilon^2 \wedge \dots \wedge \epsilon^N.$$

positive orientation

Note that if  $\{e_i\}$  is ordered as  $\{e^j\}$ , then  $\mu(e_1, e_2, \dots, e_N) = +1/N!$ , and we say that  $\{e_i\}$  is *positively oriented* with respect to  $\mu$ . In general,  $\{v_i\}$  is positively oriented with respect to  $\mu$  if  $\mu(v_1, v_2, \dots, v_N) > 0$ .

The volume element of  $\mathcal{V}$  is defined in terms of a basis for  $\mathcal{V}^*$ . The reason for this will become apparent later, when we shall see that  $dx, dy$ , and  $dz$  form a basis for  $(\mathbb{R}^3)^*$ , and  $dx dy dz \equiv dx \wedge dy \wedge dz$ .

### 25.3.3 Symplectic Vector Spaces

Mechanics was a great contributor to the development of tensor analysis. It provided examples of manifolds that went beyond mere subspaces of  $\mathbb{R}^n$ . The phase space of Hamiltonian mechanics is a paradigm of manifolds that are not “hypersurfaces” of some Euclidean space. We shall have more to say about such manifolds in Chapter 26. Here, we shall be content with the algebraic structure underlying classical mechanics.

symplectic form, symplectic vector space, and symplectic transformation

**25.3.17. Definition.** A 2-form  $\omega \in \Lambda^2(\mathcal{V}^*)$  is called *nondegenerate* if  $\omega(v_1, v_2) = 0$  for all  $v_1 \in \mathcal{V}$  implies  $v_2 = 0$ . A *symplectic form* on  $\mathcal{V}$  is a nondegenerate 2-form  $\omega \in \Lambda^2(\mathcal{V}^*)$ . The pair  $(\mathcal{V}, \omega)$  is called a **symplectic vector space**. If  $(\mathcal{V}, \omega)$  and  $(\mathcal{W}, \rho)$  are symplectic vector spaces, a linear transformation  $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$  is called a **symplectic transformation** or a **symplectic map** if  $\mathbf{T}^* \rho = \omega$ .

Any 2-form (degenerate or nondegenerate) leads to other quantities that are also of interest. For instance, given any basis  $\{v_i\}$  in  $\mathcal{V}$ , one defines the **matrix** of

The flat  $\flat$  and sharp  $\sharp$  maps defined

the 2-form  $\omega \in \Lambda^2(\mathcal{V}^*)$  by  $\omega_{ij} \equiv \omega(\mathbf{v}_i, \mathbf{v}_j)$ . Similarly, one can define the useful linear map  $\omega^\flat : \mathcal{V} \rightarrow \mathcal{V}^*$  by

$$\underbrace{[\omega^\flat(\mathbf{v})]}_{\in \mathcal{V}^*} \mathbf{v}' \equiv \omega(\mathbf{v}, \mathbf{v}'), \quad (25.15)$$

rank of a symplectic form

The rank of  $\omega^\flat$  is called the **rank** of  $\omega$ . The reader may check that

**25.3.18. Box.** A 2-form  $\omega$  is nondegenerate if and only if the determinant of  $(\omega_{ij})$  is nonzero, if and only if  $\omega^\flat$  is an isomorphism, in which case the inverse of  $\omega^\flat$  is denoted by  $\omega^\sharp$ .

canonical basis of a symplectic vector space

**25.3.19. Proposition.** Let  $\mathcal{V}$  be an  $N$ -dimensional vector space and  $\omega \in \Lambda^2(\mathcal{V}^*)$ . If the rank of  $\omega$  is  $r$ , then  $r = 2n$  for some integer  $n$  and there exists a basis  $\{\mathbf{e}_i\}$  of  $\mathcal{V}$ , called a **canonical basis** of  $\mathcal{V}$ , and a dual basis  $\{\mathbf{e}^j\}$ , such that  $\omega = \sum_{j=1}^n \mathbf{e}^j \wedge \mathbf{e}^{j+n}$ , or, equivalently, the  $N \times N$  matrix of  $\omega$  is given by

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where  $\mathbf{1}$  is the  $n \times n$  identity matrix and  $\mathbf{0}$  is the  $(N - 2n) \times (N - 2n)$  zero matrix.

*Proof.* Since  $\omega \neq 0$ , there exist a pair of vectors  $\mathbf{e}_1, \mathbf{e}'_1 \in \mathcal{V}$  such that  $\omega(\mathbf{e}_1, \mathbf{e}'_1) \neq 0$ . Dividing  $\mathbf{e}_1$  by a constant, we can assume  $\omega(\mathbf{e}_1, \mathbf{e}'_1) = 1$ . Because of its antisymmetry, the matrix of  $\omega$  in the plane  $\mathcal{P}_1$  spanned by  $\mathbf{e}_1$  and  $\mathbf{e}'_1$  is  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ . Let  $\mathcal{V}_1$  be the  $\omega$ -orthogonal complement of  $\mathcal{P}_1$ , i.e.,

$$\mathcal{V}_1 = \{\mathbf{v} \in \mathcal{V} \mid \omega(\mathbf{v}, \mathbf{v}_1) = 0 \quad \forall \mathbf{v}_1 \in \mathcal{P}_1\}.$$

Then the reader may check that  $\mathcal{P}_1 \cap \mathcal{V}_1 = \mathbf{0}$ . Moreover,  $\mathcal{V} = \mathcal{P}_1 + \mathcal{V}_1$  because

$$\mathbf{v} = \underbrace{\omega(\mathbf{v}, \mathbf{e}'_1)\mathbf{e}_1 - \omega(\mathbf{v}, \mathbf{e}_1)\mathbf{e}'_1}_{\in \mathcal{P}_1} + \underbrace{\mathbf{v} - \omega(\mathbf{v}, \mathbf{e}'_1)\mathbf{e}_1 + \omega(\mathbf{v}, \mathbf{e}_1)\mathbf{e}'_1}_{\in \mathcal{V}_1 \text{ (Reader, verify!)}}$$

for any  $\mathbf{v} \in \mathcal{V}$ . Thus,  $\mathcal{V} = \mathcal{P}_1 \oplus \mathcal{V}_1$ . If  $\omega$  is zero on all pairs of vectors in  $\mathcal{V}_1$ , then we are done, and the rank of  $\omega$  is 2; otherwise, let  $\mathbf{e}_2, \mathbf{e}'_2 \in \mathcal{V}_1$  be such that  $\omega(\mathbf{e}_2, \mathbf{e}'_2) \neq 0$ . Proceeding as above, we obtain

$$\mathcal{V}_1 = \mathcal{P}_2 \oplus \mathcal{V}_2 \Rightarrow \mathcal{V} = \mathcal{P}_1 \oplus \mathcal{P}_2 \oplus \mathcal{V}_2,$$

where  $\mathcal{P}_2$  is the plane spanned by  $\mathbf{e}_2$ , and  $\mathbf{e}'_2$  and  $\mathcal{V}_2$  its  $\omega$ -orthogonal complement in  $\mathcal{V}_1$ . Continuing this process yields

$$\mathcal{V} = \mathcal{P}_1 \oplus \mathcal{P}_2 \oplus \dots \oplus \mathcal{P}_n \oplus \mathcal{V}_n,$$

where  $\mathcal{V}_n$  is the subspace of  $\mathcal{V}$  on which  $\omega$  is zero. This shows that the rank of  $\omega$  is  $2n$ . By reordering the basis vectors such that  $\mathbf{e}'_k \equiv \mathbf{e}_{n+k}$ , we construct a new basis  $\{\mathbf{e}_i\}_{i=1}^N$  in which  $\omega$  has the desired matrix.

To conclude the proposition, it is sufficient to show that  $\sum_{j=1}^n \mathbf{e}^j \wedge \mathbf{e}^{j+n}$ , in which  $\{\mathbf{e}^j\}_{j=1}^N$  is dual to  $\{\mathbf{e}_i\}_{i=1}^N$ , has the same matrix as  $\omega$ . This is left as an exercise for the reader.  $\square$

We note that in the canonical basis,

$$\omega_{ij} = \begin{cases} 0 & \text{if } i, j \leq n, \\ \delta_{ik} & \text{if } j = n + k, k \leq n, \\ 0 & \text{if } i \geq 2n \text{ or } j \geq 2n. \end{cases}$$

If we write  $\mathbf{v} \in \mathcal{V}$  as  $\mathbf{v} = \sum_{i=1}^n (x_i \mathbf{e}_i + y_i \mathbf{e}_{n+i}) + \sum_{i=1}^{N-2n} z_i \mathbf{e}_{2n+i}$  in the canonical basis of  $\mathcal{V}$ , with a corresponding expansion for  $\mathbf{v}'$ , then the reader may verify that

$$\omega(\mathbf{v}, \mathbf{v}') = \sum_{i=1}^n (x_i y'_i - x'_i y_i).$$

The following proposition gives a useful criterion for nondegeneracy of  $\omega$ :

**25.3.20. Proposition.** *Let  $\omega$  be a 2-form in the finite-dimensional vector space  $\mathcal{V}$ . Then  $\omega$  is nondegenerate iff  $\mathcal{V}$  has even dimension, say  $2n$ , and  $\omega^n \equiv \omega \wedge \dots \wedge \omega$  is a volume element of  $\mathcal{V}$ .*

*Proof.* Suppose  $\omega$  is nondegenerate. Then,  $\omega^b$  is an isomorphism. Therefore, the rank of  $\omega$ , an even number by Proposition 25.3.19, must equal  $\dim \mathcal{V}^* = \dim \mathcal{V}$ . Moreover, by taking successive powers of  $\omega$  and using mathematical induction, one can show that  $\omega^n$  is proportional to  $\epsilon^1 \wedge \dots \wedge \epsilon^{2n}$ .

Conversely, if  $\omega^n \propto \epsilon^1 \wedge \dots \wedge \epsilon^{2n}$  is a volume element, then by Proposition 25.3.8, the  $\{\epsilon^j\}$  are linearly independent. Furthermore,  $\dim \mathcal{V}^*$  must equal the number of linearly independent factors in the wedge product of a volume element. Thus,  $\dim \mathcal{V}^* = 2n$ . But  $2n$  is also the rank of  $\omega$ . It follows that  $\omega^b$  is onto. Since  $\mathcal{V}$  is finite-dimensional, the dimension theorem implies that  $\omega^b$  is an isomorphism.  $\square$

**25.3.21. Example.** Let  $\mathcal{V}$  be a vector space and  $\mathcal{V}^*$  its dual. The direct sum  $\mathcal{V} \oplus \mathcal{V}^*$  can be turned into a symplectic vector space if we define  $\omega \in \Lambda^2(\mathcal{V} \oplus \mathcal{V}^*)$  by

$$\omega(\mathbf{v} + \varphi, \mathbf{v}' + \varphi') \equiv \varphi'(\mathbf{v}) - \varphi(\mathbf{v}'),$$

where  $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$  and  $\varphi, \varphi' \in \mathcal{V}^*$ . The reader may verify that  $(\mathcal{V} \oplus \mathcal{V}^*, \omega)$  is a symplectic vector space. This construction of symplectic vector spaces is closely related to Hamiltonian dynamics, to which we shall return in Chapter 26.  $\blacksquare$

Suppose  $(\mathcal{V}, \omega)$  and  $(\mathcal{W}, \rho)$  are  $2n$ -dimensional symplectic vector spaces. Then, by Proposition 25.3.7, any symplectic map  $\mathbf{T} : (\mathcal{V}, \omega) \rightarrow (\mathcal{W}, \rho)$  is volume-preserving, i.e.,  $(\mathbf{T}^* \rho)^n$  is a volume element of  $\mathcal{V}$ . It follows that the rank of  $\mathbf{T}^*$  is  $2n$ , and by Proposition 1.3.18, so is the rank of  $\mathbf{T}$ . The dimension theorem now implies that  $\mathbf{T}$  is an isomorphism. Symplectic transformations on a single vector space have an interesting property:

**25.3.22. Proposition.** *Let  $(\mathcal{V}, \omega)$  be a symplectic vector space. Then the set of symplectic mappings  $\mathbf{T} : (\mathcal{V}, \omega) \rightarrow (\mathcal{V}, \omega)$  forms a group under composition, called the **symplectic group** and denoted by  $Sp(\mathcal{V}, \omega)$ .*

*Proof.* Clearly,  $Sp(\mathcal{V}, \omega)$  is a subset of  $GL(\mathcal{V})$ . One need only show that the inverse of a symplectic transformation is also such a transformation and that the product of two symplectic transformations is a symplectic transformation. The details are left for the reader.  $\square$

A matrix is called symplectic if it is the representation of a symplectic transformation in a canonical basis of the underlying symplectic vector space. The reader may check that the condition for a matrix  $A$  to be symplectic is  $A^t J A = J$ , where  $J$  is the representation of  $\omega$  in the canonical basis:

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

where  $1$  and  $0$  are the  $n \times n$  identity and zero matrices, respectively.

## 25.4 Inner Product Revisited

The inner product was defined in Chapter 1 in terms of a metric function that took two vectors as input and manufactured a number. We now know what kind of machine this is in the language of tensors.

**25.4.1. Definition.** *A symmetric bilinear form  $\mathbf{g}$  on  $\mathcal{V}$  is a symmetric tensor of type  $(0, 2)$ .*

If  $\{\mathbf{e}_j\}_{j=1}^N$  is a basis of  $\mathcal{V}$  and  $\{\epsilon^i\}_{i=1}^N$  is its dual basis, then  $\mathbf{g} = g_{ij} \epsilon^i \epsilon^j$  (recall Einstein's summation convention), because  $\epsilon^i \epsilon^j = \epsilon^i \otimes \epsilon^j + \epsilon^j \otimes \epsilon^i$  form a basis of  $\mathcal{S}_2(\mathcal{V})$ . For any vector  $\mathbf{v} \in \mathcal{V}$ , we can write

$$\mathbf{g}(\mathbf{v}) = g_{ij} \epsilon^i \epsilon^j (\mathbf{v}) = g_{ij} \epsilon^i \epsilon^j (v^k \mathbf{e}_k) = v^k g_{ij} \epsilon^i \underbrace{\epsilon^j(\mathbf{e}_k)}_{=\delta_k^j} = g_{ij} v^j \epsilon^i. \tag{25.16}$$

Thus,  $\mathbf{g}(\mathbf{v}) \in \mathcal{V}^*$ . This shows that  $\mathbf{g}$  can be thought of as a mapping,  $\mathbf{g} : \mathcal{V} \rightarrow \mathcal{V}^*$ , given by Equation (25.16). For this equation to make sense, it should not matter which factor in the symmetric product  $\mathbf{v}$  contracts with. But this is a trivial

consequence of the symmetries  $g_{ij} = g_{ji}$  and  $\epsilon^i \epsilon^j = \epsilon^j \epsilon^i$ . The components  $g_{ij} v^j$  of  $\mathbf{g}(\mathbf{v})$  in the basis  $\{\epsilon^i\}_{i=1}^N$  of  $\mathcal{V}^*$  are denoted by  $v_i$ , so

$$\mathbf{g}(\mathbf{v}) = v_i \epsilon^i, \quad \text{where } v_i = g_{ij} v^j. \tag{25.17}$$

We have thus *lowered* the index of  $v^j$  by the use of the symmetric bilinear form  $\mathbf{g}$ . In applications  $v_i$  is uniquely defined; furthermore, there is a one-to-one correspondence between  $v_i$  and  $v^i$ . This can happen if and only if the mapping  $\mathbf{g} : \mathcal{V} \rightarrow \mathcal{V}^*$  is invertible, in which case there must exist a unique  $\mathbf{g}^{-1} : \mathcal{V}^* \rightarrow \mathcal{V}$ , or  $\mathbf{g}^{-1} \in \mathcal{S}_2(\mathcal{V}^*) = \mathcal{S}^2(\mathcal{V})$ , such that

$$\begin{aligned} v^j \mathbf{e}_j &= \mathbf{v} = \mathbf{g}^{-1} \mathbf{g}(\mathbf{v}) = \mathbf{g}^{-1}(v_i \epsilon^i) = v_i \mathbf{g}^{-1}(\epsilon^i) \\ &= v_i [(g^{-1})^{jk} \mathbf{e}_j \mathbf{e}_k](\epsilon^i) = v_i (g^{-1})^{jk} \mathbf{e}_j \underbrace{\mathbf{e}_k(\epsilon^i)}_{=\delta_k^i} = v_i (g^{-1})^{ji} \mathbf{e}_j. \end{aligned}$$

Comparison of the LHS and the RHS yields  $v^j = v_i (g^{-1})^{ji}$ . It is customary to omit the  $-1$  and simply write

$$v^j = g^{ji} v_i, \tag{25.18}$$

where it is understood that  $\mathbf{g}$  with upper indices is the inverse of  $\mathbf{g}$  (with lower indices).

nondegenerate  
bilinear forms and  
inner products

**25.4.2. Definition.** *An invertible bilinear form is called **nondegenerate**. A symmetric bilinear form that is nondegenerate is called an **inner product**.*

raising and lowering  
indices

We therefore see that the presence of a nondegenerate symmetric bilinear form (or an inner product) naturally connects the vectors in  $\mathcal{V}$  and  $\mathcal{V}^*$  in a unique way. Going from a vector in  $\mathcal{V}$  to its unique image in  $\mathcal{V}^*$  is done by simply **lowering the index** using Equation (25.17), and going the other way involves using Equation (25.18) to **raise the index**. This process can be generalized to all tensors. For instance, although in general, there is no connection among  $\mathcal{T}_0^2(\mathcal{V})$ ,  $\mathcal{T}_1^1(\mathcal{V})$ , and  $\mathcal{T}_2^0(\mathcal{V})$ , the introduction of an inner product connects all these spaces in a natural way and establishes a one-to-one correspondence among them. For example, to a tensor in  $\mathcal{T}_0^2(\mathcal{V})$  with components  $t^{ij}$  there corresponds a unique tensor in  $\mathcal{T}_1^1(\mathcal{V})$ , given, in component form, by  $l_j^i = g_{jk} t^{ik}$ , and another unique tensor in  $\mathcal{T}_2^0(\mathcal{V})$ , given by  $t_{ij} = g_{il} t_j^l = g_{il} g_{jk} t^{lk}$ .

Let us apply this technique to  $g^{ij}$ , which is also a tensor and for which the lowering process is defined. We have

$$g_j^i = g_{ik} g^{kj} = (g^{-1})^{ik} g_{kj} = \delta_j^i. \tag{25.19}$$

This relation holds, of course, in all bases.

The inner product has been defined as a nondegenerate symmetric bilinear form. The important criterion of nondegeneracy has equivalences:



**25.4.3. Proposition.** A symmetric bilinear form  $\mathbf{g}$  is nondegenerate if and only if

1. the matrix of components  $g_{ij}$  has a nonvanishing determinant, or
2. for every nonzero  $\mathbf{v} \in \mathcal{V}$ , there exists  $\mathbf{w} \in \mathcal{V}$  such that  $\mathbf{g}(\mathbf{v}, \mathbf{w}) \neq 0$ .

*Proof.* The first part is a direct consequence of the definition of nondegeneracy. The second part follows from the fact that  $\mathbf{g} : \mathcal{V} \rightarrow \mathcal{V}^*$  is invertible iff the nullity of  $\mathbf{g}$  is zero. It follows that if  $\mathbf{v} \in \mathcal{V}$  is nonzero, then  $\mathbf{g}(\mathbf{v}) \neq 0$ , i.e.,  $\mathbf{g}(\mathbf{v})$  is not the zero functional. Thus, there must exist a vector  $\mathbf{w} \in \mathcal{V}$  such that  $[\mathbf{g}(\mathbf{v})](\mathbf{w}) \neq 0$ . The proposition is proved once we note that  $[\mathbf{g}(\mathbf{v})](\mathbf{w}) \equiv \mathbf{g}(\mathbf{v}, \mathbf{w})$ .  $\square$

**25.4.4. Definition.** A general (not necessarily nondegenerate) symmetric bilinear form  $\mathbf{g}$  can be categorized as follows:

1. **positive (negative) definite:**  $\mathbf{g}(\mathbf{v}, \mathbf{v}) > 0$  [ $\mathbf{g}(\mathbf{v}, \mathbf{v}) < 0$ ] for every nonzero vector  $\mathbf{v}$ ;
2. **definite:**  $\mathbf{g}$  is either positive definite or negative definite;
3. **positive (negative) semidefinite:**  $\mathbf{g}(\mathbf{v}, \mathbf{v}) \geq 0$  [ $\mathbf{g}(\mathbf{v}, \mathbf{v}) \leq 0$ ] for every  $\mathbf{v}$ ;
4. **semidefinite:**  $\mathbf{g}$  is either positive semidefinite or negative semidefinite;
5. **indefinite:**  $\mathbf{g}$  is not definite.

**25.4.5. Example.** Some of the categories of the definition above can be illustrated in  $\mathbb{R}^2$  with  $\mathbf{v}_1 = (x_1, y_1)$ ,  $\mathbf{v}_2 = (x_2, y_2)$ , and  $\mathbf{v} = (x, y)$ .

(a) Positive definite:  $\mathbf{g}(\mathbf{v}_1, \mathbf{v}_2) = x_1x_2 + y_1y_2$  because if  $\mathbf{v} \neq 0$ , then one of its components is nonzero, and  $\mathbf{g}(\mathbf{v}, \mathbf{v}) = x^2 + y^2 > 0$ .

(b) Negative definite:  $\mathbf{g}(\mathbf{v}_1, \mathbf{v}_2) = \frac{1}{2}(x_1y_2 + x_2y_1) - x_1x_2 - y_1y_2$  because  $\mathbf{g}(\mathbf{v}, \mathbf{v}) = xy - x^2 - y^2 = -\frac{1}{2}(x - y)^2 - \frac{1}{2}x^2 - \frac{1}{2}y^2$ , which is negative for nonzero  $\mathbf{v}$ .

(c) Indefinite:  $\mathbf{g}(\mathbf{v}_1, \mathbf{v}_2) = x_1x_2 - y_1y_2$ . For  $\mathbf{v} = (x, x)$ ,  $\mathbf{g}(\mathbf{v}, \mathbf{v}) = 0$ . However,  $\mathbf{g}$  is nondegenerate, because it has the invertible matrix  $\mathbf{g} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  in the standard basis of  $\mathbb{R}^2$ .

(d) Positive semidefinite:  $\mathbf{g}(\mathbf{v}_1, \mathbf{v}_2) = x_1x_2 \Rightarrow \mathbf{g}(\mathbf{v}, \mathbf{v}) = x^2$  and  $\mathbf{g}(\mathbf{v}, \mathbf{v})$  is never negative. However,  $\mathbf{g}$  is degenerate because its matrix in the standard basis of  $\mathbb{R}^2$  is  $\mathbf{g} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ , which is not invertible.  $\blacksquare$

$g$ -orthogonal and null vectors

Two vectors  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$  are said to be  $g$ -orthogonal if  $\mathbf{g}(\mathbf{u}, \mathbf{v}) = 0$ . A null vector of  $\mathbf{g}$  is a vector that is  $g$ -orthogonal to itself. If  $\mathbf{g}$  is definite, then the only null vector is the zero vector. The converse is also true, as the following proposition shows.

**25.4.6. Proposition.** If  $\mathbf{g}$  is not definite, then there exists a nonzero null vector.

*Proof.* That  $\mathbf{g}$  is not positive definite implies that there exists a nonzero vector  $\mathbf{v} \in \mathcal{V}$  such that  $\mathbf{g}(\mathbf{v}, \mathbf{v}) \leq 0$ . Similarly, that  $\mathbf{g}$  is not negative definite implies that there exists a nonzero vector  $\mathbf{w} \in \mathcal{V}$  such that  $\mathbf{g}(\mathbf{w}, \mathbf{w}) \geq 0$ . Construct the vector  $\mathbf{u} = \alpha\mathbf{v} + (1 - \alpha)\mathbf{w}$  and note that  $\mathbf{g}(\mathbf{u}, \mathbf{u})$  is a continuous function of  $\alpha$ . For  $\alpha = 0$  this function has the value  $\mathbf{g}(\mathbf{w}, \mathbf{w}) \geq 0$ , and for  $\alpha = 1$  it has the value  $\mathbf{g}(\mathbf{v}, \mathbf{v}) \leq 0$ . Thus, there must be some  $\alpha$  for which  $\mathbf{g}(\mathbf{u}, \mathbf{u}) = 0$ .  $\square$

**25.4.7. Example.** In the special theory of relativity, the inner product of two “position” four-vectors,  $\mathbf{r}_1 = (x_1, y_1, z_1, ct_1)$  and  $\mathbf{r}_2 = (x_2, y_2, z_2, ct_2)$ , where  $c$  is the speed of light, is defined as

$$\mathbf{g}(\mathbf{r}_1, \mathbf{r}_2) = -x_1x_2 - y_1y_2 - z_1z_2 + c^2t_1t_2.$$

This is clearly an indefinite symmetric bilinear form. Proposition 25.4.6 tells us that there must exist a nonzero null vector. Such a vector  $\mathbf{r}$  satisfies

$$\mathbf{g}(\mathbf{r}, \mathbf{r}) = c^2t^2 - x^2 - y^2 - z^2 = 0,$$

or  $c^2 = \frac{x^2 + y^2 + z^2}{t^2} \Rightarrow c = \pm \frac{\sqrt{x^2 + y^2 + z^2}}{t} = \pm \frac{\text{distance}}{\text{time}}$ . This corresponds to a particle moving with the speed of light. Thus, light rays are the null vectors in the special theory of relativity. ■

Whenever there is an inner product on a vector space, there is the possibility of orthogonal basis vectors. However, since  $\mathbf{g}(\mathbf{v}, \mathbf{v})$  is allowed to be negative or zero, it is impossible to demand normality for some vectors.

*g*-orthonormal vectors, diagonal components of *g*, index and signature of *g*

**25.4.8. Definition.** A basis  $\{\mathbf{e}_i\}_{i=1}^N$  of  $\mathcal{V}$  is *g*-orthonormal if  $\mathbf{g}(\mathbf{e}_i, \mathbf{e}_j) = 0$  for  $i \neq j$ , and  $\mathbf{g}(\mathbf{e}_i, \mathbf{e}_i)$  (no sum!) is  $+1, -1$ , or  $0$ . The  $\mathbf{g}(\mathbf{e}_i, \mathbf{e}_i)$  are called the **diagonal components** of  $\mathbf{g}$ . We use  $n_+, n_-,$  and  $n_0$  to denote the number of vectors  $\mathbf{e}_i$  for which  $\mathbf{g}(\mathbf{e}_i, \mathbf{e}_i)$  is, respectively,  $+1, -1$ , or  $0$ . The integer  $n_-$  is called the **index** of  $\mathbf{g}$ , and  $s = n_+ - n_-$  is called the **signature** of  $\mathbf{g}$ .

The existence of orthonormal bases was established for positive definite  $\mathbf{g}$  by the Gram–Schmidt orthonormalization process in Chapter 1. One of the steps in this process is division by (the square root of)  $\mathbf{g}(\mathbf{v}, \mathbf{v})$ , which could be zero in the present context. A slight modification of the Gram–Schmidt process, in conjunction with the procedure used in the symplectic case, can help construct an orthonormal basis for a general (possibly indefinite)  $\mathbf{g}$ . To be specific, start with the **polarization identity**,

polarization identity

$$\mathbf{g}(\mathbf{v}, \mathbf{v}') = \frac{1}{4}[\mathbf{g}(\mathbf{v} + \mathbf{v}', \mathbf{v} + \mathbf{v}') - \mathbf{g}(\mathbf{v} - \mathbf{v}', \mathbf{v} - \mathbf{v}')],$$

and use it to convince yourself that  $\mathbf{g}$  is identically zero unless there exists a vector  $\mathbf{e}_1$  such that  $\mathbf{g}(\mathbf{e}_1, \mathbf{e}_1) \neq 0$ . Rescaling, we can assume that  $\mathbf{g}(\mathbf{e}_1, \mathbf{e}_1) \equiv \eta_1 = \pm 1$ . Let  $\mathcal{V}_1$  be the span of  $\mathbf{e}_1$  and  $\mathcal{V}_2$  the  $\mathbf{g}$ -orthogonal complement of  $\mathcal{V}_1$ . Clearly,  $\mathcal{V}_1 \cap \mathcal{V}_2 = \{0\}$ ; moreover,  $\mathcal{V}_1 + \mathcal{V}_2 = \mathcal{V}$ , for if  $\mathbf{v} \in \mathcal{V}$ , we can write

$$\mathbf{v} = \underbrace{\eta_1 \mathbf{g}(\mathbf{v}, \mathbf{e}_1) \mathbf{e}_1}_{\in \mathcal{V}_1} + \underbrace{\mathbf{v} - \eta_1 \mathbf{g}(\mathbf{v}, \mathbf{e}_1) \mathbf{e}_1}_{\in \mathcal{V}_2}.$$

Now, if  $\mathbf{g} = 0$  on  $\mathcal{V}_2$ , we are done; otherwise, there is a vector  $\mathbf{e}_2$  such that  $\mathbf{g}(\mathbf{e}_2, \mathbf{e}_2) \equiv \eta_2 = \pm 1$ . We continue inductively to complete the construction. We therefore have the following:

Sylvester's theorem **25.4.9. Theorem.** For every symmetric bilinear form  $\mathbf{g}$  on  $\mathcal{V}$ , there is an orthonormal basis. Furthermore,  $n_+$ ,  $n_-$ , and  $n_0$  are the same in all orthonormal bases.

Orthonormal bases give the same oriented volume element. For a proof of the last statement, also known as **Sylvester's theorem**, see [Bish 80, p. 104]. Orthonormal bases allow us to speak of the oriented volume element. Suppose  $\{\epsilon^j\}_{j=1}^N$  is an oriented basis of  $\mathcal{V}^*$ . If  $\{\varphi^k\}_{k=1}^N$  is another orthonormal basis in the same orientation and related to  $\{\epsilon^j\}$  by a matrix  $\mathbf{R}$ , then

$$\varphi^1 \wedge \varphi^2 \wedge \cdots \wedge \varphi^N = (\det \mathbf{R}) \epsilon^1 \wedge \epsilon^2 \wedge \cdots \wedge \epsilon^N.$$

Since  $\{\varphi^k\}$  and  $\{\epsilon^j\}$  are orthonormal, the determinant of  $\mathbf{g}$  is  $(-1)^{n_-}$  in both of them. Problem 25.17 then implies that  $(\det \mathbf{R})^2 = 1$  or  $\det \mathbf{R} = \pm 1$ . However,  $\{\varphi^k\}$  and  $\{\epsilon^j\}$  belong to the same orientation. Thus,  $\det \mathbf{R} = +1$ , and  $\{\varphi^k\}$  and  $\{\epsilon^j\}$  give the same volume element.

**25.4.10. Example.** Let  $\mathcal{V} = \mathbb{R}^3$  and  $\mathbf{v}_1 = (x_1, y_1, z_1)$ ,  $\mathbf{v}_2 = (x_2, y_2, z_2)$ , and  $\mathbf{v} = (x, y, z)$ . Define the symmetric bilinear form

$$\mathbf{g}(\mathbf{v}_1, \mathbf{v}_2) = \frac{1}{2}(x_1 y_2 + x_2 y_1 + y_1 z_2 + y_2 z_1 + x_1 z_2 + x_2 z_1)$$

so that  $\mathbf{g}(\mathbf{v}, \mathbf{v}) = xy + yz + xz$ . We wish to find a set of vectors in  $\mathbb{R}^3$  that are orthonormal with respect to this bilinear form. Clearly,  $\mathbf{e}_1 = (1, 1, 0)$  is such that  $\mathbf{g}(\mathbf{e}_1, \mathbf{e}_1) = 1$ . So  $\mathbf{e}_1$  is one of our vectors. Consider  $\mathbf{v} = (1, 0, 1)$  and note that the vector  $\mathbf{r}_2 = \mathbf{v} - [\mathbf{g}(\mathbf{v}, \mathbf{e}_1)/\mathbf{g}(\mathbf{e}_1, \mathbf{e}_1)]\mathbf{e}_1$ , suggested by the Gram-Schmidt process, is orthogonal to  $\mathbf{e}_1$ . Furthermore, it is easily verified that  $\mathbf{g}(\mathbf{r}_2, \mathbf{r}_2) = -\frac{5}{4}$ . Therefore, our second vector is

$$\mathbf{e}_2 = \frac{\mathbf{r}_2}{\sqrt{|\mathbf{g}(\mathbf{r}_2, \mathbf{r}_2)|}} = \left(-\frac{1}{\sqrt{5}}, -\frac{3}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$

with  $\mathbf{g}(\mathbf{e}_2, \mathbf{e}_2) = -1$ . Finally, we take  $\mathbf{w} = (0, 1, 1)$ . Then

$$\mathbf{r}_3 = \mathbf{w} - \frac{\mathbf{g}(\mathbf{w}, \mathbf{e}_1)}{\mathbf{g}(\mathbf{e}_1, \mathbf{e}_1)}\mathbf{e}_1 - \frac{\mathbf{g}(\mathbf{w}, \mathbf{e}_2)}{\mathbf{g}(\mathbf{e}_2, \mathbf{e}_2)}\mathbf{e}_2 = \frac{4}{10}(-3, 1, 1)$$

will be orthogonal to both  $\mathbf{e}_1$  and  $\mathbf{e}_2$  with  $\mathbf{g}(\mathbf{r}_3, \mathbf{r}_3) = -\frac{4}{5}$ . Thus, the third vector can be chosen to be

$$\mathbf{e}_3 = \frac{\mathbf{r}_3}{\sqrt{|\mathbf{g}(\mathbf{r}_3, \mathbf{r}_3)|}} = \left(-\frac{3}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right),$$

and we obtain  $\mathbf{g}(\mathbf{e}_1, \mathbf{e}_1) = 1$ ,  $\mathbf{g}(\mathbf{e}_2, \mathbf{e}_2) = -1$ ,  $\mathbf{g}(\mathbf{e}_3, \mathbf{e}_3) = -1$ ,  $\mathbf{g}(\mathbf{e}_i, \mathbf{e}_j) = 0$  for  $i \neq j$ . We also have  $n_+ = 1$ ,  $n_- = 2$ , and  $n_0 = 0$ , i.e., the index of  $\mathbf{g}$  is 2, and its signature is  $-1$ . Although we have worked in a particular basis, Theorem 25.4.9 guarantees that  $n_+$ ,  $n_-$ , and  $n_0$  are (orthonormal) basis-independent. ■

We should emphasize that the invariance of  $n_+$ ,  $n_-$ , and  $n_0$  is true for  $g$ -orthonormal bases. As a counterexample, consider  $\mathbf{g}$  of Example 25.4.10 applied to the standard basis of  $\mathbb{R}^3$ , which we designate with a prime. It is readily verified that

$$\mathbf{g}(\mathbf{e}'_i, \mathbf{e}'_i) = 0 \quad \text{for } i = 1, 2, 3.$$

So it might appear that  $n_0 = 3$  for this basis. However, the standard basis is *not*  $g$ -orthonormal. In fact,

$$\mathbf{g}(\mathbf{e}'_1, \mathbf{e}'_2) = \frac{1}{2} = \mathbf{g}(\mathbf{e}'_1, \mathbf{e}'_3) = \mathbf{g}(\mathbf{e}'_2, \mathbf{e}'_3).$$

That is why the nonstandard vectors  $\mathbf{e}_1, \mathbf{v}$ , and  $\mathbf{w}$  were chosen in Example 25.4.10.

In an orthonormal basis the matrix of  $\mathbf{g}$  is diagonal, with entries  $+1, -1$ , and  $0$ . In particular, if  $\mathbf{g}$  is to be nondegenerate, that is, to be an inner product,  $n_0$  must be zero. Thus, a general inner product on an  $N$ -dimensional vector space satisfies the conditions  $n_+ + n_- = N$  and  $n_+ - n_- = s$  which give  $s = N - 2n_-$ .

**25.4.11. Definition.** An inner product space with  $n_- = 1$  or  $n_- = N - 1$  is called a *Minkowski space*. For  $N = 4$  this is the space of the special theory of relativity. An inner product space with  $n_- = 0$  (or  $n_- = N$ ) is called a *Euclidean space*.

**25.4.12. Example.** Let  $\{\mathbf{e}_i\}_{i=1}^N$  be a basis of  $\mathcal{V}$  and  $\{\epsilon^i\}_{i=1}^N$  its dual. We can define the permutation tensor

$$\delta_{j_1 j_2 \dots j_N}^{i_1 i_2 \dots i_N} = \epsilon^{i_1} \wedge \epsilon^{i_2} \wedge \dots \wedge \epsilon^{i_N}(\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_N}). \tag{25.20}$$

It is clear from this definition that  $\delta_{j_1 j_2 \dots j_N}^{i_1 i_2 \dots i_N}$  is completely skew-symmetric in all upper indices. That it is also skew-symmetric in the lower indices can be seen as follows. Assume that two of the lower indices are equal. This means having two  $\mathbf{e}_j$ 's equal in (25.20). These two  $\mathbf{e}_j$ 's will contract with two  $\epsilon^i$ 's, say  $\epsilon^k$  and  $\epsilon^l$ . Thus, in the expansion there will be a term  $C\epsilon^k(\mathbf{e}_j)\epsilon^l(\mathbf{e}_j)$ , where  $C$  is the product of all the other factors. Since the product is completely skew-symmetric in the upper indices, there must also exist another term, with a minus sign and in which the upper indices  $k$  and  $l$  are interchanged:  $-C\epsilon^l(\mathbf{e}_j)\epsilon^k(\mathbf{e}_j)$ . This makes the sum zero, and by Theorem 25.2.7, (25.20) is antisymmetric in the lower indices as well.

This suggests that  $\delta_{j_1 j_2 \dots j_N}^{i_1 i_2 \dots i_N} \propto \epsilon^{i_1 i_2 \dots i_N} \epsilon_{j_1 j_2 \dots j_N}$ . To find the proportionality constant, we note that (see Problem 25.16)

$$\delta_{12 \dots N}^{12 \dots N} = \sum_{\pi} \epsilon_{\pi(1)\pi(2)\dots\pi(N)} \delta_{\pi(1)}^1 \delta_{\pi(2)}^2 \dots \delta_{\pi(N)}^N.$$

The only contribution to the sum comes from the permutation with the property  $\pi(i) = i$ . This is the identity permutation for which  $\epsilon_{\pi} = 1$ . Thus, we have  $\delta_{12 \dots N}^{12 \dots N} = 1$ . On the other hand,

$$\epsilon^{12 \dots N} \epsilon_{12 \dots N} = \epsilon^{12 \dots N} = (-1)^{n_-}.$$

Therefore, the proportionality constant is  $(-1)^{n_-}$ . Thus

$$\epsilon^{i_1 i_2 \dots i_N} \epsilon_{j_1 j_2 \dots j_N} = (-1)^{n_-} \delta_{j_1 j_2 \dots j_N}^{i_1 i_2 \dots i_N}. \quad \blacksquare$$

We can find an explicit expression for the permutation tensor of Example 25.4.12. Expanding the RHS of Equation (25.20) using Equation (25.10), we obtain

$$\begin{aligned} \delta_{j_1 j_2 \dots j_N}^{i_1 i_2 \dots i_N} &= \sum_{\pi} \epsilon_{\pi} \delta_{\pi(j_1)}^{i_1} \delta_{\pi(j_2)}^{i_2} \dots \delta_{\pi(j_N)}^{i_N}, \\ \epsilon^{i_1 i_2 \dots i_N} \epsilon_{j_1 j_2 \dots j_N} &= (-1)^{n_-} \sum_{\pi} \epsilon_{\pi} \delta_{\pi(j_1)}^{i_1} \delta_{\pi(j_2)}^{i_2} \dots \delta_{\pi(j_N)}^{i_N}. \end{aligned} \tag{25.21}$$

Furthermore, the first equation of (25.21) can be written concisely as a determinant. To see this, first note that

$$\delta_{12\dots N}^{i_1 i_2 \dots i_N} = \sum_{\pi} \epsilon_{\pi(1)\pi(2)\dots\pi(N)} \delta_{\pi(1)}^{i_1} \delta_{\pi(2)}^{i_2} \cdots \delta_{\pi(N)}^{i_N}.$$

The RHS is clearly the determinant of a matrix (expanded with respect to the  $i$ th row) whose elements are  $\delta_k^{ik}$ . The same holds true if  $1, 2, \dots, N$  is replaced by  $j_1, j_2, \dots, j_N$ ; thus,

$$\delta_{j_1 j_2 \dots j_N}^{i_1 i_2 \dots i_N} = \det \begin{pmatrix} \delta_{j_1}^{i_1} & \delta_{j_2}^{i_1} & \cdots & \delta_{j_N}^{i_1} \\ \delta_{j_1}^{i_2} & \delta_{j_2}^{i_2} & \cdots & \delta_{j_N}^{i_2} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{j_1}^{i_N} & \delta_{j_2}^{i_N} & \cdots & \delta_{j_N}^{i_N} \end{pmatrix} \quad (25.22)$$

**25.4.13. Example.** Let us apply the second equation of (25.21) to Euclidean  $\mathbb{R}^3$ :

$$\epsilon^{ijk} \epsilon_{lmn} = \delta_i^j \delta_m^j \delta_n^k - \delta_i^j \delta_n^j \delta_m^k - \delta_m^i \delta_l^j \delta_n^k + \delta_m^i \delta_n^j \delta_l^k + \delta_n^i \delta_l^j \delta_m^k - \delta_n^i \delta_m^j \delta_l^k.$$

From this fundamental relation, we can obtain other useful formulas. For example, setting  $n = k$  and summing over  $k$ , we get

$$\epsilon^{ijk} \epsilon_{lmk} = 3\delta_l^j \delta_m^j - \delta_l^i \delta_m^j - 3\delta_m^i \delta_l^j + \delta_m^i \delta_l^j + \delta_m^i \delta_l^j - \delta_l^i \delta_m^j = \delta_l^i \delta_m^j - \delta_m^i \delta_l^j.$$

Now set  $m = j$  in this equation and sum over  $j$ :

$$\epsilon^{ijk} \epsilon_{ljk} = 3\delta_l^i - \delta_l^i = 2\delta_l^i.$$

Finally, let  $l = i$  and sum over  $i$ :

$$\epsilon^{ijk} \epsilon_{ijk} = 2\delta_i^i = 2 \cdot 3 = 3!.$$

In general,

$$\epsilon^{i_1 i_2 \dots i_N} \epsilon_{i_1 i_2 \dots i_N} = (-1)^{n-N} N!,$$

$$\epsilon^{i_1 i_2 \dots i_{N-1} i_N} \epsilon_{i_1 i_2 \dots i_{N-1} j_N} = (-1)^{n-(N-1)} \delta_{j_N}^{i_N},$$

$$\epsilon^{i_1 i_2 \dots i_{N-2} i_{N-1} i_N} \epsilon_{i_1 i_2 \dots i_{N-2} j_{N-1} j_N} = (-1)^{n-(N-2)} (\delta_{j_{N-1}}^{i_{N-1}} \delta_{j_N}^{i_N} - \delta_{j_N}^{i_{N-1}} \delta_{j_{N-1}}^{i_N}),$$

and so on. ■

**25.4.14. Example.** As an application of the foregoing formalism, we can express the determinant of a  $2 \times 2$  matrix in terms of traces. Let  $A$  be such a matrix with elements  $A_j^i$ . Then

$$\begin{aligned} \det A &= \epsilon_{ij} A_1^i A_2^j = \frac{1}{2} (\epsilon_{ij} A_1^i A_2^j - \epsilon_{ij} A_2^i A_1^j) = \frac{1}{2} (\epsilon_{ij} \epsilon^{kl} A_k^i A_l^j) \\ &= \frac{1}{2} A_k^i A_l^j (\delta_i^k \delta_j^l - \delta_j^k \delta_i^l) = \frac{1}{2} (A_i^i A_j^j - A_j^i A_i^j) \\ &= \frac{1}{2} [(\operatorname{tr} A)(\operatorname{tr} A) - (A^2)_i^i] = \frac{1}{2} [(\operatorname{tr} A)^2 - \operatorname{tr}(A^2)]. \end{aligned} \quad \blacksquare$$

We can generalize the result of the example above and express the determinant of an  $N \times N$  matrix as

$$\begin{aligned} \det A &= \frac{1}{N!} \epsilon^{i_1 i_2 \dots i_N} \epsilon_{j_1 j_2 \dots j_N} A_{i_1}^{j_1} A_{i_2}^{j_2} \dots A_{i_N}^{j_N} \\ &= \frac{1}{N!} \sum_{\pi} \epsilon_{\pi} \delta_{\pi(j_1)}^{i_1} \delta_{\pi(j_2)}^{i_2} \dots \delta_{\pi(j_N)}^{i_N} A_{i_1}^{j_1} A_{i_2}^{j_2} \dots A_{i_N}^{j_N}. \end{aligned} \tag{25.23}$$

### 25.5 The Hodge Star Operator

It was established in Chapter 2 that all vector spaces of the same dimension are isomorphic (identical). Therefore, the two vector spaces  $\Lambda^p(\mathcal{V})$  and  $\Lambda^{N-p}(\mathcal{V})$  having the same dimension,  $\binom{N}{p} = \binom{N}{N-p}$ , must be isomorphic. In fact, there is a *natural* isomorphism between the two spaces:

**Hodge star operator** **25.5.1. Definition.** Let  $\mathbf{g}$  be an inner product and  $\{\mathbf{e}_i\}_{i=1}^N$  an ordered g-orthonormal basis of  $\mathcal{V}$ . The **Hodge star operator** is a linear mapping,  $*$  :  $\Lambda^p(\mathcal{V}) \rightarrow \Lambda^{N-p}(\mathcal{V})$ , given by (remember Einstein's summation convention!)

$$*(\mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_p}) \equiv \frac{1}{(N-p)!} \epsilon^{i_{p+1} \dots i_N} \epsilon_{i_1 \dots i_p} \mathbf{e}_{i_{p+1}} \wedge \dots \wedge \mathbf{e}_{i_N}, \tag{25.24}$$

where all the indices of the  $\epsilon$  tensor are raised by  $g^{ij}$ .

Although this definition is based on a choice of basis, it can be shown that the operator is in fact basis-independent. We note that  $\epsilon_{i_1 \dots i_p}^{i_{p+1} \dots i_N} = \pm \epsilon_{i_1 \dots i_p i_{p+1} \dots i_N}$  because each index raising introduces either a +1 or a -1. In particular, for Euclidean spaces, in which  $n_- = 0$ , the two epsilon symbols are the same.

**25.5.2. Example.** Let us apply Definition 25.5.1 to  $\Lambda^p(\mathbb{R}^3)$  for  $p = 0, 1, 2, 3$ . Let  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  be an oriented orthonormal basis of  $\mathbb{R}^3$ .

(a) For  $\Lambda^0(\mathbb{R}^3) = \mathbb{R}$  a basis is 1, and (25.24) gives

$$*1 = \frac{1}{3!} \epsilon^{ijk} \mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k = \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3.$$

(b) For  $\Lambda^1(\mathbb{R}^3) = \mathbb{R}^3$  a basis is  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ , and (25.24) gives  $*\mathbf{e}_i = \frac{1}{2!} \epsilon_i^{jk} \mathbf{e}_j \wedge \mathbf{e}_k$ , or  $*\mathbf{e}_1 = \mathbf{e}_2 \wedge \mathbf{e}_3$ ,  $*\mathbf{e}_2 = \mathbf{e}_3 \wedge \mathbf{e}_1$ ,  $*\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$ .

(c) For  $\Lambda^2(\mathbb{R}^3)$  a basis is  $\{\mathbf{e}_1 \wedge \mathbf{e}_2, \mathbf{e}_1 \wedge \mathbf{e}_3, \mathbf{e}_2 \wedge \mathbf{e}_3\}$ , and (25.24) gives  $*\mathbf{e}_i \wedge \mathbf{e}_j = \epsilon_{ij}^k \mathbf{e}_k$ , or  $*(\mathbf{e}_1 \wedge \mathbf{e}_2) = \mathbf{e}_3$ ,  $*(\mathbf{e}_1 \wedge \mathbf{e}_3) = -\mathbf{e}_2$ ,  $*(\mathbf{e}_2 \wedge \mathbf{e}_3) = \mathbf{e}_1$ .

(d) For  $\Lambda^3(\mathbb{R}^3)$  a basis is  $\{\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3\}$ , and (25.24) yields

$$*(\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3) = \epsilon_{123} = 1. \quad \blacksquare$$

The preceding example may suggest that applying the Hodge star operator twice (composition of  $*$  with itself, or  $* \circ *$ ) is equivalent to applying the identity operator. This is partially true. The following theorem is a precise statement of this conjecture. (For a proof, see [Bish 80, p. 111].)

**25.5.3. Theorem.** Let  $\mathcal{V}$  be an oriented space with an inner product  $\mathbf{g}$ . For  $\mathbf{A} \in \Lambda^p(\mathcal{V})$ , we have

$$* \circ * \mathbf{A} \equiv ** \mathbf{A} = (-1)^{n-} (-1)^{p(N-p)} \mathbf{A}, \tag{25.25}$$

where  $n_-$  is the index of  $\mathbf{g}$  and  $N = \dim \mathcal{V}$ .

In particular, for Euclidean spaces with an odd number of dimensions (such as  $\mathbb{R}^3$ ),  $** \mathbf{A} = \mathbf{A}$ .

One can extend the star operation to any  $\mathbf{A} \in \Lambda^p(\mathcal{V})$  by writing  $\mathbf{A}$  as a linear combination of basis vectors of  $\Lambda^p(\mathcal{V})$  constructed out of  $\{\mathbf{e}_i\}_{i=1}^N$ , and using the linearity of  $*$ .

antisymmetric tensors with numerical coefficients

In the discussion of exterior algebra one encounters sums of the form

$$A^{i_1 \dots i_p} \mathbf{v}_{i_1} \wedge \dots \wedge \mathbf{v}_{i_p}.$$

It is important to keep in mind that  $A^{i_1 \dots i_p}$  is assumed skew-symmetric. For example, if  $\mathbf{A} = \mathbf{e}_1 \wedge \mathbf{e}_2$ , then in the sum  $\mathbf{A} = A^{ij} \mathbf{e}_i \wedge \mathbf{e}_j$ , the nonzero components consist of  $A^{12} = \frac{1}{2}$  and  $A^{21} = -\frac{1}{2}$ . Similarly, when  $\mathbf{B} = \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$  is written in the form  $\mathbf{B} = B^{ijk} \mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k$ , it is understood that the nonzero components of  $\mathbf{B}$  are not restricted to  $B^{123}$ . Other components, such as  $B^{132}$ ,  $B^{231}$ , and so on, are also nonzero. In fact, we have

$$B^{123} = -B^{132} = -B^{213} = B^{231} = B^{312} = -B^{321} = \frac{1}{6}.$$

This should be kept in mind when sums over exterior products with numerical coefficients are encountered.

Cross product is defined only in three dimensions!

**25.5.4. Example.** Let  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$  and  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  an oriented orthonormal basis of  $\mathbb{R}^3$ . Then  $\mathbf{a} = a^i \mathbf{e}_i$  and  $\mathbf{b} = b^j \mathbf{e}_j$ . Let us calculate  $\mathbf{a} \wedge \mathbf{b}$  and  $*(\mathbf{a} \wedge \mathbf{b})$ . We assume a Euclidean  $\mathbf{g}$  on  $\mathbb{R}^3$ . Then  $\mathbf{a} \wedge \mathbf{b} = (a^i \mathbf{e}_i) \wedge (b^j \mathbf{e}_j) = a^i b^j \mathbf{e}_i \wedge \mathbf{e}_j$ , and

$$*(\mathbf{a} \wedge \mathbf{b}) = *(a^i \mathbf{e}_i) \wedge (b^j \mathbf{e}_j) = a^i b^j *(\mathbf{e}_i \wedge \mathbf{e}_j) = a^i b^j (\epsilon_{ij}^k \mathbf{e}_k) = (\epsilon_{ij}^k a^i b^j) \mathbf{e}_k.$$

We see that  $*(\mathbf{a} \wedge \mathbf{b})$  is a vector with components  $[(\mathbf{a} \times \mathbf{b})]^k = \epsilon_{ij}^k a^i b^j$ , which are precisely the components of  $\mathbf{a} \times \mathbf{b}$ .

The correspondence between  $\mathbf{a} \wedge \mathbf{b}$  and  $\mathbf{a} \times \mathbf{b}$  holds *only in three dimensions*, because  $\dim \Lambda^1(\mathcal{V}) = \dim \Lambda^2(\mathcal{V})$  only if  $\dim \mathcal{V} = 3$ . That is why the cross product can be defined—as a “machine” that takes two vectors in  $\mathcal{V}$  and manufactures a *vector* in  $\mathcal{V}$ —only if  $\mathcal{V}$  is three-dimensional. ■

**25.5.5. Example.** We can use the results of Examples 25.4.13 and 25.5.4 to establish a sample of familiar vector identities componentwise.

(a) For the triple cross product, we have

$$\begin{aligned} [\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]^k &= \epsilon_{ij}^k a^i (\mathbf{b} \times \mathbf{c})^j = \epsilon_{ij}^k a^i (\epsilon_{lm}^j b^l c^m) = a_i b^l c^m \epsilon^{kij} \epsilon_{jlm} \\ &= a_i b^l c^m \epsilon^{kij} \epsilon_{lmj} = a_i b^l c^m (\delta_i^k \delta_m^j - \delta_m^k \delta_i^j) \\ &= a_i b^k c^l - a_i b^l c^k = (\mathbf{a} \cdot \mathbf{c}) b^k - (\mathbf{a} \cdot \mathbf{b}) c^k, \end{aligned}$$

which is the  $k$ th component of  $\mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$ . In deriving the above "bac cab" rule, we used the fact that one can swap an upper index with the same lower index:  $a^i b_i = a_i b^i$

(b) Next we show the familiar statement that the divergence of curl is zero. Let  $\partial_i$  denote differentiation with respect to  $x_i$ . Then

$$\begin{aligned}\nabla \cdot (\nabla \times \mathbf{a}) &= \partial_i (\nabla \times \mathbf{a})^i = \partial_i \epsilon_{ijk}^i \partial^j a^k = \epsilon^{ijk} \partial_i \partial_j a_k \\ &= -\epsilon^{jik} \partial_i \partial_j a_k = -\epsilon^{jik} \partial_j \partial_i a_k = -\partial_j (\epsilon^{jik} \partial_i a_k) \\ &= -\partial_j (\nabla \times \mathbf{a})^j = -\nabla \cdot (\nabla \times \mathbf{a}) \Rightarrow \nabla \cdot (\nabla \times \mathbf{a}) = 0.\end{aligned}$$

The above steps show in general that

**25.5.6. Box.** When the product of two tensors is summed over a pair of indices in which one of the tensors is symmetric and the other antisymmetric, the result is zero.

(c) Finally, we show that curl of gradient is zero:

$$[\nabla \times (\nabla f)]^k = \epsilon_{ijk}^i \partial^j \partial^k f = \epsilon^{ijk} \partial_j \partial_k f = 0,$$

because  $\epsilon^{ijk}$  is antisymmetric in  $jk$ , while  $\partial_j \partial_k f$  is symmetric in  $jk$ . ■

## 25.6 Problems

**25.1.** Show that the mapping  $\mathbf{v} : \mathcal{V}^* \rightarrow \mathbb{R}$  given by  $\mathbf{v}(\tau) = \tau(\mathbf{v})$  is linear.

**25.2.** Show that the components of a tensor product are the products of the components of the factors:

$$(\mathbf{U} \otimes \mathbf{T})_{j_1 \dots j_{s+l}}^{i_1 \dots i_{r+k}} = U_{j_1 \dots j_s}^{i_1 \dots i_r} T_{j_{s+1} \dots j_{s+l}}^{i_{r+1} \dots i_{r+k}}.$$

**25.3.** Show that  $\mathbf{e}_{j_1} \otimes \dots \otimes \mathbf{e}_{j_r} \otimes \epsilon^{i_1} \otimes \dots \otimes \epsilon^{i_s}$  are linearly independent. Hint: Consider  $A_{i_1 \dots i_s}^{j_1 \dots j_r} \mathbf{e}_{j_1} \otimes \dots \otimes \mathbf{e}_{j_r} \otimes \epsilon^{i_1} \otimes \dots \otimes \epsilon^{i_s} = 0$  and evaluate the LHS on appropriate tensors to show that all coefficients are zero.

**25.4.** What is the tensor product of  $\mathbf{A} = 2\mathbf{e}_x - \mathbf{e}_y + 3\mathbf{e}_z$  with itself?

**25.5.** If  $\mathbf{A} \in \mathcal{L}(\mathcal{V})$  is represented by  $A_j^i$  in the basis  $\{\mathbf{e}_i\}$  and by  $A_i'^k$  in  $\{\mathbf{e}'_k\}$ , then show that

$$A_i'^k \mathbf{e}'_k \otimes \epsilon'^l = A_j^i \mathbf{e}_i \otimes \epsilon^j,$$

where  $\{\epsilon^j\}$  and  $\{\epsilon'^l\}$  are dual to  $\{\mathbf{e}_i\}$  and  $\{\mathbf{e}'_k\}$ , respectively.

**25.6.** Prove that the linear functional  $\mathbf{F} : \mathcal{V} \rightarrow \mathbb{R}$  is a linear invariant, i.e., basis-independent, function.



- 25.7. Show that  $\text{tr} : \mathcal{T}_1^1 \rightarrow \mathbb{R}$  is an invariant linear function.
- 25.8. If  $\mathbf{A}$  is skew-symmetric in some pair of variables, show that  $\mathbb{S}(\mathbf{A}) = 0$ .
- 25.9. Using the exterior product show whether the following three vectors are linearly dependent or independent:

$$\begin{aligned} \mathbf{v}_1 &= 2\mathbf{e}_1 - \mathbf{e}_2 + 3\mathbf{e}_3 - \mathbf{e}_4, \\ \mathbf{v}_2 &= -\mathbf{e}_1 + 3\mathbf{e}_2 - 2\mathbf{e}_4, \\ \mathbf{v}_3 &= 3\mathbf{e}_1 + 2\mathbf{e}_2 - 4\mathbf{e}_3 + \mathbf{e}_4. \end{aligned}$$

25.10. Let  $\mathbf{A} \in \mathcal{T}_0^r(\mathcal{V})$  be skew-symmetric. Show that if  $\tau^1, \dots, \tau^r \in \mathcal{V}^*$  are linearly dependent, then  $\mathbf{A}(\tau^1, \dots, \tau^r) = 0$ .

25.11. Show that  $\{\mathbf{e}_k \wedge \mathbf{e}_i\}$  with  $k < i$  are linearly independent.

25.12. Let  $\mathbf{v} \in \mathcal{V}$  be nonzero, and let  $\mathbf{A} \in \Lambda^p(\mathcal{V})$ . Show that  $\mathbf{v} \wedge \mathbf{A} = 0$  if and only if there exists  $\mathbf{B} \in \Lambda^{p-1}(\mathcal{V})$  such that  $\mathbf{A} = \mathbf{v} \wedge \mathbf{B}$ . Hint: Let  $\mathbf{v}$  be the first vector of a basis; separate out  $\mathbf{v}$  in the expansion of  $\mathbf{A}$  in terms of the  $p$ -fold wedge products of basis vectors, and multiply the result by  $\mathbf{v}$ .

25.13. Let  $\mathbf{A} \in \Lambda^2(\mathcal{V})$  with components  $A^{ij}$ . Show that  $\mathbf{A} \wedge \mathbf{A} = 0$  if and only if  $A^{ij}A^{kl} - A^{ik}A^{jl} + A^{il}A^{jk} = 0$  for all  $i, j, k, l$  in any basis.

25.14. A linear operator acting on a vector in one dimension simply multiplies that vector by some constant. Show that this constant is independent of the vector chosen. That is, the constant is an intrinsic property of the operator.

25.15. Let  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  be any basis in  $\mathbb{R}^3$ . Define an operator  $\mathbf{E} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  that permutes any set of three vectors  $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$  to  $\{\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_k\}$ . Find the matrix representation of this operator and show that  $\det \mathbf{E} = \epsilon_{ijk}$ .

25.16. (a) Starting with the definition of the permutation tensor  $\delta_{j_1 j_2 \dots j_N}^{i_1 i_2 \dots i_N}$ , and writing the wedge product in terms of the antisymmetrized tensor product, show that

$$\delta_{j_1 j_2 \dots j_N}^{i_1 i_2 \dots i_N} = \sum_{\pi} \epsilon_{\pi(j_1) \pi(j_2) \dots \pi(j_N)} \delta_{\pi(j_1)}^{i_1} \delta_{\pi(j_2)}^{i_2} \dots \delta_{\pi(j_N)}^{i_N}.$$

(b) Show that the inverse of the (diagonal) matrix of  $\mathbf{g}$  in an orthonormal basis is the same as the matrix of  $\mathbf{g}$ .

(c) Now show that  $\epsilon^{12 \dots N} = (-1)^{n-} \epsilon_{12 \dots N} = (-1)^{n-}$ .

25.17. Let  $\{\mathbf{e}_i\}_{i=1}^N$  be a  $g$ -orthonormal basis of  $\mathcal{V}$ . Let  $\eta_{ij} = \pm \delta_j^i$  be the matrix of  $\mathbf{g}$  in this orthonormal basis. Let  $\{\mathbf{v}_j\}_{j=1}^N$  be another (not necessarily orthonormal) basis of  $\mathcal{V}$  with a transformation matrix  $\mathbf{R}$ . Using  $\mathbf{G}$  to denote the matrix of  $\mathbf{g}$  in  $\{\mathbf{v}_j\}$ , show that

$$\det \mathbf{G} = \det \eta (\det \mathbf{R})^2 = (-1)^{n-} (\det \mathbf{R})^2.$$

In particular, the sign of this determinant is invariant. Why is  $\det \mathbf{G}$  not equal to  $\det \boldsymbol{\eta}$ ? Is there any conflict with the statement that the determinant is basis-independent?

**25.18.** Show that the kernel of  $\mathbf{g} : \mathcal{V} \rightarrow \mathcal{V}^*$  consists of all vectors  $\mathbf{u} \in \mathcal{V}$  such that  $\mathbf{g}(\mathbf{u}, \mathbf{v}) = 0$  for all  $\mathbf{v} \in \mathcal{V}$ . Show also that in the  $\mathbf{g}$ -orthonormal basis  $\{\mathbf{e}_j\}$ , the set  $\{\mathbf{e}_i \mid \mathbf{g}(\mathbf{e}_i, \mathbf{e}_i) = 0\}$  is a basis of  $\ker \mathbf{g}$ , and therefore  $n_0$  is the nullity of  $\mathbf{g}$ .

**25.19.** Use Equation (25.23) to show that for a  $3 \times 3$  matrix  $\mathbf{A}$ ,

$$\det \mathbf{A} = \frac{1}{3!} [(\operatorname{tr} \mathbf{A})^3 - 3 \operatorname{tr} \mathbf{A} \operatorname{tr}(\mathbf{A}^2) + 2 \operatorname{tr}(\mathbf{A}^3)].$$

**25.20.** Show that a 2-form  $\omega$  is nondegenerate if and only if the determinant of  $(\omega_{ij})$  is nonzero if and only if  $\omega^b$  is an isomorphism.

**25.21.** Let  $\mathcal{V}$  be a finite-dimensional vector space and  $\omega \in \Lambda^2(\mathcal{V}^*)$ . Suppose there exist a pair of vectors  $\mathbf{e}_1, \mathbf{e}'_1 \in \mathcal{V}$  such that  $\omega(\mathbf{e}_1, \mathbf{e}'_1) \neq 0$ . Let  $\mathcal{P}_1$  be the plane spanned by  $\mathbf{e}_1$  and  $\mathbf{e}'_1$ , and  $\mathcal{V}_1$  the  $\omega$ -orthogonal complement of  $\mathcal{P}_1$ . Show that  $\mathcal{V}_1 \cap \mathcal{P}_1 = 0$ , and that  $\mathbf{v} - \omega(\mathbf{v}, \mathbf{e}'_1)\mathbf{e}_1 + \omega(\mathbf{v}, \mathbf{e}_1)\mathbf{e}'_1$  is in  $\mathcal{V}_1$ .

**25.22.** Show that  $\sum_{j=1}^n \epsilon^j \wedge \epsilon^{j+n}$ , in which  $\{\epsilon^j\}_{j=1}^N$  is dual to  $\{\mathbf{e}_i\}_{i=1}^N$ , the canonical basis of  $\mathcal{V}$ , has the same matrix as  $\omega$ .

**25.23.** Suppose that  $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$  are expressed in a canonical basis of  $\mathcal{V}$  with coefficients  $\{x_i, y_i, z_i\}$  and  $\{x'_i, y'_i, z'_i\}$ . Show that

$$\omega(\mathbf{v}, \mathbf{v}') = \sum_{i=1}^n (x_i y'_i - x'_i y_i).$$

**25.24.** Let  $\mathcal{V}$  be a vector space and  $\mathcal{V}^*$  its dual. Define  $\omega \in \Lambda^2(\mathcal{V} \oplus \mathcal{V}^*)$  by

$$\omega(\mathbf{v} + \boldsymbol{\varphi}, \mathbf{v}' + \boldsymbol{\varphi}') \equiv \boldsymbol{\varphi}'(\mathbf{v}) - \boldsymbol{\varphi}(\mathbf{v}')$$

where  $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$  and  $\boldsymbol{\varphi}, \boldsymbol{\varphi}' \in \mathcal{V}^*$ . Show that  $(\mathcal{V} \oplus \mathcal{V}^*, \omega)$  is a symplectic vector space.

**25.25.** By taking successive powers of  $\omega$  show that

$$\omega^k = \sum_{j_1 \dots j_k=1}^n \epsilon^{j_1} \wedge \epsilon^{j_1+n} \wedge \dots \wedge \epsilon^{j_k} \wedge \epsilon^{j_k+n}.$$

Conclude that

$$\omega^n = n!(-1)^{[n/2]} \epsilon^1 \wedge \dots \wedge \epsilon^{2n},$$

where  $[n/2]$  is the largest integer less than or equal to  $n/2$ .

**25.26.** Show that the condition for a matrix  $A$  to be symplectic is  $A^t J A = J$  where  $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  is the representation of  $\omega$  in the canonical basis.

**25.27.** Show that  $Sp(\mathcal{V}, \omega)$  is a subgroup of  $GL(\mathcal{V})$ .

**25.28.** Find the index and the signature for the bilinear form  $\mathbf{g}$  on  $\mathbb{R}^3$  given by  $\mathbf{g}(\mathbf{v}_1, \mathbf{v}_2) = x_1 y_2 + x_2 y_1 - y_1 z_2 - y_2 z_1$ .

**25.29.** In relativistic electromagnetic theory the current  $\mathbf{J}$  and the electromagnetic field tensor  $\mathbf{F}$  are, respectively, a four-vector<sup>6</sup> and an antisymmetric tensor of rank 2. That is,  $\mathbf{J} = J^k \mathbf{e}_k$  and  $\mathbf{F} = F^{ij} \mathbf{e}_i \wedge \mathbf{e}_j$ . Find the components of  $*\mathbf{J}$  and  $*\mathbf{F}$ . Recall that the space of relativity is a 4D Minkowski space.

**25.30.** Show that where there is a sum over an upper index and a lower index, swapping the upper index to a lower index, and vice versa, does not change the sum. In other words,  $A^i B_i = A_i B^i$ .

**25.31.** Show the following vector identities, using the definition of cross products in terms of  $\epsilon_{ijk}$ .

(a)  $\mathbf{A} \times \mathbf{A} = \mathbf{0}$ .

(b)  $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = (\nabla \times \mathbf{A}) \cdot \mathbf{B} - (\nabla \times \mathbf{B}) \cdot \mathbf{A}$ .

(c)  $\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla) \mathbf{A} + \mathbf{A} (\nabla \cdot \mathbf{B}) - (\mathbf{A} \cdot \nabla) \mathbf{B} - \mathbf{B} (\nabla \cdot \mathbf{A})$ .

(d)  $\nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$ .

**25.32.** A vector operator  $\mathbf{V}$  is defined as  $\{\mathbf{V}^1, \mathbf{V}^2, \mathbf{V}^3\}$ , a set of three operators, satisfying the following commutation relations with angular momentum:  $[\mathbf{V}^i, \mathbf{J}^j] = i \epsilon^{ijk} \mathbf{V}_k$ . Show that  $\mathbf{V}^k \mathbf{V}_k$  commutes with all components of angular momentum.

**25.33.** The Pauli spin matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

describe a particle with spin  $\frac{1}{2}$  in nonrelativistic quantum mechanics. Verify that these matrices satisfy

$$[\sigma^i, \sigma^j] \equiv \sigma^i \sigma^j - \sigma^j \sigma^i = 2i \epsilon_k^{ij} \sigma^k, \quad \{\sigma^i, \sigma^j\} \equiv \sigma^i \sigma^j + \sigma^j \sigma^i = 2\delta_j^i 1_2,$$

where  $1_2$  is the unit  $2 \times 2$  matrix. Show also that  $\sigma^i \sigma^j = i \epsilon_k^{ij} \sigma^k + \delta_j^i 1_2$ , and for any two vectors  $\mathbf{a}$  and  $\mathbf{b}$ ,  $(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} 1_2 + i \boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b})$ .

**25.34.** Show that any contravariant tensor of rank two can be written as the sum of a symmetric tensor and an antisymmetric tensor. Can this be generalized to tensors of arbitrary rank?

<sup>6</sup>It turns out to be more natural to consider  $\mathbf{J}$  as a 3-form. However, such a fine distinction is not of any consequence for the present discussion.

## Additional Reading

1. Abraham, R., Marsden, J. and Ratiu, T. *Manifolds, Tensor Analysis, and Applications*, 2nd ed., Springer-Verlag, 1988. A comprehensive textbook on tensors with many examples drawn from physics.
2. Bishop, R. and Goldberg, S. *Tensor Analysis on Manifolds*, Dover, 1980. This little and lucid book is one of the earliest ones on index-free tensor analysis.
3. Flanders, H. *Differential Forms with Applications to Physical Sciences*, Dover, 1989. One of the first books on exterior algebra written for physicists. It has many examples drawn from various areas of physics.

# Analysis of Tensors

Tensor *algebra* deals with lifeless vectors and tensors—objects that do not move, do not change, possess no dynamics. Whenever there is a need for tensors in physics, there is also a need to know the way these tensors change with position and time. Tensors that depend on position and time are called tensor fields and are the subject of this chapter.

In studying the algebra of tensors, we learned that they are generalizations of vectors. Once we have a vector space  $\mathcal{V}$  and its dual space  $\mathcal{V}^*$ , we can take the tensor products of factors of  $\mathcal{V}$  and  $\mathcal{V}^*$  and create tensors of various kinds. Thus, once we know what a vector is, we can make up tensors from it.

In the previous chapter, we did not concern ourselves with what a vector was; we simply assumed that it existed. Because all the vectors considered there were stationary, their mere existence was enough. However, in tensor analysis, where things keep changing from point to point (and over time), the existence of vectors at one point does not guarantee their existence at all points. Therefore, we now have to demand more from vectors than their mere existence. Tied to the concept of vectors is the notion of space, or space-time. Let us consider this first.

## 26.1 Differentiable Manifolds

Space is one of the undefinables in elementary physics. Length and time intervals are concepts that are “God given,” and any definitions of these concepts will be circular. This is true as long as we are confined within a single space. In classical physics, this space is the three-dimensional Euclidean space in which every motion takes place. In special relativity, space is changed to Minkowski space-time. In non-relativistic quantum mechanics, the underlying space is the (infinite-dimensional) Hilbert space, and time is the only dynamical parameter. In the general theory

of relativity, gravitation and space-time are intertwined through the concept of curvature.

Mathematicians have invented a unifying theme that brings the common features of all spaces together. This unifying theme is the theory of differentiable manifolds. A rigorous understanding of differentiable manifolds is beyond the scope of this book. However, a working knowledge of manifold theory is surprisingly simple. Let us begin with a crude definition of a differentiable manifold.

differentiable  
manifold  
provisionally defined

**26.1.1. Definition.** A *differentiable manifold* is a collection of objects called *points* that are connected to each other in a smooth fashion such that the neighborhood of each point looks like the neighborhood of an  $m$ -dimensional (Cartesian) space;  $m$  is called the *dimension* of the manifold.

As is customary in the literature, we use “manifold” to mean “differentiable manifold.”

**26.1.2. Example.** The following are examples of differentiable manifolds.

- (a) The space  $\mathbb{R}^n$  is an  $n$ -dimensional manifold.
- (b) The surface of a sphere is a two-dimensional manifold.
- (c) A torus is a two-dimensional manifold.
- (d) The collection of all  $n \times n$  real matrices whose elements are real functions having derivatives of all orders is an  $n^2$ -dimensional manifold. Here a point is an  $n \times n$  matrix.
- (e) The collection of all rotations in  $\mathbb{R}^3$  is a three-dimensional manifold. (Here a point is a rotation.)
- (f) Any smooth surface in  $\mathbb{R}^3$  is a two-dimensional manifold.
- (g) The unit  $n$ -sphere  $S^n$ , which is the collection of points in  $\mathbb{R}^{n+1}$  satisfying

$$x_1^2 + \dots + x_{n+1}^2 = 1,$$

is a manifold.

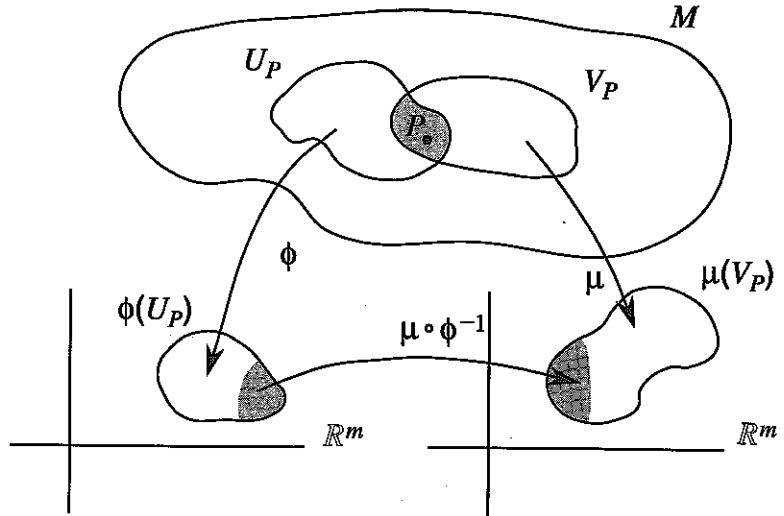
Any surface with sharp kinks, edges, or points cannot be a manifold. Thus, neither a cone nor a finite cylinder is a two-dimensional manifold. However, an infinitely long cylinder is a manifold. ■

coordinate functions  
and charts

Let  $U_P$  denote a neighborhood of  $P$ . When we say that this neighborhood looks like an  $m$ -dimensional Cartesian space, we mean that there exists a bijective map  $\varphi : U_P \rightarrow \mathbb{R}^m$  from a neighborhood  $U_P$  of  $P$  to a neighborhood  $\varphi(U_P)$  of  $\varphi(P)$  in  $\mathbb{R}^m$ , such that as we move the point  $P$  continuously in  $U_P$ , its image moves *continuously* in  $\varphi(U_P)$ . Since  $\varphi(P) \in \mathbb{R}^m$ , we can define functions  $x^i : U_P \rightarrow \mathbb{R}$  such that  $\varphi(P) = (x^1(P), x^2(P), \dots, x^m(P))$ . These functions are called **coordinate functions** of  $\varphi$ . The numbers  $x^i(P)$  are called **coordinates** of  $P$ . The neighborhood  $U_P$  together with its mapping  $\varphi$  form a **chart**, denoted by  $(U_P, \varphi)$ .

$C^\infty$ -related charts  
and atlases

Now let  $(V_P, \mu)$  be another chart at  $P$  with coordinate functions  $\mu(P) = (y^1(P), y^2(P), \dots, y^m(P))$  (see Figure 26.1). It is assumed that the function  $\mu \circ \varphi^{-1} : \varphi(U_P \cap V_P) \rightarrow \mu(U_P \cap V_P)$ , which maps a subset of  $\mathbb{R}^m$  to another subset of  $\mathbb{R}^m$ , possesses derivatives of all orders. Then, we say that the two charts



**Figure 26.1** Two charts  $(U_P, \phi)$  and  $(V_P, \mu)$ , containing  $P$  are mapped into  $\mathbb{R}^m$ . The function  $\mu \circ \phi^{-1}$  is an ordinary function from  $\mathbb{R}^m$  to  $\mathbb{R}^m$ .

$\mu$  and  $\phi$  are  $C^\infty$ -related. Such a relation underlies the concept of smoothness in the definition of a manifold. A collection of charts that cover the manifold and of which each pair is  $C^\infty$ -related is called a  $C^\infty$  atlas.

**26.1.3. Example.** For the two-dimensional unit sphere  $S^2$  we can construct an atlas as follows. Let  $P = (x, y, z)$  be a point in  $S^2$ . Then  $x^2 + y^2 + z^2 = 1$ , or

$$z = \pm\sqrt{1 - x^2 - y^2}.$$

The plus sign corresponds to the upper hemisphere, and the minus sign to the lower hemisphere. Let  $U_3^+$  be the upper hemisphere with the equator removed. Then a chart  $(U_3^+, \phi_3)$  with  $\phi_3 : U_3^+ \rightarrow \mathbb{R}^2$  can be constructed by projecting on the  $xy$ -plane:  $\phi_3(P) = (x, y)$ . Similarly,  $(U_3^-, \mu_3)$  with  $\mu_3 : U_3^- \rightarrow \mathbb{R}^2$  given by  $\mu_3(P) = (x, y)$  is a chart for the lower hemisphere.

Construction of an atlas for the sphere  $S^2$ .

In manifold theory the neighborhoods on which mappings of charts are defined have no boundaries (thus the word “open”). This is because it is more convenient to define limits on boundaryless (open) neighborhoods. Thus, in the above two charts the equator, which is the boundary for both hemispheres, must be excluded. With this exclusion  $U_3^+$  and  $U_3^-$  cannot cover the entire  $S^2$ ; hence, they do not form an atlas. More charts are needed to cover the unit two-sphere. Two such charts are the right and left hemispheres  $U_2^+$  and  $U_2^-$ , for which  $y > 0$  and  $y < 0$ , respectively. However, these two neighborhoods leave two points uncovered, the points  $(1, 0, 0)$  and  $(-1, 0, 0)$ . Again this is because boundaries of the right and left hemispheres must be excluded. Adding the front and back hemispheres  $U_1^\pm$  to the collection covers these two points. Then  $S^2$  is completely covered and we have

an atlas. There is, of course, a lot of overlap among charts. We now show that these overlaps are  $C^\infty$ -related.

As an illustration, we consider the overlap between  $U_3^+$  and  $U_2^+$ . This is the upper-right quarter of the sphere. Let  $(U_3^+, \varphi_3)$  and  $(U_2^+, \varphi_2)$  be charts with

$$\varphi_3(x, y, z) = (x, y), \quad \varphi_2(x, y, z) = (x, z).$$

The inverses are therefore given by

$$\varphi_3^{-1}(x, y) = (x, y, z) = (x, y, \sqrt{1 - x^2 - y^2}),$$

$$\varphi_2^{-1}(x, z) = (x, y, z) = (x, \sqrt{1 - x^2 - z^2}, z),$$

and

$$\varphi_2 \circ \varphi_3^{-1}(x, y) = \varphi_2(x, y, \sqrt{1 - x^2 - y^2}) = (x, \sqrt{1 - x^2 - y^2}).$$

Let us denote  $\varphi_2 \circ \varphi_3^{-1}$  by  $F$ , so that  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is described by two functions, the components of  $F$ :

$$F_1(x, y) = x \quad \text{and} \quad F_2(x, y) = \sqrt{1 - x^2 - y^2}.$$

The first component has derivatives of all orders at all points. The second component has derivatives of all orders at all points except at  $x^2 + y^2 = 1$ , which is excluded from the region of overlap of  $U_3^+$  and  $U_2^+$ , for which  $z$  can never be zero. Thus,  $F$  has derivatives of all orders at all points of its domain of definition.

One can similarly show that all regions of overlap for all charts have this property, i.e., all charts are  $C^\infty$ -related. ■

**26.1.4. Example.** For  $S^2$  of the preceding example, we can find a new atlas in terms of new coordinate functions. Since  $x_1^2 + x_2^2 + x_3^2 = 1$ , we can use spherical coordinates  $\theta = \cos^{-1} x_3$ ,  $\varphi = \tan^{-1}(x_2/x_1)$ . A chart is then given by  $(S^2 - \{1\} - \{-1\}, \mu)$ , where  $\mu(P) = (\theta, \varphi)$  maps a point of  $S^2$  onto a region in  $\mathbb{R}^2$ . This is schematically shown in Figure 26.2. The singletons  $\{1\}$  and  $\{-1\}$  are the north and the south poles, respectively. This chart cannot cover all of  $S^2$ , however, because when  $\theta = 0$  (or  $\pi$ ), the value of the azimuthal angle  $\varphi$  is not determined. In other words,  $\theta = 0$  (or  $\pi$ ) determines *one* point of the sphere (the north pole or the south pole), but its image in  $\mathbb{R}^2$  is the whole range of  $\varphi$  values. Therefore, we must exclude  $\theta = 0$  (or  $\pi$ ) from the chart  $(S^2, \mu)$ . To cover these two points, we need more charts. ■

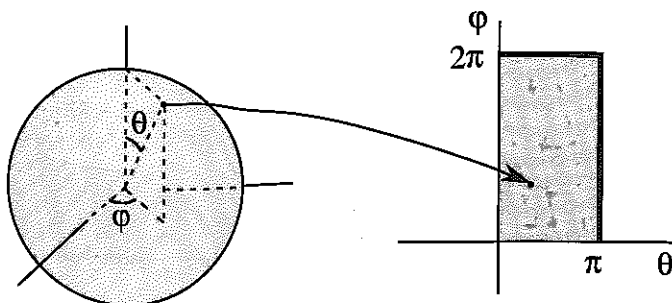
Illustration of  
stereographic  
projection

**26.1.5. Example.** A third atlas for  $S^2$  is the so-called **stereographic projection** shown in Figure 26.3. In such a mapping the image of a point is obtained by drawing a line from the north pole to that point and extending it, if necessary, until it intersects the  $x_1x_2$ -plane. It can be verified that the mapping  $\varphi: S^2 - \{1\} \rightarrow \mathbb{R}^2$  is given by

$$\varphi(x_1, x_2, x_3) = \left( \frac{x_1}{1 - x_3}, \frac{x_2}{1 - x_3} \right).$$

We see that this mapping fails for  $x_3 = 1$ , that is, the north pole. Therefore, the north pole must be excluded (thus, the domain  $S^2 - \{1\}$ ). To cover the north pole we need another stereographic projection—this time from the south pole. Then the two mappings will cover all of  $S^2$ , and it can be shown that the two charts are  $C^\infty$ -related (see Example 26.1.12). ■





**Figure 26.2** A chart mapping points of  $S^2$  into  $\mathbb{R}^2$ . Note that the map is not defined for  $\theta = 0, \pi$ , and therefore at least one more chart is required to cover the whole sphere.

The three foregoing examples illustrate the following fact, which can be shown to hold rigorously:

**26.1.6. Box.** *It is impossible to cover the whole  $S^2$  with just one chart.*

vector spaces are manifolds

**26.1.7. Example.** Let  $\mathcal{V}$  be an  $m$ -dimensional real vector space. Fix any basis  $\{e_i\}$  in  $\mathcal{V}$  with dual basis  $\{\epsilon^i\}$ . Define  $\phi : \mathcal{V} \rightarrow \mathbb{R}^m$  by  $\phi(v) = (\epsilon^1(v), \dots, \epsilon^m(v))$ . Then the reader may verify that  $(\mathcal{V}, \phi)$  is an atlas. Linearity of  $\phi$  ensures that it has derivatives of all orders. This construction shows that  $\mathcal{V}$  is a manifold of dimension  $m$ . ■

product manifold defined

If  $M$  and  $N$  are manifolds of dimensions  $m$  and  $n$ , respectively, we can construct their **product manifold**  $M \times N$ , a manifold of dimension  $m + n$ . A typical chart on  $M \times N$  is obtained from charts on  $M$  and  $N$  as follows. Let  $(U, \varphi)$  be a chart on  $M$  and  $(V, \mu)$  one on  $N$ . Then a chart on  $M \times N$  is  $(U \times V, \varphi \times \mu)$  where

$$\varphi \times \mu(P, Q) = (\varphi(P), \mu(Q)) \in \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n} \quad \text{for } P \in U, Q \in V.$$

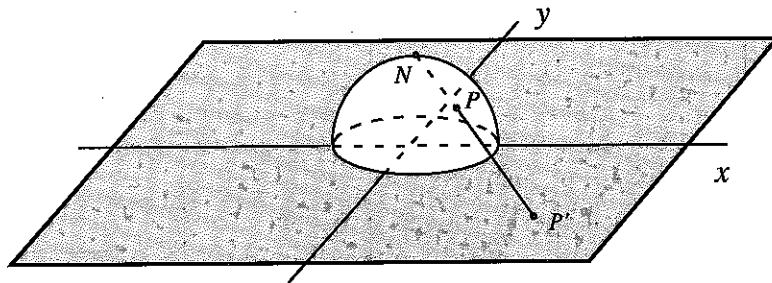
submanifold

**26.1.8. Definition.** Let  $M$  be a manifold. A subset  $N$  of  $M$  is called a **submanifold** of  $M$  if  $N$  is a manifold in its own right.

open submanifolds

A trivial, but important, example of submanifolds is the so-called **open submanifold**. If  $M$  is a manifold and  $U$  is an open subset<sup>1</sup> of  $M$ , then  $U$  inherits a manifold structure from  $M$  by taking any chart  $(U_\alpha, \varphi_\alpha)$  and restricting  $\varphi_\alpha$  to  $U \cap U_\alpha$ . It is clear that  $\dim U = \dim M$ . Having gained familiarity with manifolds, it is now appropriate to consider maps between them that are compatible with their structure.

<sup>1</sup>Recall that an open subset  $U$  is one each of whose points is the center of an open ball lying entirely in  $U$ .



**Figure 26.3** Stereographic projection of  $S^2$  into  $R^2$ . Note that the north pole has no image under this map; another chart is needed to cover the whole sphere.

differentiable maps  
and their coordinate  
expressions

**26.1.9. Definition.** Let  $M$  and  $N$  be manifolds of dimensions  $m$  and  $n$ , respectively. Let  $f : M \rightarrow N$  be a map. We say that  $f$  is  $C^\infty$ , or **differentiable**, if for every chart  $(U, \varphi)$  in  $M$  and every chart  $(V, \mu)$  in  $N$ , the composite map  $\mu \circ f \circ \varphi^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ , called the **coordinate expression for  $f$** , is  $C^\infty$  wherever it is defined.<sup>2</sup>

function as a special  
kind of map

The content of this definition is illustrated in Figure 26.4. A particularly important special case occurs when  $N = \mathbb{R}$ ; then we call  $f$  a (real-valued) **function**. The collection of all  $C^\infty$  functions at a point  $P \in M$  is denoted by  $F^\infty(P)$ : If  $f \in F^\infty(P)$ , then  $f : U_P \rightarrow \mathbb{R}$  is  $C^\infty$  for some neighborhood  $U_P$  of  $P$ . Let  $f : M \rightarrow N$  be a differentiable map. Then  $f$  is automatically continuous. Now let  $V$  be an open subset of  $N$ . The set  $f^{-1}(V)$  is an open subset of  $M$  by Proposition<sup>3</sup> 16.4.6.

**26.1.10. Proposition.** Let  $M$  be an  $m$ -dimensional manifold,  $f : M \rightarrow N$  a differentiable map, and  $V$  an open subset of  $N$ . Then  $f^{-1}(V)$ , the set of points of  $M$  mapped onto  $V$ , is an open  $m$ -dimensional submanifold of  $M$ .

Just as the concept of isomorphism identified all vector spaces, algebras, and groups that were equivalent to one another, it is desirable to introduce a notion that brings together those manifolds that “look alike.”

diffeomorphism and  
local diffeomorphism  
defined

**26.1.11. Definition.** A bijective differentiable map whose inverse is also differentiable is called a **diffeomorphism**. Two manifolds between which a diffeomorphism exists are called **diffeomorphic**. Let  $M$  and  $N$  be manifolds.  $M$  is said to be **diffeomorphic to  $N$  at  $P \in M$**  if there is a neighborhood  $U$  of  $P$  and a diffeomorphism  $f : U \rightarrow f(U)$ . Then  $f$  is called a **local diffeomorphism at  $P$** .

<sup>2</sup>The domain of  $\mu \circ f \circ \varphi^{-1}$  is not all of  $\mathbb{R}^m$ , but only its open subset  $\varphi(U)$ . However, we shall continue to abuse the notation and write  $\mathbb{R}^m$  instead of  $\varphi(U)$ . This way, we do not have to constantly change the domain as  $U$  changes. The domain is always clear from the context.

<sup>3</sup>Although Proposition 16.4.6 was shown for normed linear spaces, it really holds for all “spaces” for which the concept of open set is defined.

diffeomorphisms of a manifold form a group

$n$ -sphere and its stereographic projection

In our discussion of groups, we saw that the set of linear isomorphisms of a vector space  $\mathcal{V}$  onto itself forms a group  $GL(\mathcal{V})$ . The set of diffeomorphisms of a manifold  $M$  onto itself also forms a group, which is denoted by  $\text{Diff}(M)$ .

**26.1.12. Example.** The generalization of a sphere is the unit  $n$ -sphere, which is a subset of  $\mathbb{R}^{n+1}$  defined by

$$S^n = \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \dots + x_{n+1}^2 = 1\}.$$

The stereographic projection defines an atlas for  $S^n$  as follows. For all points of  $S^n$  except  $(0, 0, \dots, 1)$ , the north pole, define the chart  $\varphi_+ : S^n - \{1\} \equiv U^+ \rightarrow \mathbb{R}^n$  by

$$\varphi_+(x_1, \dots, x_{n+1}) = \left( \frac{x_1}{1 - x_{n+1}}, \dots, \frac{x_n}{1 - x_{n+1}} \right) \quad \text{for } (x_1, \dots, x_{n+1}) \in U^+.$$

To include the north pole, consider a second chart  $\varphi_- : S^n - \{-1\} \equiv U^- \rightarrow \mathbb{R}^n$  defined by

$$\varphi_-(x_1, \dots, x_{n+1}) = \left( \frac{x_1}{1 + x_{n+1}}, \dots, \frac{x_n}{1 + x_{n+1}} \right) \quad \text{for } (x_1, \dots, x_{n+1}) \in U^-.$$

Next, let us find the inverses of these maps. We find the inverse of  $\varphi_+$ ; that of  $\varphi_-$  can be found similarly. Let  $\xi_k \equiv x_k/(1 - x_{n+1})$ . Then one can readily show that

$$\sum_{k=1}^n \xi_k^2 = \frac{1 + x_{n+1}}{1 - x_{n+1}} \Rightarrow x_{n+1} = \frac{\sum_{k=1}^n \xi_k^2 + 1}{\sum_{k=1}^n \xi_k^2 - 1}$$

and

$$x_i = \frac{2\xi_i}{1 - \sum_{k=1}^n \xi_k^2} \quad \text{for } i = 1, 2, \dots, n.$$

From the definition of  $\varphi_+$ , we have

$$\begin{aligned} \varphi_+^{-1}(\xi_1, \dots, \xi_n) &= (x_1, \dots, x_n, x_{n+1}) \\ &= \left( \frac{2\xi_1}{1 - \sum_{k=1}^n \xi_k^2}, \dots, \frac{2\xi_n}{1 - \sum_{k=1}^n \xi_k^2}, \frac{\sum_{k=1}^n \xi_k^2 + 1}{\sum_{k=1}^n \xi_k^2 - 1} \right). \end{aligned} \tag{26.1}$$

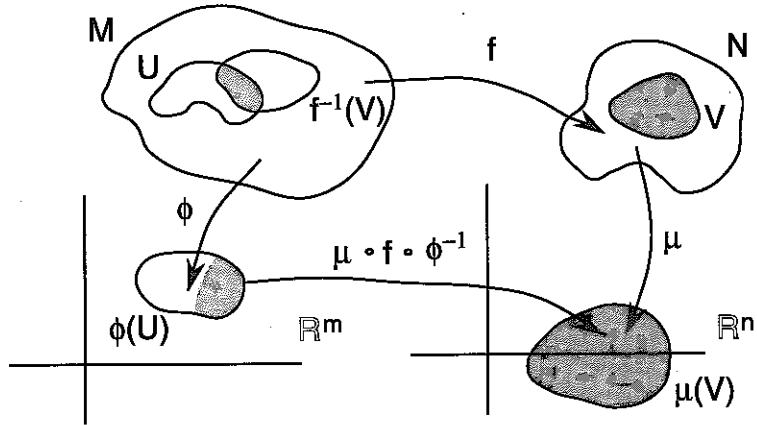
On the overlap of  $U^+$  and  $U^-$ , i.e., on all points of  $S^n$  except the north and the south poles,  $\varphi_- \circ \varphi_+^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  can be calculated by noting that  $\varphi_-$  has the following effect on a typical entry of Equation (26.1):

$$x_j \mapsto \frac{x_j}{1 + x_{n+1}} = \frac{\frac{2\xi_j}{1 - \sum_{k=1}^n \xi_k^2}}{1 + \frac{\sum_{k=1}^n \xi_k^2 + 1}{\sum_{k=1}^n \xi_k^2 - 1}} = -\frac{\xi_j}{\sum_{k=1}^n \xi_k^2}.$$

Therefore,

$$\varphi_- \circ \varphi_+^{-1}(\xi_1, \dots, \xi_n) = \left( -\frac{\xi_1}{\sum_{k=1}^n \xi_k^2}, \dots, -\frac{\xi_n}{\sum_{k=1}^n \xi_k^2} \right).$$

It is clear that  $\varphi_- \circ \varphi_+^{-1}$  has derivatives of all orders except possibly at a point for which  $\xi_i = 0$  for all  $i$ . But this would correspond to  $x_{n+1} = 1$ , which is excluded from the region of overlap. ■



**Figure 26.4** Corresponding to every map  $f : M \rightarrow N$  there exists a coordinate map  $\mu \circ f \circ \phi^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ .

## 26.2 Curves and Tangent Vectors

We noted above that *functions* are special cases of Definition 26.1.9. Another special case occurs when  $M = \mathbb{R}$ . This is important enough to warrant a separate definition.

differentiable curve

**26.2.1. Definition.** A *differentiable curve in the manifold  $M$*  is a  $\mathcal{C}^\infty$  map of an interval of  $\mathbb{R}$  to  $M$ .

initial and final points of a curve

This definition should be familiar from calculus, where  $M = \mathbb{R}^3$  and a curve is given by its *parametric equation*  $(f_1(t), f_2(t), f_3(t))$ , or simply by  $\mathbf{r}(t)$ . The point  $\gamma(a) \in M$  is called the **initial point**, and  $\gamma(b) \in M$  is called the **final point** of the curve  $\gamma$ . A curve is closed if  $\gamma(a) = \gamma(b)$ .

We are now ready to consider what a vector at a point is. All the familiar vectors in classical physics, such as displacement, velocity, momentum, and so forth, are based on the displacement vector. Let us see how we can generalize such a vector so that it is compatible with the concept of a manifold.

In  $\mathbb{R}^2$ , we define the displacement vector from  $P$  to  $Q$  as a directed straight line that starts at  $P$  and ends at  $Q$ . Furthermore, the direction of the vector remains the same if we connect  $P$  to any other final point on the line  $PQ$  located beyond  $Q$ . This is because  $\mathbb{R}^2$  is a flat space, a straight line is well-defined, and there is no ambiguity in the direction of the vector from  $P$  to  $Q$ . Things change, however, if we move to a two-dimensional spherical surface such as the globe. How do we define the straight line from New York to Beijing? There is no satisfactory definition of the word "straight" on a curved surface. Let us say that "straight" means shortest distance. Then our shortest path would lie on a great circle passing through New York and Beijing. Define the "direction" of the trip as the "straight"

arrow, say 1 km in length, connecting our present position to the next point 1 km away. As we move from New York to Beijing, going westward, the tip of the arrow keeps changing direction. Its direction in New York is slightly different from its direction in Chicago. In San Francisco the direction is changed even more, and by the time we reach Beijing, the tip of the arrow will be almost opposite to its original direction.

The reason for such a changing arrow is, of course, the curvature of the manifold. We can minimize this curvature effect if we do not go too far from New York. If we stay close to New York, the surface of the earth appears flat, and we can draw arrows between points. The closer the two points, the better the approximation to flatness. Clearly, the concept of a vector is a *local* concept, and the process of constructing a vector is a *limiting* process.

The limiting process in the globe example entailed the notions of “closeness.” Such a notion requires the concept of distance, which is natural for a globe but not necessary for a general manifold. For most manifolds it is possible to define a metric that gives the “distance” between two points of the manifold. However, the concept of a vector is too general to require such an elaborate structure as a metric. The abstract usefulness of a metric is a result of its real-valuedness: given two points  $P_1$  and  $P_2$ , the distance between them,  $d(P_1, P_2)$ , is a nonnegative real number. Thus, distances between different points can be compared.

We have already defined two concepts for manifolds (more basic than the concept of a metric) that together can replace the concept of a metric in defining a vector as a limit. These are the concepts of (real-valued) functions and curves. Let us see how functions and curves can replace metrics.

Let  $\gamma : [a, b] \rightarrow M$  be a curve in the manifold  $M$ . Let  $P \in M$  be a point of  $M$  that lies on  $\gamma$  such that  $\gamma(c) = P$  for some  $c \in [a, b]$ . Let  $f \in F^\infty(P)$ . Restrict  $f$  to the neighboring points of  $P$  that lie on  $\gamma$ . Then the composite function  $f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}$  is a real-valued function on  $\mathbb{R}$ .

We can compare values of  $f \circ \gamma$  for various real numbers close to  $c$ —as in calculus. If  $u \in [a, b]$  denotes<sup>4</sup> the variable, then  $f \circ \gamma(u) = f(\gamma(u))$  gives the value of  $f \circ \gamma$  at various  $u$ 's. In particular, the difference  $\Delta(f \circ \gamma) \equiv f(\gamma(u)) - f(\gamma(c))$  is a measure of how close the point  $\gamma(u) \in M$  is to  $P$ . Going one step further, we define

$$\left. \frac{d(f \circ \gamma)}{du} \right|_{u=c} = \lim_{u \rightarrow c} \frac{f(\gamma(u)) - f(\gamma(c))}{u - c}, \quad (26.2)$$

the usual derivative of an ordinary function of one variable. However, this derivative depends on  $\gamma$  and on the point  $P$ . The function  $f$  is merely a *test function*. We could choose any other function to test how things change with movement along  $\gamma$ . What is important is not which function we choose, but how the curve  $\gamma$  causes it to change with movement along  $\gamma$  away from  $P$ . This change is determined by

<sup>4</sup>We usually use  $u$  or  $t$  to denote the (real) argument of the map  $\gamma : [a, b] \rightarrow M$ .

the directional derivative along  $\gamma$  at  $P$ , as given by (26.2). A directional derivative determines a tangent which, in turn, suggests a tangent *vector*. That is why the tangent vector at  $P$  along  $\gamma$  is defined to be the directional derivative itself!

The use of derivative as tangent vector may appear strange to the novice, especially physicists encountering it for the first time, but it has been familiar to mathematicians for a long time. It is hard for the beginner to imagine vectors being charged with the responsibility of measuring the rate of change of functions. It takes some mental adjustment to get used to this idea. The following simple illustration may help with establishing the vector-derivative connection.

illustration of the equality of vectors and directional derivatives

**26.2.2. Example.** Let us take the familiar case of a plane and consider the vector  $\mathbf{a} = a_x \hat{e}_x + a_y \hat{e}_y$ . What kind of a directional derivative can correspond to  $\mathbf{a}$ ? First we need a curve  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$  that is somehow associated with  $\mathbf{a}$ . It is not hard to convince oneself that the most natural association is that of vectors to tangents. Thus, we seek a curve whose tangent is (parallel to)  $\mathbf{a}$ . The easiest (but not the only) way is simply to take the straight line along  $\mathbf{a}$ ; that is, let  $\gamma(u) = (a_x u, a_y u)$ . The directional derivative at  $t = 0$  for an arbitrary function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  is given by

$$\left. \frac{d(f \circ \gamma)}{du} \right|_{u=0} = \lim_{u \rightarrow 0} \frac{f(\gamma(u)) - f(\gamma(0))}{u} = \lim_{u \rightarrow 0} \frac{f(a_x u, a_y u) - f(0, 0)}{u}. \tag{26.3}$$

Taylor expansion in two dimensions yields

$$f(a_x u, a_y u) = f(0, 0) + a_x u \left. \frac{\partial f}{\partial x} \right|_{u=0} + a_y u \left. \frac{\partial f}{\partial y} \right|_{u=0} + \dots$$

Substituting in (26.3), we obtain

$$\begin{aligned} \left. \frac{d(f \circ \gamma)}{du} \right|_{u=0} &= \lim_{u \rightarrow 0} \frac{a_x u (\partial f / \partial x)_{u=0} + a_y u (\partial f / \partial y)_{u=0} + \dots}{u} \\ &= a_x \frac{\partial f}{\partial x} + a_y \frac{\partial f}{\partial y} = \left( a_x \frac{\partial}{\partial x} + a_y \frac{\partial}{\partial y} \right) f. \end{aligned}$$

This clearly shows the connection between directional derivatives and vectors. In fact, the correspondences  $\partial/\partial x \leftrightarrow \hat{e}_x$  and  $\partial/\partial y \leftrightarrow \hat{e}_y$  establish this connection very naturally.

Note that the curve  $\gamma$  chosen above is by no means unique. In fact, there are infinitely many curves that have the same tangent at  $t = 0$  and give the same directional derivative. ■

Since vectors are the same as derivatives, we expect them to have the properties shared by derivatives:

tangent vector defined

**26.2.3. Definition.** Let  $M$  be a differentiable manifold. A *tangent vector* at  $P \in M$  is an operator  $\mathbf{t} : F^\infty(P) \rightarrow \mathbb{R}$  such that for every  $f, g \in F^\infty(P)$  and  $\alpha, \beta \in \mathbb{R}$

1.  $\mathbf{t}$  is linear:  $\mathbf{t}(\alpha f + \beta g) = \alpha \mathbf{t}(f) + \beta \mathbf{t}(g)$ ;

2.  $\mathbf{t}$  satisfies the *derivation property*:

$$\mathbf{t}(fg) = g(P)\mathbf{t}(f) + f(P)\mathbf{t}(g).$$

derivation property of tangent vectors

The operator  $t$  is an abstraction of the derivative operator. Note that  $t(f)$ ,  $g(P)$ ,  $f(P)$ , and  $t(g)$  are all real numbers.

tangent space defined  
The reader may easily check that if addition and scalar multiplication of tangent vectors are defined in an obvious way, the set of all tangent vectors at  $P \in M$  becomes a vector space, called the **tangent space** at  $P$  and denoted by  $\mathcal{T}_P(M)$ . If  $U$  is an open subset of  $M$  (therefore, an open submanifold of  $M$ ), then it is clear that

$$\mathcal{T}_P(U) = \mathcal{T}_P(M) \quad \text{for all } P \in U. \tag{26.4}$$

Definition 26.2.3 was motivated by Equations (26.2) and (26.3). Let us go backwards and see if (26.2) is indeed a tangent, that is, if it satisfies the two conditions of Definition 26.2.3.

vectors tangent to a curve

**26.2.4. Proposition.** *Let  $\gamma$  be a  $C^\infty$  curve in  $M$  such that  $\gamma(c) = P$ . Define  $\vec{\gamma}(c)$  by*

$$(\vec{\gamma}(c))(f) \equiv \left. \frac{d}{du} f \circ \gamma \right|_{u=c}$$

for every  $f \in F^\infty(P)$ . Then  $\vec{\gamma}(c)$  is a tangent vector at  $P$  called the vector **tangent to  $\gamma$  at  $c$** .

*Proof.* We have to show that the two conditions of Definition 26.2.3 are satisfied for  $f, g \in F^\infty(P)$  and  $\alpha, \beta \in \mathbb{R}$ . The first condition is trivial. For the second condition, we use the product rule for ordinary differentiation as follows:

$$\begin{aligned} (\vec{\gamma}(c))(fg) &= \left. \frac{d}{du} (fg) \circ \gamma \right|_{u=c} \equiv \left. \frac{d}{du} [(f \circ \gamma)(g \circ \gamma)] \right|_{u=c} \\ &= \left[ \left. \frac{d}{du} (f \circ \gamma) \right|_{u=c} \right] (g \circ \gamma)_{u=c} + (f \circ \gamma)_{u=c} \left[ \left. \frac{d}{du} (g \circ \gamma) \right|_{u=c} \right] \\ &= [(\vec{\gamma}(c))(f)]g(\gamma(c)) + f(\gamma(c))[(\vec{\gamma}(c))(g)] \\ &= [(\vec{\gamma}(c))(f)]g(P) + f(P)[(\vec{\gamma}(c))(g)]. \end{aligned}$$

Note that in going from the first equality to the second, we used the fact that by definition, the product of two functions evaluated at a point is the product of the values of the two functions at that point. □

Let us now consider a special curve and corresponding tangent vector that is of extreme importance in applications. Let  $\varphi = (x^1, x^2, \dots, x^m)$  be a coordinate system at  $P$ , where  $x^i : M \rightarrow \mathbb{R}$  is the  $i$ th coordinate function. Then  $\varphi$  is a bijective  $C^\infty$  mapping from the manifold  $M$  into  $\mathbb{R}^m$ . Its inverse,  $\varphi^{-1} : \mathbb{R}^m \rightarrow M$ , is also a  $C^\infty$  mapping. Now, the  $i$ th coordinate of  $P$  is the real number  $u \equiv x^i(P)$ . Suppose that all coordinates of  $P$  are held fixed except the  $i$ th one, which is allowed to vary with  $u$  describing this variation.

coordinate curve,  
coordinate vector  
field, and coordinate  
frames

**26.2.5. Definition.** Let  $(U_P, \varphi)$  be a chart at  $P \in M$ . Then the curve  $\gamma^i : \mathbb{R} \rightarrow M$ , defined by

$$\gamma^i(u) = \varphi^{-1}(x^1(P), \dots, x^{i-1}(P), u, x^{i+1}(P), \dots, x^m(P))$$

is called the  *$i$ th coordinate curve* through  $P$ . The tangent vector to this curve at  $P$  is denoted by  $\partial_i|_P$  and is called the  *$i$ th coordinate vector field* at  $P$ . The collection of all vector fields at  $P$  is called a *coordinate frame* at  $P$ . The variable  $u$  is arbitrary in the sense that it can be replaced by any (good) function of  $u$ .

Let  $c = x^i(P)$ . Then for  $f \in F^\infty(P)$ , we have

$$\begin{aligned} (\partial_i|_P)f &= (\vec{\gamma}_i(c))(f) = \left. \frac{d}{du} f \circ \gamma^i \right|_{u=c} \\ &= \left. \frac{d}{du} f(\varphi^{-1}(x^1(P), \dots, x^{i-1}(P), u, x^{i+1}(P), \dots, x^m(P))) \right|_{u=c} \\ &\equiv \left. \frac{\partial f}{\partial x^i} \right|_P \Rightarrow \partial_i|_P = \left. \frac{\partial}{\partial x^i} \right|_P, \end{aligned} \tag{26.5}$$

where the last equality is a (natural) *definition* of the partial derivative of  $f$  with respect to the  $i$ th coordinate evaluated at the point  $P$ . This partial derivative is again a  $C^\infty$  function at  $P$ . We therefore have the following:

**26.2.6. Proposition.** The coordinate frame  $\{\partial_i|_P\}_{i=1}^m$  at  $P$  is a set of operators  $\partial_i(P) : F^\infty(P) \rightarrow \mathbb{R}$  given by

$$\begin{aligned} (\partial_i|_P)f &= \left. \frac{\partial f}{\partial x^i} \right|_P \\ &\equiv \left. \frac{d}{du} f(\varphi^{-1}(x^1(P), \dots, x^{i-1}(P), u, x^{i+1}(P), \dots, x^m(P))) \right|_{u=c}. \end{aligned} \tag{26.6}$$

Another common notation for  $\partial f / \partial x^i$  is  $f_{,i}$

**26.2.7. Example.** Pick a point  $P = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$  on the sphere  $S^2$  in a chart  $(U_P, \mu)$  given by  $\mu(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) = (\theta, \varphi)$ . If  $\theta$  is kept constant and  $\varphi$  is allowed to vary over values given by  $u$ , then the coordinate curve associated with  $\varphi$  is given by

$$\gamma_\varphi(u) = \mu^{-1}(\theta, u) = (\sin \theta \cos u, \sin \theta \sin u, \cos \theta).$$

As  $u$  varies,  $\gamma_\varphi(u)$  describes a curve on  $S^2$ . This curve is simply a circle of radius  $\sin \theta$ . The tangent to this curve at any point is  $\partial / \partial \varphi$ , or simply  $\partial_\varphi$ , the derivative with respect to the coordinate  $\varphi$ .

Similarly, the curve  $\gamma_\theta(u)$  describes a great circle on  $S^2$  with tangent  $\partial_\theta \equiv \partial / \partial \theta$ . ■

The vector space  $\mathcal{T}_P(M)$  of all tangents at  $P$  was mentioned earlier. In the case of  $S^2$  this tangent space is simply a plane tangent to the sphere at a point. Also, the two vectors,  $\partial_\theta$  and  $\partial_\varphi$  encountered in Example 26.2.7 are clearly linearly independent. Thus, they form a basis for the tangent plane. This argument can be generalized to any manifold. The following theorem is such a generalization (for a proof, see [Bish 80, pp. 51–53]):



Remember Einstein's summation convention!

**26.2.8. Theorem.** Let  $M$  be an  $m$ -dimensional manifold and  $P \in M$ . Then the set  $\{\partial_i|_P\}_{i=1}^m$  forms a basis of  $\mathcal{T}_P(M)$ . In particular,  $\mathcal{T}_P(M)$  is  $m$ -dimensional. An arbitrary vector,  $\mathbf{t} \in \mathcal{T}_P(M)$ , can be written as

$$\mathbf{t} = \alpha^i \partial_i|_P, \quad \text{where } \alpha^i = \mathbf{t}(x^i).$$

The last statement can be derived by letting both sides operate on  $x^j$  and using Equation (26.6). Let  $M = \mathcal{V}$ , a vector space. Choose a basis  $\{e_i\}$  in  $\mathcal{V}$  with its dual considered as coordinate functions. Then, at every  $\mathbf{v} \in \mathcal{V}$ , there is a natural isomorphism  $\phi : \mathcal{V} \rightarrow \mathcal{T}_{\mathbf{v}}(\mathcal{V})$  mapping a vector  $\mathbf{u} = \alpha^i e_i \in \mathcal{V}$  onto  $\alpha^i \partial_i|_{\mathbf{v}} \in \mathcal{T}_{\mathbf{v}}(\mathcal{V})$ . The reader may verify that this isomorphism is coordinate independent; i.e., if one chooses any other basis of  $\mathcal{V}$  with its corresponding dual, then  $\phi(\mathbf{v})$  will be the same vector as before, expressed in the new coordinate basis. Thus,

**26.2.9. Box.** If  $\mathcal{V}$  is a vector space, then for all  $\mathbf{v} \in \mathcal{V}$ , one can identify  $\mathcal{T}_{\mathbf{v}}(\mathcal{V})$  with  $\mathcal{V}$  itself.

Suppose we have two coordinate systems at  $P$ ,  $\{x^i\}$  with tangents  $\partial_i|_P$  and  $\{y^j\}$  with tangents  $\nabla_j|_P$ . Any  $\mathbf{t} \in \mathcal{T}_P(M)$  can be expressed either in terms of  $\partial_i|_P$  or in terms of  $\nabla_j|_P$ :  $\mathbf{t} = \alpha^i \partial_i|_P = \beta^j \nabla_j|_P$ . We can use this relation to obtain  $\alpha^i$  in terms of  $\beta^j$ : From Theorem 26.2.8, we have

$$\alpha^i = \mathbf{t}(x^i) = (\beta^j \nabla_j|_P)(x^i) \equiv \left[ \beta^j \frac{\partial}{\partial y^j} \Big|_P \right] (x^i) = \beta^j \frac{\partial x^i}{\partial y^j} \Big|_P. \tag{26.7}$$

In particular, if  $\mathbf{t} = \nabla_k|_P$ , then  $\beta^j = \mathbf{t}(y^j) = [\nabla_k|_P](y^j) = \delta_k^j$ , and (26.7) gives  $\alpha^i = \partial x^i / \partial y^k$ . Thus, using Equation (26.5),

$$\frac{\partial}{\partial y^j} \Big|_P = \frac{\partial x^i}{\partial y^j} \frac{\partial}{\partial x^i} \Big|_P. \tag{26.8}$$

For any function  $f \in F^\infty(P)$ , Equation (26.8) yields

$$\left[ \frac{\partial}{\partial y^j} \Big|_P \right] f = \frac{\partial f}{\partial y^j} \Big|_P = \frac{\partial x^i}{\partial y^j} \Big|_P \left[ \frac{\partial}{\partial x^i} \Big|_P \right] f = \frac{\partial x^i}{\partial y^j} \Big|_P \frac{\partial f}{\partial x^i} \Big|_P.$$

This is the chain rule for differentiation.

**26.2.10. Example.** Let us find the coordinate curves and the coordinate frame at  $P = (x, y, z)$  on  $S^2$ . We use the coordinates of Example 26.1.3. In particular, consider  $\varphi_3$ , whose inverse is given by

$$\varphi_3^{-1}(x, y) = (x, y, \sqrt{1 - x^2 - y^2}).$$

The coordinate curve  $\gamma_2(u)$  along  $y$  is obtained by letting  $y$  be a function<sup>5</sup> of  $u$ :

$$\gamma_2(u) = \varphi_3^{-1}(x, h(u)) = (x, h(u), \sqrt{1 - x^2 - h^2(u)}),$$

where  $h(0) = y$  and  $h'(0) = \alpha$ , a constant. To find the coordinate vector field at  $P$ , let  $f \in F^\infty(P)$ , and note that

$$\begin{aligned} \partial_2 f &= \left. \frac{d}{du} f(\gamma_2(u)) \right|_{u=0} = \left. \frac{d}{du} f(x, h(u), \sqrt{1 - x^2 - h^2(u)}) \right|_{u=0} \\ &= \left. \frac{\partial f}{\partial y} \frac{dh}{du} \right|_{u=0} + \left. \frac{\partial f}{\partial z} \left[ \frac{1}{2}(-2h(u)) \frac{dh}{du} \frac{1}{\sqrt{1 - x^2 - h^2(u)}} \right] \right|_{u=0} \\ &= \alpha \left( \frac{\partial f}{\partial y} - \frac{y}{z} \frac{\partial f}{\partial z} \right) = \alpha \left( \frac{\partial}{\partial y} - \frac{y}{z} \frac{\partial}{\partial z} \right) f. \end{aligned}$$

So, choosing the function  $h$  in such a way that  $\alpha = 1$ ,

$$\partial_2 = \partial_y - \frac{y}{z} \partial_z,$$

where  $\partial_y$  and  $\partial_z$  are the coordinate vector fields of  $\mathbb{R}^3$ . The coordinate vector field  $\partial_1$  can be obtained similarly. ■

### 26.3 Differential of a Map

Now that we have constructed tangent spaces and defined bases for them, we are ready to consider the notion of the differential (derivative) of a map between manifolds.

**26.3.1. Definition.** *Let  $M$  and  $N$  be manifolds of dimensions  $m$  and  $n$ , respectively, and let  $\psi : M \rightarrow N$  be a  $C^\infty$  map. Let  $P \in M$ , and let  $Q = \psi(P) \in N$  be the image of  $P$ . Then there is induced a map  $\psi_{*P} : \mathcal{T}_P(M) \rightarrow \mathcal{T}_Q(N)$ , called the **differential of  $\psi$  at  $P$**  and given as follows. Let  $\mathbf{t} \in \mathcal{T}_P(M)$  and  $f \in F^\infty(Q)$ . The action of  $\psi_{*P}(\mathbf{t}) \in \mathcal{T}_Q(N)$  on  $f$  is defined as*

differential of a map  
at a point

$$(\psi_{*P}(\mathbf{t}))(f) \equiv \mathbf{t}(f \circ \psi). \tag{26.9}$$

The reader may check that the differential of a composite map is the composite of the corresponding differentials, i.e.,

$$(\psi \circ \phi)_{*P} = \psi_{*\phi(P)} \circ \phi_{*P}. \tag{26.10}$$

Furthermore, if  $\psi$  is a local diffeomorphism at  $P$ , then  $\psi_{*P}$  is a vector space isomorphism. The inverse of this statement—which is called the **inverse mapping theorem**, and is much harder to prove (see [Abra 88, pp. 116 and 196])—is also true:

inverse mapping theorem

**26.3.2. Theorem.** (inverse mapping theorem) *If  $\psi : M \rightarrow N$  is a map and  $\psi_{*P} : \mathcal{T}_P(M) \rightarrow \mathcal{T}_{\psi(P)}(N)$  is a vector space isomorphism, then  $\psi$  is a local diffeomorphism at  $P$ .*

Let us see how Equation (26.9) looks in terms of coordinate functions. Suppose that  $\{x^i\}_{i=1}^m$  are coordinates at  $P$  and  $\{y^a\}_{a=1}^n$  are coordinates at  $Q = \psi(P)$ . We note that  $y^a \circ \psi$  is a real-valued  $C^\infty$  function on  $M$ . Thus, we may write (with the function expressed in terms of coordinates)

$$y^a \circ \psi \equiv f^a(x^1, \dots, x^m).$$

We also have  $\mathbf{t} = \alpha^i \partial_i|_P$ . Similarly,  $\psi_{*P}(\mathbf{t}) = \beta^a (\partial/\partial y^a)|_Q$  because  $\{\partial/\partial y^a|_Q\}$  form a basis. Theorem 26.2.8 and Definition 26.3.1 now give

$$\begin{aligned} \beta^a &= \psi_{*P}(\mathbf{t})(y^a) = \mathbf{t}(y^a \circ \psi) = \mathbf{t}(f^a) \\ &= [\alpha^i \partial_i|_P](f^a) = \alpha^i \left. \frac{\partial f^a}{\partial x^i} \right|_P \equiv \sum_{i=1}^m \alpha^i \left. \frac{\partial f^a}{\partial x^i} \right|_P. \end{aligned}$$

This can be written in matrix form as

$$\begin{pmatrix} \beta^1 \\ \beta^2 \\ \vdots \\ \beta^n \end{pmatrix} = \begin{pmatrix} \partial f^1/\partial x^1 & \partial f^1/\partial x^2 & \dots & \partial f^1/\partial x^m \\ \partial f^2/\partial x^1 & \partial f^2/\partial x^2 & \dots & \partial f^2/\partial x^m \\ \vdots & \vdots & \ddots & \vdots \\ \partial f^n/\partial x^1 & \partial f^n/\partial x^2 & \dots & \partial f^n/\partial x^m \end{pmatrix} \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^m \end{pmatrix}. \tag{26.11}$$

Jacobian matrix of a differentiable map

The  $n \times m$  matrix is denoted by  $J$  and is called the **Jacobian matrix of  $\psi$**  with respect to the coordinates  $x^i$  and  $y^a$ . On numerous occasions the two manifolds are simply Cartesian spaces, so that  $\psi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ . In such a case,  $f^a$  is naturally written as  $\psi^a$ , and the Jacobian matrix will have elements of the form  $\partial \psi^a / \partial x^i$ .

Differential of a constant map is the zero map.

An important special case of the differential of a map is that of a constant map. Let  $\psi : M \rightarrow \{Q\} \in N$  be such a map; it maps all points of  $M$  onto a single point  $Q$  of  $N$ . For any  $f \in F^\infty(Q)$ , the function  $f \circ \psi \in F^\infty(P)$  is constant for all  $P \in M$ . Let  $\mathbf{t} \in \mathcal{T}_P(M)$  be an arbitrary vector. Then

$$(\psi_{*P}(\mathbf{t}))(f) \equiv \mathbf{t}(f \circ \psi) = 0 \quad \forall f \Rightarrow \psi_{*P}(\mathbf{t}) = 0 \quad \forall \mathbf{t} \tag{26.12}$$

because  $\mathbf{t}(c) = 0$  for any constant  $c$ . So,

**26.3.3. Box.** *If  $\psi : M \rightarrow \{Q\} \in N$  is a constant map, so that it maps the entire manifold  $M$  onto a point  $Q$  of  $N$ , then  $\psi_{*P} : \mathcal{T}_P(M) \rightarrow \mathcal{T}_Q(N)$  is the zero map.*

<sup>5</sup>See the last statement of Definition 26.2.5.

Two other special cases merit closer attention:  $M = \mathbb{R}$  for arbitrary  $N$ , and  $N = \mathbb{R}$  for arbitrary  $M$ . In either case  $\mathcal{T}_c(\mathbb{R})$  is one-dimensional with the basis vector  $(d/du)|_c$ . When  $M = \mathbb{R}$ , the mapping becomes a curve,  $\gamma : \mathbb{R} \rightarrow N$ . The only vector whose image we are interested in is  $\mathbf{t} = (d/du)|_c$ , with  $\gamma(c) = P$ . From (26.9) using Proposition 26.2.4 in the last step, we have

$$\left[ \gamma_{*c} \frac{d}{du} \Big|_c \right] f = \frac{d}{du} f \circ \gamma \Big|_{u=c} = (\tilde{\gamma}(c))(f).$$

This tells us that the differential of a curve at  $c$  is simply its tangent vector at  $\gamma(c)$ . It is common to leave out the constant vector  $(d/du)|_c$ , and write  $\gamma_{*c}$  for the LHS.

components of  
tangents to curves

**26.3.4. Example.** It is useful to have an expression for the components of the tangent to a curve  $\gamma$  at an arbitrary point on it. Since  $\gamma$  maps the real line to  $M$ , with a coordinate patch established on  $M$ , we can write  $\gamma$  as  $\gamma = (\gamma^1, \dots, \gamma^m)$  where  $\gamma^i = x^i \circ \gamma$  are ordinary functions of one variable. Proposition 26.2.4 then yields

$$\begin{aligned} \gamma_{*t} f &= \frac{d}{du} f \circ \gamma \Big|_{u=t} = \frac{d}{du} f(\gamma(u)) \Big|_{u=t} = \frac{d}{du} f(\gamma^1(u), \dots, \gamma^m(u)) \Big|_{u=t} \\ &= \frac{\partial f}{\partial x^i} \frac{d\gamma^i}{du} \Big|_{u=t} \equiv \frac{\partial f}{\partial x^i} \frac{d\gamma^i}{dt} = \dot{\gamma}^i \partial_i f, \end{aligned}$$

or

$$\boxed{\gamma_{*t} \equiv \dot{\gamma}^i \partial_i, \quad \text{where } \dot{\gamma}^i = \frac{d\gamma^i}{dt}.} \quad (26.13)$$

For this reason,  $\gamma_{*t}$  is sometimes denoted by  $\dot{\gamma}$ . ■

When  $N = \mathbb{R}$ , we are dealing with a real-valued function  $f : M \rightarrow \mathbb{R}$ . The differential of  $f$  at  $P$  is  $f_{*P} : \mathcal{T}_P(M) \rightarrow \mathcal{T}_c(\mathbb{R})$ , where  $c = f(P)$ . Since  $\mathcal{T}_c(\mathbb{R})$  is one-dimensional, for a tangent  $\mathbf{t} \in \mathcal{T}_P(M)$ , we have  $f_{*P}(\mathbf{t}) = a(d/du)|_c$ . Let  $g : \mathbb{R} \rightarrow \mathbb{R}$  be an arbitrary function on  $\mathbb{R}$ . Then  $[f_{*P}(\mathbf{t})](g) = a(dg/du)_c$ , or, by definition of the LHS,  $\mathbf{t}(f \circ g) = a(dg/du)_c$ . To find  $a$  we choose the function  $g(u) = u$ , i.e., the identity function; then  $dg/du = 1$  and  $\mathbf{t}(f \circ g) = \mathbf{t}(f) = a$ . We thus obtain  $f_{*P}(\mathbf{t}) = \mathbf{t}(f)(d/du)|_c$ . Since  $\mathcal{T}_c(\mathbb{R})$  is a flat one-dimensional vector space, all vectors are the same, and there is no need to write  $(d/du)|_c$ . Thus, we define the **differential of  $f$** , denoted by  $df \equiv f_*$ , as a map  $df : \mathcal{T}_P(M) \rightarrow \mathbb{R}$  given by

differential of a  
real-valued function

$$df(\mathbf{t}) = \mathbf{t}(f). \quad (26.14)$$

In particular, if  $f$  is the coordinate function  $x^i$  and  $\mathbf{t}$  is the tangent to the  $j$ th coordinate curve  $\partial_j|_P$ , we obtain

$$dx^i|_P(\partial_j|_P) = [\partial_j|_P](x^i) = \frac{\partial x^i}{\partial x^j} \Big|_P = \delta_j^i. \quad (26.15)$$

This shows that

**26.3.5. Box.**  $\{dx^i|_P\}_{i=1}^m$  is dual to the basis  $\{\partial_j|_P\}_{j=1}^m$  of  $\mathcal{T}_P(M)$ .

Remember Einstein's summation convention!

**26.3.6. Example.** Let  $f : M \rightarrow \mathbb{R}$  be a real-valued function on  $M$ . Let  $x^i$  be coordinates at  $P$ . We want to express  $df$  in terms of coordinate functions. For  $\mathbf{t} \in \mathcal{T}_P(M)$  we can write  $\mathbf{t} = \alpha^i \partial_i|_P$  and

$$df(\mathbf{t}) = \mathbf{t}(f) = \alpha^i [\partial_i|_P](f) = \alpha^i \partial_i(f),$$

where in the last step, we suppressed the  $P$ . Theorem 26.2.8 and Equation (26.14) yield  $\alpha^i = \mathbf{t}(x^i) = (dx^i)(\mathbf{t})$ . We thus have

$$df(\mathbf{t}) = \partial_i(f)[(dx^i)(\mathbf{t})] = [\partial_i(f)(dx^i)](\mathbf{t}).$$

Since this is true for all  $\mathbf{t}$ , we get

$$df = \partial_i(f)(dx^i) \equiv \sum_{i=1}^m \partial_i(f)(dx^i) = \sum_{i=1}^m \frac{\partial f}{\partial x^i} dx^i. \tag{26.16}$$

This is the classical formula for the differential of a function  $f$ . If we choose  $y^j$ , the  $j$ th member of a new coordinate system, for  $f$ , we obtain

$$dy^j = \sum_{i=1}^m \frac{\partial y^j}{\partial x^i} dx^i \equiv \frac{\partial y^j}{\partial x^i} dx^i, \tag{26.17}$$

which is the transformation dual to Equation (26.8). ■

The following is a powerful theorem that constructs a submanifold out of a differentiable map (for a proof, see [Warn 83, p. 31]):

**26.3.7. Theorem.** Assume that  $\psi : M \rightarrow N$  is a  $C^\infty$  map, that  $Q$  is a point in the range of  $\psi$ , and that  $\psi_* : \mathcal{T}_P(M) \rightarrow \mathcal{T}_Q(N)$  is surjective for all  $P \in \psi^{-1}(Q)$ . Then  $\psi^{-1}(Q)$  is a submanifold of  $M$  and  $\dim \psi^{-1}(Q) = \dim M - \dim N$ .

Compare this theorem with Proposition 26.1.10. There,  $V$  was an open subset of  $N$ , and since  $f^{-1}(V)$  is open, it is automatically an open submanifold. The difficulty in proving Theorem 26.3.7 lies in the fact that  $\psi^{-1}(Q)$  is closed because  $\{Q\}$ , a single point of  $N$ , is closed.

We can justify the last statement of the theorem as follows. From Equation (26.12), we readily conclude that  $\mathcal{T}_P(\psi^{-1}(Q)) = \ker \psi_{*P}$ . The dimension theorem, applied to  $\psi_{*P} : \mathcal{T}_P(M) \rightarrow \mathcal{T}_Q(N)$ , now gives

$$\dim \mathcal{T}_P(M) = \dim \ker \psi_{*P} + \text{rank } \psi_{*P} \Rightarrow \dim M = \dim \psi^{-1}(Q) + \dim N,$$

where the last equality follows from the surjectivity of  $\psi_{*P}$ .

**26.3.8. Example.** Consider a  $C^\infty$  map  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Let  $c \in \mathbb{R}$  such that the partial derivatives of  $f$  are defined and not all zero for all points of  $f^{-1}(c)$ . Then, according to Equation (26.11), a vector  $\alpha^i \partial_i \in \mathcal{T}_P(\mathbb{R}^n)$  is mapped by  $f_*$  to the vector  $\alpha^i (\partial f / \partial x^i)_{f=c} d/dt$ . Since  $\partial f / \partial x^i$  are not all zero, by properly choosing  $\alpha^i$ , we can make  $\alpha^i (\partial f / \partial x^i)_{f=c} d/dt$  sweep over all real numbers. Therefore,  $f_*$  is surjective, and by Theorem 26.3.7,  $f^{-1}(c)$  is an  $(n - 1)$ -dimensional submanifold of  $\mathbb{R}^n$ . A noteworthy special case is the function defined by

$$f(x^1, x^2, \dots, x^n) = (x^1)^2 + (x^2)^2 + \dots + (x^n)^2$$

and  $c = r^2 > 0$ . Then,  $f^{-1}(c)$ , an  $(n - 1)$ -sphere of radius  $r$ , is a submanifold of  $\mathbb{R}^n$ . ■

## 26.4 Tensor Fields on Manifolds

So far we have studied vector spaces, learned how to construct tensors out of vectors, touched on manifolds (the abstraction of spaces), seen how to construct vectors at a single point in a manifold by the use of the tangent-at-a-curve idea, and even found the dual vectors  $dx^i|_P$  to the coordinate vectors  $\partial_i|_P$  at a point  $P$  of a manifold. We have everything we need to study the analysis of tensors.

### 26.4.1 Vector Fields

We are familiar with the concept of a vector field in 3D: Electric field, magnetic field, gravitational field, velocity field, and so forth are all familiar notions. We now want to generalize the concept so that it is applicable to a general manifold. To begin with, let us consider the following definition.

**26.4.1. Definition.** *The union of all tangent spaces at different points of a manifold  $M$  is denoted by  $T(M)$  and called the **tangent bundle** of  $M$ :*

tangent bundle defined

$$T(M) = \bigcup_{P \in M} \mathcal{T}_P(M)$$

It can be shown ([Bish 80, pp. 158–164]) that  $T(M)$  is a *manifold* of dimension  $2 \dim M$ .

**26.4.2. Definition.** *A **vector field**  $X$  on a subset  $U$  of a manifold  $M$  is a mapping  $X : U \rightarrow T(M)$  such that  $X(P) \equiv X|_P \equiv X_P \in \mathcal{T}_P(M)$ . The set of vector fields on  $M$  is denoted by  $\mathcal{X}(M)$ . Let  $M$  and  $N$  be manifolds and  $F : M \rightarrow N$  a differentiable map. We say that the two vector fields  $X \in \mathcal{X}(M)$  and  $Y \in \mathcal{X}(N)$  are  **$F$ -related** if  $F_*(X_P) = Y_{F(P)}$  for all  $P \in M$ . This is sometimes written simply as  $F_*X = Y$ .*

vector field defined

vector fields related by a map

It is worthwhile to point out that  $F_*X$  is not, in general, a vector field on  $N$ . To be a vector field,  $F_*X$  must be defined at all points of  $N$ . The natural way to define  $F_*X$  at  $Q \in N$  is  $[F_*X(Q)](f) = X(f \circ F)$  where  $P$  is the preimage of  $Q$ , i.e.,

$F(P) = Q$ . But there may not exist any such  $P$  ( $F$  may not be onto), or there may be more than one  $P$  ( $F$  may not be one-to-one) with such property. Therefore, this natural construction does not lead to a vector field on  $N$ . If  $F_*\mathbf{X}$  happens to be a vector field on  $N$ , then it is clearly  $F$ -related to  $\mathbf{X}$ . In terms of the coordinates  $x^i$ , at each point  $P \in M$ ,

$$\mathbf{X}_P \equiv \mathbf{X}|_P = X^i_P \partial_i|_P,$$

where the real numbers  $X^i_P$  are components of  $\mathbf{X}_P$  in the basis  $\{\partial_i|_P\}$ . As  $P$  moves around in  $U$ , the real numbers  $X^i_P$  keep changing. Thus, we can think of  $X^i_P$  as a function of  $P$  and define the real-valued function  $X^i : M \rightarrow \mathbb{R}$  by  $X^i(P) \equiv X^i_P$ . Therefore, the components of a vector field are real-valued functions on  $M$ .

**26.4.3. Example.** Let  $M = \mathbb{R}^3$ . At each point  $P = (x, y, z) \in \mathbb{R}^3$ , let  $(\hat{e}_x, \hat{e}_y, \hat{e}_z)$  be a basis for  $\mathbb{R}^3$ . Let  $\mathcal{V}_P$  be the vector space at  $P$ . Then  $T(\mathbb{R}^3)$  is the collection of all vector spaces  $\mathcal{V}_P$  for all  $P$ .

We can determine the value of an electric field at a point in  $\mathbb{R}^3$  by first specifying the point, as  $P_0 = (x_0, y_0, z_0)$ , for example. This uniquely determines the tangent space  $\mathcal{T}_{P_0}(\mathbb{R}^3)$ . Once we have the vector space, we can ask what the components of the electric field are in that space. These components are given by three numbers:  $E_x(x_0, y_0, z_0)$ ,  $E_y(x_0, y_0, z_0)$ , and  $E_z(x_0, y_0, z_0)$ . The argument is the same for any other vector field.

To specify a “point” in  $T(\mathbb{R}^3)$ , we need three numbers to determine the location in  $\mathbb{R}^3$  and another three numbers to determine the components of a vector field at that point. Thus, a “point” in  $T(\mathbb{R}^3)$  is given by six “coordinates”  $(x, y, z, E_x, E_y, E_z)$ , and  $T(\mathbb{R}^3)$  is a six-dimensional manifold. ■

We know how a tangent vector  $\mathfrak{t}$  at a point  $P \in M$  acts on a function  $f \in F^\infty(P)$  to give a real number  $\mathfrak{t}(f)$ . We can extend this, point by point, for a vector field  $\mathbf{X}$  and define a function  $\mathbf{X}(f)$  by

$$[\mathbf{X}(f)](P) \equiv \mathbf{X}_P(f), \quad P \in U, \tag{26.18}$$

where  $U$  is a subset of  $M$  on which both  $\mathbf{X}$  and  $f$  are defined. The RHS is well-defined because we know how  $\mathbf{X}_P$ , the vector at  $P$ , acts on functions at  $P$  to give the real number  $[\mathbf{X}_P](f)$ . On the LHS, we have  $\mathbf{X}(f)$ , which maps the point  $P$  onto a real number. Thus,  $\mathbf{X}(f)$  is indeed a real-valued function on  $M$ . We can therefore define vector fields directly as operators on  $C^\infty$  functions satisfying

$$\mathbf{X}(\alpha f + \beta g) = \alpha \mathbf{X}(f) + \beta \mathbf{X}(g),$$

$$\mathbf{X}(fg) = [\mathbf{X}(f)]g + [f]\mathbf{X}(g).$$

$C^\infty$  vector fields

A prototypical vector field is the coordinate vector field  $\partial_i$ . In general,  $\mathbf{X}(f)$  is not a  $C^\infty$  function even if  $f$  is. A vector field that produces a  $C^\infty$  function  $\mathbf{X}(f)$  for every  $C^\infty$  function  $f$  is called a  $C^\infty$  vector field. Such a vector field has components that are  $C^\infty$  functions on  $M$ .

The set of tangent vectors  $\mathcal{T}_P(M)$  at a point  $P \in M$  form an  $m$ -dimensional vector space. The set of vector fields  $\mathcal{X}(M)$ —which yield a vector at every point

of the manifold—also constitutes a vector space. However, this vector space is (uncountably) infinite-dimensional. A property of  $\mathcal{X}(M)$  that is absent in  $\mathcal{T}_P(M)$  is composition.<sup>6</sup> This suggests the possibility of defining a “product” on  $\mathcal{X}(M)$  to turn it into an algebra. Let  $\mathbf{X}$  and  $\mathbf{Y}$  be vector fields. For  $\mathbf{X} \circ \mathbf{Y}$  to be a vector field, it has to satisfy the derivation property. But

$$\begin{aligned}\mathbf{X} \circ \mathbf{Y}(fg) &= \mathbf{X}(\mathbf{Y}(fg)) = \mathbf{X}(\mathbf{Y}(f)g + f\mathbf{Y}(g)) \\ &= (\mathbf{X}(\mathbf{Y}(f)))g + \mathbf{Y}(f)\mathbf{X}(g) + \mathbf{X}(f)\mathbf{Y}(g) + f(\mathbf{X}(\mathbf{Y}(g))) \\ &\neq (\mathbf{X} \circ \mathbf{Y}(f))g + f(\mathbf{X} \circ \mathbf{Y}(g)).\end{aligned}$$

The set of vector fields form a Lie algebra. However, the reader may verify that  $\mathbf{X} \circ \mathbf{Y} - \mathbf{Y} \circ \mathbf{X}$  does indeed satisfy the derivation property. Therefore, by defining the binary operation  $\mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathcal{X}(M)$  as

$$[\mathbf{X}, \mathbf{Y}] \equiv \mathbf{X} \circ \mathbf{Y} - \mathbf{Y} \circ \mathbf{X},$$

$\mathcal{X}(M)$  becomes an algebra, called the **Lie algebra** of vector fields of  $M$ . The binary operation is called the **Lie bracket**. Although it was not mentioned at the time, we have encountered another example of a Lie algebra in Chapter 2, namely  $\mathcal{L}(\mathcal{V})$  under the binary operation of the commutation relation. Lie brackets have the following two properties:

$$[\mathbf{X}, \mathbf{Y}] = -[\mathbf{Y}, \mathbf{X}],$$

$$[[\mathbf{X}, \mathbf{Y}], \mathbf{Z}] + [[\mathbf{Z}, \mathbf{X}], \mathbf{Y}] + [[\mathbf{Y}, \mathbf{Z}], \mathbf{X}] = 0.$$

Jacobi identity These two relations are the defining properties of all Lie algebras. The last relation is called the **Jacobi identity**.  $\mathcal{X}(M)$  with Lie brackets is an example of an infinite-dimensional Lie algebra;  $\mathcal{L}(\mathcal{V})$  with commutators is an example of a finite-dimensional Lie algebra.

We shall have occasion to use the following theorem in our treatment of Lie groups and algebras in the next chapter:

**26.4.4. Theorem.** *Let  $M$  and  $N$  be manifolds and  $F : M \rightarrow N$  a differentiable map. Assume that  $\mathbf{X}_i \in \mathcal{X}(M)$  is  $F$ -related to  $\mathbf{Y}_i \in \mathcal{X}(N)$  for  $i = 1, 2$ . Then  $[\mathbf{X}_1, \mathbf{X}_2]$  is  $F$ -related to  $[\mathbf{Y}_1, \mathbf{Y}_2]$ , i.e.,*

$$F_*[\mathbf{X}_1, \mathbf{X}_2] = [F_*\mathbf{X}_1, F_*\mathbf{X}_2].$$

*Proof.* Let  $f$  be an arbitrary function on  $N$ . Then

$$\begin{aligned}(F_*[\mathbf{X}_1, \mathbf{X}_2])f &\equiv [\mathbf{X}_1, \mathbf{X}_2](f \circ F) = \mathbf{X}_1(\mathbf{X}_2(f \circ F)) - \mathbf{X}_2(\mathbf{X}_1(f \circ F)) \\ &= \mathbf{X}_1([F_*\mathbf{X}_2(f)] \circ F) - \mathbf{X}_2([F_*\mathbf{X}_1(f)] \circ F) \\ &= F_*\mathbf{X}_1(F_*\mathbf{X}_2(f)) - F_*\mathbf{X}_2(F_*\mathbf{X}_1(f)) \\ &= [F_*\mathbf{X}_1, F_*\mathbf{X}_2]f,\end{aligned}$$

<sup>6</sup>Recall that a typical element of  $\mathcal{T}_P(M)$  is a map  $t : F^\infty(P) \rightarrow \mathbb{R}$  for which composition is meaningless.



where we used Equation (26.9) in the first, second, and third lines, and the result of Problem 26.8 in the second line.  $\square$

It is convenient to visualize vector fields as streamlines. In fact, most of the terminology used in three-dimensional vector analysis, such as flux, divergence, and curl, have their origins in the flow of fluids and the associated velocity vector fields. The streamlines are obtained—in nonturbulent flow—by starting at one point and drawing a curve whose tangent at all points is the velocity vector field. For a smooth flow this curve is unique. There is an exact analogy in manifold theory.

integral curve of a  
vector field

**26.4.5. Definition.** Let  $\mathbf{X} \in \mathcal{X}(M)$  be defined on an open subset  $U$  of  $M$ . An **integral curve of  $\mathbf{X}$**  is a curve  $\gamma$  whose range lies in  $U$  and for every  $t$  in the domain of  $\gamma$ , the vector tangent to  $\gamma$  satisfies  $\gamma_{*t} = \mathbf{X}(\gamma(t))$ . If  $\gamma(0) = P$ , we say that  $\gamma$  **starts at  $P$** .

Let us choose a coordinate system on  $M$ . Then  $\mathbf{X} \equiv X^i \partial_i$ , where  $X^i$  are  $\mathcal{C}^\infty$  functions on  $M$ , and, by 26.13,  $\gamma_{*t} = \dot{\gamma}^i \partial_i$ . The equation for the integral curve of  $\mathbf{X}$  will therefore become

$$\dot{\gamma}^i \partial_i = X^i(\gamma(t)) \partial_i, \quad \text{or} \quad \frac{d\gamma^i}{dt} = X^i(\gamma^1(t), \dots, \gamma^m(t)), \quad i = 1, 2, \dots, m.$$

Since  $\gamma^i$  are simply coordinates of points on  $M$ , we rewrite the equation above as

$$\frac{dx^i}{dt} = X^i(x^1(t), \dots, x^m(t)), \quad i = 1, 2, \dots, m. \tag{26.19}$$

This is a system of first-order differential equations that has a unique (local) solution once the initial value  $\gamma(0)$  of the curve, i.e., the coordinates of the starting point  $P$ , is given. The precise statement for existence and uniqueness of integral curves is contained in the following theorem.

**26.4.6. Theorem.** Let  $\mathbf{X}$  be a  $\mathcal{C}^\infty$  vector field defined on an open subset  $U$  of  $M$ . Suppose  $P \in U$ , and  $c \in \mathbb{R}$ . Then there is a positive number  $\epsilon$  and a unique integral curve  $\gamma$  of  $\mathbf{X}$  defined on  $|t - c| \leq \epsilon$  such that  $\gamma(c) = P$ .

**26.4.7. Example.** EXAMPLES OF INTEGRAL CURVES

- (a) Let  $M = \mathbb{R}$  with coordinate function  $x$ . The vector field  $\mathbf{X} = x \partial_x$  has an integral curve with initial point  $x_0$  given by the DE  $dx/dt = x(t)$ , which has the solution  $x(t) = e^t x_0$ .
- (b) Let  $M = \mathbb{R}^n$  with coordinate functions  $x^i$ . The vector field  $\mathbf{X} = \sum a^i \partial_i$  has an integral curve, with initial point  $\mathbf{r}_0$ , given by the system of DEs  $dx^i/dt = a^i$ , which has the solution  $x^i(t) = a^i t + x_0^i$ , or  $\mathbf{r} = \mathbf{a}t + \mathbf{r}_0$ . The curve is therefore a straight line parallel to  $\mathbf{a}$  going through  $\mathbf{r}_0$ .
- (c) Let  $M = \mathbb{R}^n$  with coordinate functions  $x^i$ . Consider the vector field

$$\mathbf{X} = \sum_{i,j=1}^n a_j^i x^j \partial_i.$$

The integral curve of this vector field, with initial point  $\mathbf{r}_0$ , is given by the system of DEs  $dx^i/dt = \sum_{j=1}^n a_j^i x^j$ , which can be written in vector form as  $d\mathbf{r}/dt = \mathbf{A}\mathbf{r}$  where  $\mathbf{A}$  is a constant matrix. By differentiating this equation several times, one can convince oneself that  $d^k \mathbf{r}/dt^k = \mathbf{A}^k \mathbf{r}$ . The Taylor expansion of  $\mathbf{r}(t)$  then yields

$$\mathbf{r}(t) = \sum_{k=0}^{\infty} \frac{1}{k!} \left. \frac{d^k \mathbf{r}}{dt^k} \right|_{t=0} t^k = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbf{A}^k \mathbf{r}_0 = e^{t\mathbf{A}} \mathbf{r}_0.$$

(d) Let  $M = \mathbb{R}^2$  with coordinate  $x, y$ . The reader may verify that the vector field  $\mathbf{X} = -y\partial_x + x\partial_y$  has an integral curve through  $(x_0, y_0)$  given by

$$x = x_0 \cos t - y_0 \sin t,$$

$$y = x_0 \sin t + y_0 \cos t,$$

i.e., a circle centered at the origin passing through  $(x_0, y_0)$ . ■

Going back to the velocity vector field analogy, we can think of integral curves as the path of particles flowing with the fluid. If we think of the entire fluid as a manifold  $M$ , the flow of particles can be thought of as a transformation of  $M$ . To be precise, let  $M$  be an arbitrary manifold, and  $\mathbf{X} \in \mathcal{X}(M)$ . At each point  $P$  of  $M$ , there is a unique local integral curve  $\gamma_P$  of  $\mathbf{X}$  starting at  $P$  defined on an open subset  $U$  of  $M$ . The map  $F_t : U \rightarrow M$  defined by  $F_t(P) = \gamma_P(t)$  is a (local) transformation of  $M$ . The collection of such maps with different  $t$ 's is called the **flow** of the vector field  $\mathbf{X}$ . The uniqueness of the integral curve  $\gamma_P$  implies that  $F_t$  is a local diffeomorphism. In fact, the collection of maps  $\{F_t\}_{t \in \mathbb{R}}$  forms a (local) **one-parameter group of transformations** in the sense that

$$F_t \circ F_s = F_{t+s}, \quad F_0 = \text{id}, \quad (F_t)^{-1} = F_{-t}. \quad (26.20)$$

One has to keep in mind that  $F_t$  at a point  $P \in M$  is, in general, defined only *locally* in  $t$ , i.e., only for  $t$  in some open interval that depends on  $P$ . For some special, but important, cases this interval can be taken to be the entire  $\mathbb{R}$  for all  $P$ , in which case we speak of a **global one-parameter group of transformations**, and  $\mathbf{X}$  is called a **complete vector field** on  $M$ .

The symbol  $F_t$  used for the flow of the vector field  $\mathbf{X}$  does not contain its connection to  $\mathbf{X}$ . In order to make this connection, it is common to define

$$F_t \equiv \exp(t\mathbf{X}). \quad (26.21)$$

This definition, with no significance attached to "exp" at this point, converts Equation (26.20) into

$$\begin{aligned} \exp(t\mathbf{X}) \circ \exp(s\mathbf{X}) &= \exp[(t+s)\mathbf{X}], \\ \exp(0\mathbf{X}) &= \text{id}, \\ [\exp(t\mathbf{X})]^{-1} &= \exp(-t\mathbf{X}), \end{aligned} \quad (26.22)$$

flow of a vector field  
one-parameter group  
of transformations

Global 1-parameter  
group of  
transformations;  
complete vector  
fields

which notationally justifies the use of “exp.” We shall see in our discussion of Lie groups that this choice of notation is not accidental.

Using this notation, we can write

$$\mathbf{X}_P(f) \equiv \left. \frac{d}{dt} f \circ F_t(P) \right|_{t=0} = \left. \frac{d}{dt} f \circ \exp(t\mathbf{X}) \right|_{t=0}.$$

One usually leaves out the function  $f$  and writes

$$\mathbf{X}_P = \left. \frac{d}{dt} \exp(t\mathbf{X}) \right|_{t=0}, \tag{26.23}$$

where it is understood that the LHS acts on some  $f$  that must compose on the RHS to the left of the exponential. Similarly, we have

$$\begin{aligned} (F_*\mathbf{X})_{F(P)} &= \left. \frac{d}{dt} F(\exp t\mathbf{X}) \right|_{t=0}, \\ G_{*F(P)} \left( \underbrace{\left. \frac{d}{dt} F(\exp t\mathbf{X}) \right|_{t=0}}_{=F_*(\mathbf{X})} \right) &= \left. \frac{d}{dt} G \circ F(\exp t\mathbf{X}) \right|_{t=0}, \end{aligned} \tag{26.24}$$

where  $F : M \rightarrow N$  and  $G : N \rightarrow K$  are maps between manifolds.

**26.4.8. Example.** In this example, we derive a useful formula that gives the value of a function at a neighboring point of  $P \in M$  located on the integral curve of  $\mathbf{X} \in \mathcal{X}(M)$  going through  $P$ . We first note that since  $\mathbf{X}_P$  is tangent to  $\gamma_P$  at  $P = \gamma(0)$ , by Proposition 26.2.4 we have

$$\mathbf{X}_P(f) = \left. \frac{d}{dt} f(\gamma_P(t)) \right|_{t=0} = \left. \frac{d}{dt} f(F_t(P)) \right|_{t=0}.$$

Next we use the definition of derivative and the fact that  $F_0(P) = P$  to write

$$\lim_{t \rightarrow 0} \frac{1}{t} [f(F_t(P)) - f(P)] = \mathbf{X}_P(f).$$

Now, if we assume that  $t$  is very small, we have

$$f(F_t(P)) = f(P) + t\mathbf{X}_P(f) + \dots, \tag{26.25}$$

which is a Taylor series with only the first two terms kept. ■

### 26.4.2 Tensor Fields

We have defined vector spaces  $\mathcal{T}_P(M)$  at each point of  $M$ . We have also constructed coordinate bases,  $\{\partial_i|_P\}_{i=1}^m$ , for these vector spaces. At the end of Section 26.2, we showed that the differentials  $\{dx^i|_P\}_{i=1}^m$  form a basis that is dual to  $\{\partial_i|_P\}_{i=1}^m$ . Let us concentrate on this dual space, which we will denote by  $\mathcal{T}_P^*(M)$ .

Taking the union of all  $\mathcal{T}_P^*(M)$  at all points of  $M$ , we obtain the cotangent

cotangent bundle of a manifold

bundle of  $M$ :

$$T^*(M) = \bigcup_{P \in M} \mathcal{T}_P^*(M). \tag{26.26}$$

This is the dual space of  $T(M)$  at each point of  $M$ . We can now define the analogue of the vector field for the cotangent bundle.

differential one-form

**26.4.9. Definition.** A *differential one-form*  $\theta$  on a subset  $U$  of a manifold  $M$  is a mapping  $\theta : U \rightarrow T^*(M)$  such that  $\theta(P) \equiv \theta_P \in \mathcal{T}_P^*(M)$ . The collection of all one-forms on  $M$  is denoted by  $\mathcal{X}^*(M)$ .

If  $\theta$  is a one-form and  $\mathbf{X}$  is a vector field on  $M$ , then  $\theta(\mathbf{X})$  is a real-valued function on  $M$  defined naturally by  $[\theta(\mathbf{X})](P) \equiv (\theta_P)(\mathbf{X}_P)$ . The first factor on the RHS is a linear functional at  $P$ , and the second factor is a vector at  $P$ . So, the pairing of the two factors produces a real number. A prototypical one-form is the coordinate differential,  $dx^i$ .

pullback of a differentiable map

Associated with a differentiable map  $\psi : M \rightarrow N$ , we defined a differential  $\psi_*$  that mapped a tangent space of  $M$  to a tangent space of  $N$ . The dual of  $\psi_*$  (Definition 1.3.17) is denoted by  $\psi^*$  and is called the **pullback** of  $\psi$ . It takes a one-form on  $N$  to a one-form on  $M$ . In complete analogy to the case of vector fields,  $\theta$  can be written in terms of the basis  $\{dx^i\}$ :  $\theta = \theta_i dx^i$ . Here  $\theta_i$ , the components of  $\theta$ , are real-valued functions on  $M$ .

With the vector spaces  $\mathcal{T}_P(M)$  and  $\mathcal{T}_P^*(M)$  at our disposal, we can construct various kinds of tensors at each point  $P$ . The union of all these tensors is a manifold, and a tensor field can be defined as usual. Thus, we have the following definition.

bundle of tensors and tensor fields

**26.4.10. Definition.** Let  $\mathcal{T}_P(M)$  and  $\mathcal{T}_P^*(M)$  be the tangent and cotangent spaces at  $P \in M$ . Then the set of tensors of type  $(r, s)$  on  $\mathcal{T}_P(M)$  is denoted by  $\mathcal{T}_{s,P}^r(M)$ . The *bundle of tensors* of type  $(r, s)$  over  $M$ , denoted by  $T_s^r(M)$ , is

$$T_s^r(M) = \bigcup_{P \in M} \mathcal{T}_{s,P}^r(M).$$

A *tensor field*  $\mathbf{T}$  of type  $(r, s)$  over a subset  $U$  of  $M$  is a mapping  $\mathbf{T} : U \rightarrow T_s^r(M)$  such that  $\mathbf{T}(P) \equiv \mathbf{T}_P \equiv \mathbf{T}|_P \in \mathcal{T}_{s,P}^r(M)$ .

In particular,  $T_0^0(M)$  is the set of real-valued functions on  $M$ ,  $T_0^1(M) = T(M)$ , and  $T_1^0(M) = T^*(M)$ . Furthermore, since  $\mathbf{T}$  is a multilinear map, the parentheses are normally reserved for vectors and their duals, and as indicated in Definition 26.4.10, the value of  $\mathbf{T}$  at  $P \in M$  is written as  $\mathbf{T}_P$  or  $\mathbf{T}|_P$ . The reader may check that the map

$$\mathbf{T} : \underbrace{\mathcal{X}^*(M) \times \cdots \times \mathcal{X}^*(M)}_{r \text{ times}} \times \underbrace{\mathcal{X}(M) \times \cdots \times \mathcal{X}(M)}_{s \text{ times}} \rightarrow T_0^0(M)$$

defined by

$$[\mathbf{T}(\omega^1, \dots, \omega^r, \mathbf{v}_1, \dots, \mathbf{v}_s)](P) = \mathbf{T}_P(\omega^1|_P, \dots, \omega^r|_P, \mathbf{v}_1|_P, \dots, \mathbf{v}_s|_P)$$

A crucial property of tensors has the property that

$$\begin{aligned} \mathbf{T}(\dots, f\omega^j + g\theta^j, \dots) &= f\mathbf{T}(\dots, \omega^j, \dots) + g\mathbf{T}(\dots, \theta^j, \dots), \\ \mathbf{T}(\dots, f\mathbf{v}_k + g\mathbf{u}_k, \dots) &= f\mathbf{T}(\dots, \mathbf{v}_k, \dots) + g\mathbf{T}(\dots, \mathbf{u}_k, \dots) \end{aligned} \tag{26.27}$$

for any two functions  $f$  and  $g$  on  $M$ . Thus,<sup>7</sup>

**26.4.11. Box.** *A tensor is linear in vector fields and 1-forms, even when the coefficients of linear expansion are functions.*

The components of  $\mathbf{T}$  with respect to coordinates  $x^i$  are the  $m^{r+s}$  real-valued functions

$$T_{j_1 j_2 \dots j_s}^{i_1 i_2 \dots i_r} = \mathbf{T}(dx^{i_1}, dx^{i_2}, \dots, dx^{i_r}, \partial_{j_1}, \partial_{j_2}, \dots, \partial_{j_s}).$$

If tensor fields are to be of any use, we must be able to differentiate them. We shall consider three types of derivatives with different applications. We study one of them here, another in the next section, and the third in Chapter 28.

Derivatives can be defined only for objects that can be added (really, subtracted). For functions of a single (real or complex) variable, this is done almost subconsciously: We take the difference between the values of the function at two nearby points and divide by the length of the interval between the two points. We extended this definition to operators in Chapter 2 with practically no change. For functions of more than one variable, one chooses a direction (a vector) and considers change in the function along that direction. This leads to the concept of *directional derivative*, or partial derivative when the vector happens to be along one of the axes.

In all the above cases, the objects being differentiated reside in the *same* space:  $f(t)$  and  $f(t + \Delta t)$  are both real (complex) numbers;  $\mathbf{H}(t)$  and  $\mathbf{H}(t + \Delta t)$  both belong to  $\mathcal{L}(\mathcal{V})$ . When we try to define derivatives of tensor fields, however, we run immediately into trouble:  $\mathbf{T}_P$  and  $\mathbf{T}_{P'}$  cannot be compared because they belong to two different spaces, one to  $\mathcal{J}_{s,P}^r(M)$  and the other to  $\mathcal{J}_{s,P'}^r(M)$ . To make comparisons, we need first to establish a “connection” between the two spaces. This connection has to be a vector space isomorphism so that there is one and only one vector in the second space that is to be compared with a given vector in the first space. The problem is that there are infinitely many isomorphisms between any given two vector spaces. No “natural” isomorphism exists between  $\mathcal{J}_{s,P}^r(M)$  and  $\mathcal{J}_{s,P'}^r(M)$ ; thus the diversity of tensor “derivatives!” We narrow down this diversity

difficulty associated with differentiating tensors

<sup>7</sup>In mathematical jargon,  $\mathcal{X}(M)$  and  $\mathcal{X}^*(M)$  are called **modules** over the (ring of) real-valued functions on  $M$ . Rings are a generalization of the real numbers (field of real numbers) whose elements have all the properties of a field except that they may have no inverse. A module over a field is a vector space.

by choosing a specific vector at  $\mathcal{T}_{s,P}^r(M)$  and seeking a natural way of defining the derivative along that vector by associating a “natural” isomorphism corresponding to the vector. There are a few methods of doing this. We describe one of them here. First, let us see what happens to tensor fields under a diffeomorphism of  $M$  onto itself. Let  $F : M \rightarrow M$  be such a diffeomorphism. The differential  $F_{*P}$  of this diffeomorphism is an isomorphism of  $\mathcal{T}_P(M)$  and  $\mathcal{T}_{F(P)}(M)$ . This isomorphism induces an isomorphism of the vector spaces  $\mathcal{T}_{s,P}^r(M)$  and  $\mathcal{T}_{s,F(P)}^r(M)$ —also denoted by  $F_{*P}$ —by Equation (25.7). Let us denote by  $F_*$  a map of  $T(M)$  onto  $T(M)$  whose restriction to  $\mathcal{T}_P(M)$  is  $F_{*P}$ . If  $\mathbf{T}$  is a tensor field on  $M$ , then  $F_*(\mathbf{T})$  is also a tensor field, whose value at  $F(Q)$  is obtained by letting  $F_{*Q}$  act on  $\mathbf{T}(Q)$ :

$$[F_*(\mathbf{T})](F(Q)) = F_{*Q}(\mathbf{T}(Q)),$$

or, letting  $P = F(Q)$  or  $Q = F^{-1}(P)$ ,

$$[F_*(\mathbf{T})](P) = F_{*F^{-1}(P)}(\mathbf{T}(F^{-1}(P))). \tag{26.28}$$

Now, let  $\mathbf{X}$  be a vector field and  $P \in M$ . The flow of  $\mathbf{X}$  at  $P$  defines a local diffeomorphism  $F_t : U \rightarrow F_t(U)$  with  $P \in U$ . The differential  $F_{t*}$  of this diffeomorphism is an isomorphism of  $\mathcal{T}_P(M)$  and  $\mathcal{T}_{F_t(P)}(M)$ . As discussed above, this isomorphism induces an isomorphism of the vector space  $\mathcal{T}_s^r(M)$  onto itself. The derivative we are after is defined by comparing a tensor field evaluated at  $P$  with the image of the same tensor field under the isomorphism  $F_{t*}^{-1}$ . The following definition makes this procedure more precise.

**26.4.12. Definition.** Let  $P \in M$ ,  $\mathbf{X} \in \mathcal{X}(M)$ , and  $F_t$  the flow of  $\mathbf{X}$  defined in a neighborhood of  $P$ . The Lie derivative of a tensor field  $\mathbf{T}$  at  $P$  with respect to  $\mathbf{X}$  is denoted by  $(L_{\mathbf{X}}\mathbf{T})_P$  and defined by

Lie derivative of tensor fields with respect of a vector field

$$(L_{\mathbf{X}}\mathbf{T})_P = \lim_{t \rightarrow 0} \frac{1}{t} [F_{t*}^{-1}\mathbf{T}_{F_t(P)} - \mathbf{T}_P] \equiv \left. \frac{d}{dt} F_{t*}^{-1}\mathbf{T}_{F_t(P)} \right|_{t=0}. \tag{26.29}$$

Let us calculate the derivative in Equation (26.29) at an arbitrary value of  $t$ . For this purpose, let  $Q \equiv F_t(P)$ . Then

$$\begin{aligned} \frac{d}{dt} F_{t*}^{-1}\mathbf{T}_{F_t(P)} &\equiv \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [F_{(t+\Delta t)*}^{-1}\mathbf{T}_{F_{t+\Delta t}(P)} - F_{t*}^{-1}\mathbf{T}_{F_t(P)}] \\ &= F_{t*}^{-1} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [F_{\Delta t*}^{-1}\mathbf{T}_{F_{t+\Delta t}(P)} - \mathbf{T}_{F_t(P)}] \\ &= F_{t*}^{-1} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [F_{\Delta t*}^{-1}\mathbf{T}_{F_{\Delta t}(Q)} - \mathbf{T}_Q] \equiv F_{t*}^{-1} (L_{\mathbf{X}}\mathbf{T})_Q. \end{aligned}$$

Since  $Q$  is arbitrary, we can remove it from the equation and write, as the generalization of Equation (26.29),

$$L_{\mathbf{X}}\mathbf{T} = F_{t*} \frac{d}{dt} F_{t*}^{-1}\mathbf{T}. \tag{26.30}$$

An important special case of the definition above is the Lie derivative of a vector field with respect to another. Let  $\mathbf{X}, \mathbf{Y} \in \mathcal{X}(M)$ . To evaluate the RHS of (26.29), we apply the first term in the brackets to an arbitrary function  $f$ ,

$$\begin{aligned} [F_{t*}^{-1}\mathbf{Y}_{F_t(P)}](f) &= \mathbf{Y}_{F_t(P)}(f \circ F_t^{-1}) = \mathbf{Y}(f \circ F_t^{-1})|_{F_t(P)} \\ &= \mathbf{Y}(f \circ F_{-t})|_{F_t(P)} = \mathbf{Y}(f - t\mathbf{X}(f))|_{F_t(P)} \\ &= (\mathbf{Y}f)_{F_t(P)} - t\mathbf{Y}(\mathbf{X}(f))|_{F_t(P)} \\ &= (\mathbf{Y}f)_P + t[\mathbf{X}(\mathbf{Y}f)]_P - t\{[\mathbf{Y}(\mathbf{X}f)]_P + t[\mathbf{X}(\mathbf{Y}(\mathbf{X}f))]_P\} \\ &= \mathbf{Y}_P(f) + t\mathbf{X}_P \circ \mathbf{Y}_P(f) - t\mathbf{Y}_P \circ \mathbf{X}_P(f) \\ &= \mathbf{Y}_P(f) + t[\mathbf{X}_P, \mathbf{Y}_P](f) = \mathbf{Y}_P(f) + t[\mathbf{X}, \mathbf{Y}]_P(f). \end{aligned}$$

The first equality on the first line follows from (26.9), the second equality from the meaning of  $\mathbf{Y}_{F_t(P)}$ ; the second equality on the second line and the fourth line follow from (26.25). Finally, the fifth line follows if we ignore the  $t^2$  term. Therefore,

$$\begin{aligned} (L_{\mathbf{X}}\mathbf{Y})_P(f) &= \lim_{t \rightarrow 0} \frac{1}{t} [F_{t*}^{-1}\mathbf{Y}_{F_t(P)} - \mathbf{Y}_P](f) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \{t[\mathbf{X}, \mathbf{Y}]_P\}(f) = [\mathbf{X}, \mathbf{Y}]_P(f). \end{aligned}$$

Lie derivative is commutator

Since this is true for all  $P$  and  $f$ , we get

$$L_{\mathbf{X}}\mathbf{Y} = [\mathbf{X}, \mathbf{Y}]. \tag{26.31}$$

This and other properties of the Lie derivative are summarized in the following proposition.

Properties of Lie derivative

**26.4.13. Proposition.** *Let  $\mathbf{T} \in T'_s(M)$  and  $\mathbf{T}'$  be arbitrary tensor fields and  $\mathbf{X}$  a given vector field. Then*

1.  $L_{\mathbf{X}}$  satisfies a derivation property in the algebra of tensors, i.e.,

$$L_{\mathbf{X}}(\mathbf{T} \otimes \mathbf{T}') = (L_{\mathbf{X}}\mathbf{T}) \otimes \mathbf{T}' + \mathbf{T} \otimes (L_{\mathbf{X}}\mathbf{T}').$$

2.  $L_{\mathbf{X}}$  is type-preserving, i.e.,  $L_{\mathbf{X}}\mathbf{T}$  is a tensor field of type  $(r, s)$ .
3.  $L_{\mathbf{X}}$  commutes with the operation of contraction of tensor fields; in particular, in combination with property 1, we have

$$L_{\mathbf{X}}\langle \theta, \mathbf{Y} \rangle = \langle L_{\mathbf{X}}\theta, \mathbf{Y} \rangle + \langle \theta, L_{\mathbf{X}}\mathbf{Y} \rangle$$

4.  $L_{\mathbf{X}}f = \mathbf{X}f$  for every function  $f$ .
5.  $L_{\mathbf{X}}\mathbf{Y} = [\mathbf{X}, \mathbf{Y}]$  for every vector field  $\mathbf{Y}$ .

*Proof.* Except for the last property, which we demonstrated above, the rest follow directly from definitions and simple manipulations. The details are left as exercises.  $\square$

Although the Lie derivative of a vector field is nicely given in terms of commutators, no such simple relation exists for the Lie derivative of a 1-form. However, if we work in a given coordinate frame, then a useful expression for the Lie derivative of a 1-form can be obtained. Applying  $L_X$  to  $\langle \theta, X \rangle$ , we obtain

Remember Einstein's summation convention!

$$\underbrace{L_X \langle \theta, Y \rangle}_{=X(\langle \theta, Y \rangle)} = \langle L_X \theta, Y \rangle + \langle \theta, L_X Y \rangle = \langle L_X \theta, Y \rangle + \langle \theta, [X, Y] \rangle. \quad (26.32)$$

In particular, if  $Y = \partial_i$  and we write  $X = X^j \partial_j$ ,  $\theta = \theta_j dx^j$ , then the LHS becomes  $X(\theta_i) = X^j \partial_j \theta_i$ , and the RHS can be written as

$$(L_X \theta)_i + \underbrace{\langle \theta, [X^j \partial_j, \partial_i] \rangle}_{-(\partial_i X^j) \theta_j}.$$

It follows that

$$L_X \theta \equiv (L_X \theta)_i dx^i = (X^j \partial_j \theta_i + \theta_j \partial_i X^j) dx^i. \quad (26.33)$$

We give two other useful properties of the Lie derivative applicable to all tensors. From the Jacobi identity one can readily deduce that

$$L_{[X, Y]} Z = L_X L_Y Z - L_Y L_X Z.$$

Similarly, Equation (26.32) yields

$$L_{[X, Y]} \theta = L_X L_Y \theta - L_Y L_X \theta$$

Putting these two equations together, recalling that a general tensor is a linear combination of tensor products of vectors and 1-forms, and that the Lie derivative obeys the product rule of differentiation, we obtain

$$L_{[X, Y]} T = L_X L_Y T - L_Y L_X T \quad (26.34)$$

for any tensor field  $T$ . Furthermore, Equation (26.33) and the linearity of the Lie bracket imply that  $L_{\alpha X + \beta Y} = \alpha L_X + \beta L_Y$  when acting on vectors and 1-forms. It follows by the same argument as above that

$$L_{\alpha X + \beta Y} T = \alpha L_X T + \beta L_Y T \quad \forall T \in \mathcal{T}_s^r(M). \quad (26.35)$$



## 26.5 Exterior Calculus

Skew-symmetric tensors are of special importance to applications. We studied these tensors in their algebraic format in the last chapter. Let us now investigate them as they reside on manifolds.

differential forms, or simply forms, defined

**26.5.1. Definition.** Let  $M$  be a manifold and  $Q$  a point of  $M$ . Let  $\Lambda_Q^p(M)$  denote the space of all antisymmetric tensors of rank  $p$  over the tangent space at  $Q$ . Let  $\Lambda^p(M)$  be the union of all  $\Lambda_Q^p(M)$  for all  $Q \in M$ . A **differential  $p$ -form**  $\omega$  is a mapping  $\omega : U \rightarrow \Lambda^p(M)$  such that  $\omega(Q) \in \Lambda_Q^p(M)$  where  $U$  is, as usual, a subset of  $M$ . To emphasize their domain of definition, we sometimes use the notation  $\Lambda^p(U)$ .

Since  $\{dx^i\}_{i=1}^m$  is a basis for  $T_Q^*(M)$  at every  $Q \in M$ ,  $\{dx^{i_1} \wedge \dots \wedge dx^{i_p}\}$  is a basis for the  $p$ -forms. All the algebraic properties established in the last chapter apply to these  $p$ -forms at every point  $Q \in M$ .

The concept of a pullback has been mentioned a number of times in connection with linear maps. The most frequent use of pullbacks takes place in conjunction with the  $p$ -forms.

defining pullback for differential forms

**26.5.2. Definition.** Let  $M$  and  $N$  be manifolds and  $\psi : M \rightarrow N$  a differentiable map. The **pullback map** on  $p$ -forms is the map  $\psi^* : \Lambda^p(N) \rightarrow \Lambda^p(M)$  defined by

$$\psi^* \rho(\mathbf{X}_1, \dots, \mathbf{X}_p) = \rho(\psi_* \mathbf{X}_1, \dots, \psi_* \mathbf{X}_p) \quad \text{for } \rho \in \Lambda^p(N).$$

For  $p = 0$ , i.e., for functions on  $M$ ,  $\psi^* \omega \equiv \omega \circ \psi$ .

It can be shown that

$$\psi^*(\omega \wedge \eta) = \psi^* \omega \wedge \psi^* \eta, \quad (\psi \circ \phi)^* = \phi^* \circ \psi^*. \tag{26.36}$$

Since  $\omega$  varies from point to point, we can define its derivatives. Recall that  $T_0^0(M)$  is the collection of real-valued functions on  $M$ . Since the dual of  $\mathbb{R}$  is  $\mathbb{R}$ , we conclude that  $\Lambda^0(M)$ , the collection of zero-forms, is the union of all real-valued functions on  $M$ . Also recall that if  $f$  is a zero-form, then  $df$ , the differential of  $f$ , is a one-form. Thus, the differential operator  $d$  creates a one-form from a zero-form. The fact that this can be generalized to  $p$ -forms is the subject of the next theorem (for a proof, see [Abra 88, pp. 111–112]).

exterior derivative and its antiderivation property

**26.5.3. Theorem.** For each point  $Q$  of  $M$ , there exists a neighborhood  $U$  and a unique operator  $d : \Lambda^p(U) \rightarrow \Lambda^{p+1}(U)$ , called the **exterior derivative** operator, such that for any  $\omega \in \Lambda^p(U)$  and  $\eta \in \Lambda^q(U)$ ,

1.  $d(\omega + \eta) = d\omega + d\eta$  if  $q = p$ ; otherwise the sum is not defined.
2.  $d(\omega \wedge \eta) = (d\omega) \wedge \eta + (-1)^p \omega \wedge (d\eta)$ ; this is called the **antiderivation** property of  $d$  with respect to the wedge product.

3.  $d(d\omega) = 0$  for any differential form  $\omega$ ; stated differently,  $d \circ d = 0$ .
4.  $df = (\partial_i f)dx^i$  for any real-valued function  $f$ .
5.  $d$  is **natural with respect to pullback**; that is,  $d_M \circ \psi^* = \psi^* \circ d_N$  for any differentiable map  $\psi : M \rightarrow N$ . Here  $d_M$  ( $d_N$ ) is the exterior derivative operating on differential forms of  $M$  ( $N$ ).

**26.5.4. Example.** Let  $M = \mathbb{R}^3$  and  $\omega = a_i dx^i$  a 1-form on  $M$ . The exterior derivative of  $\omega$  is

$$d\omega = (da_i) \wedge dx^i = (\partial_j a_i dx^j) \wedge dx^i = \sum_{j < i} (\partial_j a_i - \partial_i a_j) dx^j \wedge dx^i.$$

We see that the components of  $d\omega$  are the components of  $\nabla \times \mathbf{A}$  where  $\mathbf{A} = (a_1, a_2, a_3)$ . It follows that the curl of a vector in  $\mathbb{R}^3$  is the exterior derivative of the 1-form constructed out of the components of the vector. ■

**26.5.5. Example.** In relativistic electromagnetic theory the electric and magnetic fields are combined to form the electromagnetic field tensor. This is a skew-symmetric tensor field of rank 2, which can be written as<sup>8</sup>

$$\begin{aligned} \mathbf{F} = & -E_x dt \wedge dx - E_y dt \wedge dy - E_z dt \wedge dz \\ & + B_z dx \wedge dy - B_y dx \wedge dz + B_x dy \wedge dz, \end{aligned} \tag{26.37}$$

where  $t$  is the time coordinate and the units are such that  $c$ , the velocity of light, is equal to 1.

Let us take the exterior derivative of  $\mathbf{F}$ . In the process, we use  $df = (\partial_i f)dx^i$ ,  $d(dx^i \wedge dx^j) = 0$ , and in  $dE_i$  or  $dB_j$  we include only the terms that give a nonzero contribution:

$$\begin{aligned} d\mathbf{F} = & -\left(\frac{\partial E_x}{\partial y} dy + \frac{\partial E_x}{\partial z} dz\right) \wedge dt \wedge dx - \left(\frac{\partial E_y}{\partial x} dx + \frac{\partial E_y}{\partial z} dz\right) \wedge dt \wedge dy \\ & -\left(\frac{\partial E_z}{\partial x} dx + \frac{\partial E_z}{\partial y} dy\right) \wedge dt \wedge dz + \left(\frac{\partial B_z}{\partial t} dt + \frac{\partial B_z}{\partial z} dz\right) \wedge dx \wedge dy \\ & -\left(\frac{\partial B_y}{\partial t} dt + \frac{\partial B_y}{\partial y} dy\right) \wedge dx \wedge dz + \left(\frac{\partial B_x}{\partial t} dt + \frac{\partial B_x}{\partial x} dx\right) \wedge dy \wedge dz. \end{aligned}$$

Collecting all similar terms and taking into account changes of sign due to the antisymmetry of the exterior products gives

$$\begin{aligned} d\mathbf{F} = & \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} + \frac{\partial B_z}{\partial t}\right) dt \wedge dx \wedge dy + \left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} - \frac{\partial B_y}{\partial t}\right) dt \wedge dx \wedge dz \\ & + \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} + \frac{\partial B_x}{\partial t}\right) dt \wedge dy \wedge dz + \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}\right) dx \wedge dy \wedge dz \\ = & \left[(\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t})_z\right] dt \wedge dx \wedge dy + \left[(\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t})_y\right] dt \wedge dx \wedge dz \\ & + \left[(\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t})_x\right] dt \wedge dy \wedge dz + (\nabla \cdot \mathbf{B}) dx \wedge dy \wedge dz \end{aligned}$$

<sup>8</sup>Note how in the wedge product, the first factor has a lower index (is an “earlier” coordinate) than the second factor. If this restriction is to be removed, we need to introduce a factor of  $\frac{1}{2}$  for each component (see Example 26.5.11).

The homogeneous Maxwell's equations are written in terms of differential forms.

Each component of  $d\mathbf{F}$  vanishes because of Maxwell's equations. ■

The example above shows that

**26.5.6. Box.** *The two homogeneous Maxwell's equations can be written as  $d\mathbf{F} = 0$ , where  $\mathbf{F}$  is defined by Equation (26.37).*

The exterior derivative is a very useful concept in the theory of differential forms, as illustrated in the preceding example. However, that is not the only differentiation available to the differential forms. We have already defined the Lie derivative for arbitrary tensors. Since differential forms are (antisymmetrized) linear combinations of covariant tensors, Lie differentiation is defined for them as well. In fact, since differential forms have no contravariant parts, one uses the pullback map  $F_t^*$  in the definition of the Lie derivative instead of  $F_{t*}^{-1}$ :

$$L_X\omega = (F_t^*)^{-1} \frac{d}{dt} F_t^*\omega. \tag{26.38}$$

The two derivatives defined so far have the following convenient property, whose proof is left as an exercise for the reader:

**26.5.7. Theorem.** *The exterior derivative  $d$  is natural with respect to  $L_X$  (or commutes with  $L_X$ ) for  $X \in \mathcal{X}(M)$ ; that is,  $d \circ L_X = L_X \circ d$ .*

In our definition of exterior product in the previous chapter, we assumed that the number of vectors was equal to the number of linear functionals taken from the dual space [see Equation (25.9)]. As a result of this complete pairing, we always ended up with a number. It is useful, however, to define an "incomplete" pairing in which the number of vectors and dual vectors are not the same. In particular, if we have a  $p$ -form and a single vector field, then we can pair the vector field with one of the factors of the  $p$ -form to get a  $(p - 1)$ -form. This process is important enough to warrant the following:

**26.5.8. Definition.** *Let  $X$  be a vector field and  $\omega$  a  $p$ -form on a manifold  $M$ . Then define  $i_X : \Lambda^p(M) \rightarrow \Lambda^{p-1}(M)$  by*

$$i_X\omega(X_1, \dots, X_{p-1}) = \omega(X, X_1, \dots, X_{p-1}).$$

If  $\omega \in \Lambda^0(M)$ , i.e., if  $\omega$  is just a function, we set  $i_X\omega = 0$ .  $i_X\omega$  is called the **interior product** or **contraction** of  $X$  and  $\omega$ . Another notation commonly used for  $i_X\omega$  is  $X \lrcorner \omega$ .

Although no signature of "differentiation" appears on  $i_X$ , it does have such a property:

interior product of a vector field and a differential form

**26.5.9. Theorem.** Let  $\omega$  be a  $p$ -form and  $\eta$  a  $q$ -form on a manifold  $M$ . Then,  $i_X$  is an antiderivation with respect to the wedge product:

$$i_X(\omega \wedge \eta) = (i_X\omega) \wedge \eta + (-1)^p \omega \wedge (i_X\eta).$$

*Proof.* The proof follows directly from Definitions 25.3.3 and 26.5.8, and is left for the reader. □

We have introduced three types of derivation on the algebra of differential forms: the exterior derivative, the Lie derivative, and the interior product. The following theorem connects all three derivations in a most useful way (see A[Abra 88, pp. 115–116]):

Relation between  $d$ ,  $L_X$ , and  $i_X$

**26.5.10. Theorem.** Let  $\omega \in \Lambda^p(M)$ ,  $f \in \Lambda^0(M)$ , and  $X \in \mathcal{X}(M)$ . Let  $i_X : \Lambda^p(M) \rightarrow \Lambda^{p-1}(M)$ ,  $d : \Lambda^p(M) \rightarrow \Lambda^{p+1}(M)$ , and  $L_X : \Lambda^p(M) \rightarrow \Lambda^p(M)$  be the interior product, the exterior derivative, and the Lie derivative, respectively. Then

1.  $i_X df = L_X f$ .
2.  $L_X = i_X \circ d + d \circ i_X$ .
3.  $L_X \omega = f L_X \omega + df \wedge i_X \omega$ .

If  $X = X^j \partial_j$  and  $\omega = \omega_{i_1 i_2 \dots i_{p+1}} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_{p+1}}$ , then the reader may verify that  $i_X \omega = X^i \omega_{i i_1 \dots i_p} dx^{i_1} \wedge \dots \wedge dx^{i_p}$ . In particular, we have the useful formula

$$\begin{aligned} i_X(dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_{p+1}}) &= X^j \delta_{j i_1 \dots i_p}^{i_1 i_2 \dots i_{p+1}} dx^{j_1} \wedge dx^{j_2} \wedge \dots \wedge dx^{j_p} \\ &= X^j \left( \sum_{\pi} \epsilon_{\pi} \delta_{\pi(j)}^{i_1} \delta_{\pi(j_1)}^{i_2} \dots \delta_{\pi(j_p)}^{i_{p+1}} \right) dx^{j_1} \wedge dx^{j_2} \wedge \dots \wedge dx^{j_p}. \end{aligned} \tag{26.39}$$

analysis of the Lorentz force law in the language of forms

**26.5.11. Example.** Let  $\mathbf{p} = p_{\alpha} dx^{\alpha}$  be the momentum one-form and write the electromagnetic field tensor as<sup>9</sup>  $\mathbf{F} = \frac{1}{2} F_{\alpha\beta} dx^{\alpha} \wedge dx^{\beta}$ , where  $\alpha$  and  $\beta$  run over the values 0, 1, 2, and 3 with 0 being the time index. Let

$$\frac{d\mathbf{p}}{d\tau} \equiv \left( \frac{dp_{\alpha}}{d\tau} \right) dx^{\alpha}$$

be the derivative of momentum with respect to the proper time,  $\tau$ . Also, let  $\mathbf{u} = u^{\beta} \partial_{\beta}$  be the velocity four-vector of a charged particle. Then the Lorentz force law can be written simply as  $d\mathbf{p}/d\tau = q\mathbf{F}(\mathbf{u}) \equiv -q i_{\mathbf{u}} \mathbf{F}$ , where  $q$  is the electric charge of the particle whose 4-velocity is  $\mathbf{u}$ . Note that  $\mathbf{F}$ , a two-form, contracts with  $\mathbf{u}$ , a vector, to give a one-form on

<sup>9</sup>The factor  $\frac{1}{2}$  is introduced here to avoid restricting the sum over  $\alpha$  and  $\beta$ .

the RHS. Thus, both sides are of the same type. Let us write this equation in component form:

$$\begin{aligned}
 \frac{dp_\alpha}{d\tau} dx^\alpha &= -q \frac{1}{2} F_{\alpha\beta} i_\alpha (dx^\alpha \wedge dx^\beta) = -\frac{1}{2} q F_{\alpha\beta} (u^\gamma \delta_{\gamma\mu}^{\alpha\beta} dx^\mu) \\
 &= \frac{1}{2} q F_{\alpha\beta} u^\gamma (\delta_\gamma^\alpha \delta_\mu^\beta - \delta_\mu^\alpha \delta_\gamma^\beta) dx^\mu \\
 &= \frac{1}{2} q F_{\alpha\beta} (u^\beta dx^\alpha - u^\alpha dx^\beta) \\
 &= \frac{1}{2} q (F_{\alpha\beta} - F_{\beta\alpha}) u^\beta dx^\alpha = (q F_{\alpha\beta} u^\beta) dx^\alpha.
 \end{aligned} \tag{26.40}$$

Equating the components on both sides, we get  $dp_\alpha/d\tau = q F_{\alpha\beta} u^\beta$ , which may be familiar to the reader. To make the equation even more familiar, consider the component  $\alpha = 1$ ,

$$\frac{dp_1}{d\tau} = q F_{1\beta} u^\beta = q [F_{10} u^0 + F_{12} u^2 + F_{13} u^3], \tag{26.41}$$

and recall that  $u^\alpha = dx^\alpha/d\tau$ , where

$$(d\tau)^2 = (dt)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = (dt)^2 (1 - v^2)$$

and  $\mathbf{v} = (dx^1/dt, dx^2/dt, dx^3/dt)$  is the 3-velocity of the particle. Since  $x^0 = t$ , we get

$$u^0 = \frac{dt}{d\tau} = \frac{1}{\sqrt{1 - v^2}}, \quad u^i = \frac{dx^i}{d\tau} = \frac{v_i}{\sqrt{1 - v^2}} \quad \text{for } i = 1, 2, 3.$$

Substituting this in (26.41) and remembering that  $F_{10} = -F_{01} = E_1$ ,  $F_{12} = B_3$ , and  $F_{13} = -F_{31} = -B_2$ , we obtain

$$\frac{dp_1}{dt \sqrt{1 - v^2}} = q \left[ E_1 \frac{1}{\sqrt{1 - v^2}} + B_3 \frac{v_2}{\sqrt{1 - v^2}} - B_2 \frac{v_3}{\sqrt{1 - v^2}} \right],$$

or

$$\frac{dp_1}{dt} = q [E_1 + (v_2 B_3 - v_3 B_2)] = [q(\mathbf{E} + \mathbf{v} \times \mathbf{B})]_1.$$

The other components are obtained similarly. Thus, in vector form we have

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

where  $\mathbf{p}$  now represents the 3-momentum of the particle. This is the Lorentz force law for electromagnetism in its familiar form. Again, note the simplification offered by the language of forms.  $\blacksquare$

A combination that is very useful is that of the exterior derivative and the Hodge star operator. Recall that the latter is defined by

$$* (dx^{i_1} \wedge \cdots \wedge dx^{i_p}) = \frac{1}{(m-p)!} \epsilon_{i_{p+1} \dots i_m}^{i_1 \dots i_p} dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_m}, \tag{26.42}$$

where  $m$  is the dimension of the manifold.

**26.5.12. Example.** Let us calculate  $*F$  and  $d(*F)$  where  $F = \frac{1}{2}F_{\alpha\beta}dx^\alpha \wedge dx^\beta$  is the electromagnetic field tensor. We have

$$*F = *(\frac{1}{2}F_{\alpha\beta}dx^\alpha \wedge dx^\beta) = \frac{1}{2}F_{\alpha\beta} * (dx^\alpha \wedge dx^\beta) = \frac{1}{2}F_{\alpha\beta} \frac{1}{2!}\epsilon^{\alpha\beta}_{\mu\nu}dx^\mu \wedge dx^\nu$$

and

$$d(*F) = d(\frac{1}{4}F_{\alpha\beta}\epsilon^{\alpha\beta}_{\mu\nu}dx^\mu \wedge dx^\nu) = \frac{1}{4}\epsilon^{\alpha\beta}_{\mu\nu}F_{\alpha\beta,\gamma}dx^\gamma \wedge dx^\mu \wedge dx^\nu,$$

where  $F_{\alpha\beta,\gamma} \equiv \partial F_{\alpha\beta}/\partial x^\gamma$ . We can now use the components  $F_{j0} = E_j, F_{12} = B_3, F_{13} = -B_2,$  and  $F_{23} = B_1$  to write  $d(*F)$  in terms of  $\mathbf{E}$  and  $\mathbf{B}$ . After a long but straightforward calculation, we obtain

$$\begin{aligned} d(*F) = & \left[ \left( \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{B} \right)_z \right] dt \wedge dx \wedge dy + \left[ \left( \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{B} \right)_y \right] dt \wedge dz \wedge dx \\ & + \left[ \left( \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{B} \right)_x \right] dt \wedge dy \wedge dz + (\nabla \cdot \mathbf{E}) dx \wedge dy \wedge dz. \end{aligned} \tag{26.43}$$

The inhomogeneous pair of Maxwell's equations is

$$\nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} + 4\pi\mathbf{J}, \quad \nabla \cdot \mathbf{E} = 4\pi\rho, \tag{26.44}$$

where  $\rho$  and  $\mathbf{J}$  are charge and current densities, respectively. We can put these two densities together to form a four-current one-form with  $\rho$  as the zeroth component:  $\mathcal{J} = J_\alpha dx^\alpha$ . Thus,

Maxwell's inhomogeneous equations in the language of forms

$$\begin{aligned} * \mathcal{J} = J_\alpha (*dx^\alpha) &= J_\alpha \frac{1}{3!}\epsilon^{\alpha}_{\mu\nu\rho}dx^\mu \wedge dx^\nu \wedge dx^\rho \\ &= J_0 dx \wedge dy \wedge dz + J_x dt \wedge dy \wedge dz + J_y dt \wedge dz \wedge dx + J_z dt \wedge dx \wedge dy \\ &= \rho dx \wedge dy \wedge dz - J^x dt \wedge dy \wedge dz - J^y dt \wedge dz \wedge dx - J^z dt \wedge dx \wedge dy, \end{aligned} \tag{26.45}$$

where we have used the facts that  $\rho = J^0 = J_0$  and  $\mathbf{J} = (J^x, J^y, J^z) = -(J_x, J_y, J_z)$ . ■

Comparing Equations (26.43), (26.44), and (26.45), we note that

**26.5.13. Box.** *In the language of forms, the inhomogeneous pair of Maxwell's equations has the simple appearance  $d(*F) = 4\pi(*\mathcal{J})$ .*

Problem 26.15 shows that the relation  $d^2\omega = 0$  is equivalent—at least in  $\mathbb{R}^3$ —to the vanishing of the curl of the gradient and the divergence of the curl. It is customary in physics to try to go backwards as well, that is, given that  $\nabla \times \mathbf{E} = 0,$  to assume that  $\mathbf{E} = \nabla f$  for some function  $f$ . Similarly, we want to believe that  $\nabla \cdot \mathbf{B} = 0$  implies that  $\mathbf{B} = \nabla \times \mathbf{A}$ .

closed and exact forms

What is the analogue of the above statement for a general  $p$ -form? A form  $\omega$  that satisfies  $d\omega = 0$  is called a **closed form**. An **exact form** is one that can be written as the exterior derivative of another form. Thus, *every exact form is automatically closed*. This is the **Poincaré lemma**. The converse of this lemma is true only if the region of definition of the form is topologically simple, as explained in the following.

regions that are contractable to a point

Consider a  $p$ -form  $\omega$  defined on a region  $U$  of a manifold  $M$ . If all closed curves in  $U$  can be shrunk to a point in  $U$ , we say that  $U$  is **contractable** to a point. If  $\omega$  is not defined for a point  $P$  on  $M$ , then any  $U$  that contains  $P$  is not contractable to a point. We can now state the converse of the Poincaré lemma (for a proof, see [Bish 80, p. 175]):

converse of the Poincaré lemma

**26.5.14. Theorem.** (converse of the Poincaré lemma) *Let  $U$  be a region in a manifold  $M$ , such that  $U$  is contractable to a point. Let  $\omega$  be a  $p$ -form on  $U$  such that  $d\omega = 0$ . Then there exists a  $(p - 1)$ -form  $\eta$  on  $U$  such that  $\omega = d\eta$ .*

**26.5.15. Example.** The electromagnetic field tensor  $F = \frac{1}{2}F_{\alpha\beta}dx^\alpha \wedge dx^\beta$  is a two-form that satisfies  $dF = 0$ . The converse of the Poincaré lemma says that if  $F$  is well behaved in a region  $U$  of  $\mathbb{R}^4$ , then there must exist a one-form  $\eta$  such that  $F = d\eta$ .

Let us write this one-form in terms of coordinates as  $\eta = A_\alpha dx^\alpha$ . Then  $d\eta = A_{\alpha,\beta}dx^\beta \wedge dx^\alpha$ , and we have

$$\frac{1}{2}F_{\alpha\beta}dx^\alpha \wedge dx^\beta = A_{\beta,\alpha}dx^\alpha \wedge dx^\beta \Rightarrow \frac{1}{2}(F_{\alpha\beta} - A_{\beta,\alpha} + A_{\alpha,\beta})dx^\alpha \wedge dx^\beta = 0.$$

Since  $dx^\alpha \wedge dx^\beta$  are linearly independent and their coefficients are antisymmetric, each of the latter must vanish. Thus,

$$F_{\alpha\beta} = A_{\beta,\alpha} - A_{\alpha,\beta} = \frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta}.$$

The four-vector  $A^\alpha$  is simply the four-potential of relativistic electromagnetic theory. ■

gauge invariance in the language of forms

Note that the  $(p - 1)$ -form of Theorem 26.5.14 is not unique. In fact, if  $\alpha$  is any  $(p - 2)$ -form, then  $\omega$  can be written as

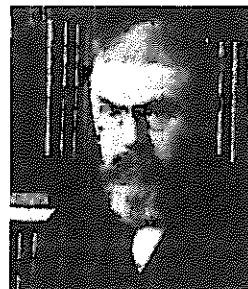
$$\omega = d(\eta + d\alpha)$$

because  $d(d\alpha)$  is identical to zero. This freedom of choice in selecting  $\eta$  is called **gauge invariance**, and its generalization plays an important role in the physics of fundamental interactions.

**Jules Henri Poincaré (1854–1912):** The development of mathematics in the nineteenth century began under the shadow of a giant, Carl Friedrich Gauss; it ended with the domination by a genius of similar magnitude, Henri Poincaré. Both were universal mathematicians in the supreme sense, and both made important contributions to astronomy and mathematical physics. If Poincaré’s discoveries in number theory do not equal those of Gauss, his achievements in the theory of functions are at least on the same level—even when one takes

into account the theory of elliptic and modular functions, which must be credited to Gauss and which represents in that field his most important discovery, although it was not published during his lifetime. If Gauss was the initiator in the theory of differentiable manifolds, Poincaré played the same role in *algebraic topology*. Finally, Poincaré remains the most important figure in the theory of differential equations and the mathematician who after Newton did the most remarkable work in celestial mechanics. Both Gauss and Poincaré had very few students and liked to work alone; but the similarity ends there. Where Gauss was very reluctant to publish his discoveries, Poincaré's list of papers approaches five hundred, which does not include the many books and lecture notes he published as a result of his teaching at the Sorbonne.

Poincaré's parents both belonged to the upper middle class, and both their families had lived in Lorraine for several generations. His paternal grandfather had two sons: Léon, Henri's father, was a physician and a professor of medicine at the University of Nancy; Antoine had studied at the École Polytechnique and rose to high rank in the engineering corps. One of Antoine's sons, Raymond, was several times prime minister and was president of the French Republic during World War I; the other son, Lucien, occupied high administrative functions in the university. Poincaré's mathematical ability became apparent while he was still a student in the *lycée*. He won first prizes in the *concours général* (a competition among students from all French *lycées*) and in 1873 entered the École Polytechnique at the top of his class; his professor at Nancy is said to have referred to him as a "monster of mathematics." After graduation, he followed courses in engineering at the École des Mines and worked briefly as an engineer while writing his thesis for the doctorate in mathematics which he obtained in 1879. Shortly afterward he started teaching at the University of Caen, and in 1881 he became a professor at the University of Paris, where he taught until his untimely death in 1912. At the early age of thirty-three he was elected to the Académie des Sciences and in 1908 to the Académie Française. He was also the recipient of innumerable prizes and honors both in France and abroad.



Before he was thirty years of age, Poincaré became world famous with his epoch-making discovery of the "automorphic functions" of one complex variable (or, as he called them, the "fuchsian" and "kleinean" functions). Much has been written on the "competition" between C. F. Klein and Poincaré in the discovery of automorphic functions. However, Poincaré's ignorance of the mathematical literature when he started his researches is almost unbelievable. He hardly knew anything on the subject beyond Hermite's work on the modular functions; he certainly had never read Riemann, and by his own account had not even heard of the "Dirichlet principle," which he was to use in such imaginative fashion a few years later. Nevertheless, Poincaré's idea of associating a fundamental domain to any fuchsian group does not seem to have occurred to Klein, nor did the idea of "using" non-Euclidean geometry, which is never mentioned in his papers on modular functions up to 1880.

Poincaré was one of the few mathematicians of his time who understood and admired the work of Lie and his continuators on "continuous groups," and in particular the only mathematician who in the early 1900s realized the depth and scope of E. Cartan's papers. In 1899 Poincaré proved what is now called the *Poincaré-Birkhoff-Witt theorem* which has become fundamental in the modern theory of Lie algebras. The theory of differential



equations and its applications to dynamics was clearly at the center of Poincaré's mathematical thought; from his first (1878) to his last (1912) paper, he attacked the theory from all possible angles and very seldom let a year pass without publishing a paper on the subject. The most extraordinary production of Poincaré's, also dating from his prodigious period of creativity (1880–1883) (reminding us of Gauss's *Tagebuch* of 1797–1801), is the qualitative theory of differential equations. It is one of the few examples of a mathematical theory that sprang apparently out of nowhere and that almost immediately reached perfection in the hands of its creator. Everything was new in the first two of the four big papers that Poincaré published on the subject between 1880 and 1886.

For more than twenty years Poincaré lectured at the Sorbonne on mathematical physics; he gave himself to that task with his characteristic thoroughness and energy, with the result that he became an expert in practically all parts of theoretical physics, and published more than seventy papers and books on the most varied subjects, with a predilection for the theories of light and of electromagnetic waves. On two occasions he played an important part in the development of the new ideas and discoveries that revolutionized physics at the end of the nineteenth century. His remark on the possible connection between X-rays and the phenomenon of phosphorescence was the starting point of H. Becquerel's experiments that led him to the discovery of radioactivity. On the other hand, Poincaré was active from 1899 on in the discussions concerning Lorentz's theory of the electron; Poincaré was the first to observe that the Lorentz transformations form a group; and many physicists consider that Poincaré shares with Lorentz and Einstein the credit for the invention of the special theory of relativity. The main leitmotiv of Poincaré's mathematical work is clearly the idea of "continuity": Whenever he attacks a problem in analysis, we almost immediately see him investigating what happens when the conditions of the problem are allowed to vary continuously. He was therefore bound to encounter at every turn what we now call *topological* problems. He himself said in 1901, "Every problem I had attacked led me to *Analysis situs*," particularly the researches on differential equations and on the periods of multiple integrals. Starting in 1894 he inaugurated in a remarkable series of six papers—written during a period of ten years—the modern methods of algebraic topology.

Whereas Poincaré has been accused of being too conservative in physics, he certainly was very open-minded regarding new mathematical ideas. The quotations in his papers show that he read extensively, if not systematically, and was aware of all the latest developments in practically every branch of mathematics. He was probably the first mathematician to use Cantor's theory of sets in analysis. Up to a certain point, he also looked with favor on the axiomatic trend in mathematics, as it was developing toward the end of the nineteenth century, and he praised Hilbert's *Grundlagen der Geometrie*. However, he obviously had a blind spot regarding the formalization of mathematics, and poked fun repeatedly at the efforts of the disciples of Peano and Russell in that direction; but, somewhat paradoxically, his criticism of the early attempts of Hilbert was probably the starting point of some of the most fruitful of the later developments of metamathematics. Poincaré stressed that Hilbert's point of view of defining objects by a system of axioms was admissible only if one could prove a priori that such a system did not imply contradiction, and it is well known that the proof of noncontradiction was the main goal of the theory that Hilbert founded after 1920. Poincaré seems to have been convinced that such attempts were hopeless, and K. Gödel's theorem proved him right.

---

### 26.5.1 Integration on Manifolds

Integration of  
differential forms in  
 $\mathbb{R}^n$

We mentioned in Chapter 25 that certain exterior products are interpreted as volume elements. We now exploit this notion and define integration on manifolds. Starting with  $\mathbb{R}^n$ , considered as a manifold, we define the integral of an  $n$ -form  $\omega$  as follows. Choose a coordinate system  $\{x^i\}_{i=1}^n$  in  $\mathbb{R}^n$ , write  $\omega = f dx^1 \wedge \cdots \wedge dx^n$ , and define the integral of the  $n$ -form as

$$\int_{\mathbb{R}^n, x} \omega \equiv \int_{\mathbb{R}^n} f(x^1, \dots, x^n) dx^1 \cdots dx^n,$$

where to avoid dealing with infinities, one assumes that  $f$  vanishes outside a finite region. The second symbol in the lower part of the integral sign indicates the variables of integration. Let us now change the coordinates, say to  $\{y^j\}_{j=1}^n$ . Using Equation (26.17), which gives the transformation rule for 1-forms when changing coordinates, and Equation (25.13), which defines the determinant in terms of  $n$ -forms, we obtain

$$\omega = f dx^1 \wedge \cdots \wedge dx^n = f \det \left( \frac{\partial x^i}{\partial y^j} \right) dy^1 \wedge \cdots \wedge dy^n,$$

where  $f$  is now understood to be a function of the  $y$ 's through the  $x$ 's. So, in terms of the new coordinates, the integral becomes

$$\int_{\mathbb{R}^n, y} \omega = \int_{\mathbb{R}^n} f(x^1(y), \dots, x^n(y)) \det \left( \frac{\partial x^i}{\partial y^j} \right) dy^1 \wedge \cdots \wedge dy^n.$$

If we had the absolute value of the Jacobian in the integral, the two sides would be equal. So, all we can say at this point is

$$\int_{\mathbb{R}^n, y} \omega = \pm \int_{\mathbb{R}^n, x} \omega.$$

This discussion is analogous to our discussion of orientation in vector spaces (see Section 25.3.2).

We therefore distinguish between two kinds of coordinate transformations: If the Jacobian determinant is positive, we say that the coordinate transformation is **orientation preserving**. Otherwise, the transformation is called **orientation reversing**.

Our ability to integrate functions on  $\mathbb{R}^n$  depends crucially on the fact that volume elements do not change sign at any point of  $\mathbb{R}^n$ . If this were not so, we could find a finite (albeit small) region of space—in the vicinity of the point at which the volume element changes sign—whose volume would be zero. This property of  $\mathbb{R}^n$  is the content of the following:

orientable manifolds

**26.5.16. Definition.** A manifold  $M$  of dimension  $n$  is called **orientable** if it has a nowhere vanishing  $n$ -form.

Any two nonvanishing  $n$ -forms  $\omega$  and  $\omega'$  on an orientable manifold are related by a nowhere-vanishing function:  $\omega' = f\omega$ . Clearly,  $f$  has to be either positive or negative everywhere.  $\omega$  and  $\omega'$  are said to be *equivalent* if  $f$  is positive. Thus, the nonvanishing  $n$ -forms on an orientable manifold fall into two classes, all members of each class being equivalent to one another, and a member of one class being related to a member of the other class via a negative function. Each class is called an **orientation** on  $M$ .

Given an orientation, an  $n$ -form  $\omega$ , and a chart  $\{U_\alpha, \phi_\alpha\}$  on  $M$ , we define

$$\int_M \omega \equiv \sum_\alpha \int_{\mathbb{R}^n} (\phi_\alpha^{-1})^*(\omega|_\alpha), \tag{26.46}$$

where  $(\phi_\alpha^{-1})^*$  is the pullback of  $\phi_\alpha^{-1} : \mathbb{R}^n \rightarrow M$ , so that it maps  $n$ -forms on  $M$  to  $n$ -forms on  $\mathbb{R}^n$ ;  $\omega|_\alpha$  is the restriction of  $\omega$  to  $U_\alpha$ , and the sum over  $\alpha$  is assumed to exist. This amounts to saying that the region in  $M$  on which  $\omega$  is defined is finite, or that  $\omega$  has *compact support*.

We note that the RHS of Equation (26.46) is an integration on  $\mathbb{R}^n$  that appears to depend on the choice of coordinate functions. However, it can be shown that the integral is independent of such choice. In practice, one chooses a coordinate patch and transfers the integration to  $\mathbb{R}^n$ , where the process is familiar.

“volume” of a manifold

If we choose a coordinate patch  $\{x^i\}_{i=1}^n$  and integrate  $dx^1 \wedge \cdots \wedge dx^n$  according to Equation (26.46), we obtain the “volume” of the manifold  $M$ . If  $M$  is compact, this volume will be finite.<sup>10</sup>

## 26.6 Symplectic Geometry

Mechanics stimulated a great deal of dialogue between physics and mathematics in the latter part of the nineteenth century and the beginning of the twentieth. The branch of mathematics that benefited the most out of this dialog is the theory of differentiable manifolds, whose tribute back to mechanics has been the most beautiful language in which the latter can express itself, the language of symplectic geometry. All the discussion of symplectic vector spaces of the last chapter can be carried over to the tangent spaces of a manifold and patched together by the differentiable structure of the manifold.

symplectic form, symplectic structure, and symplectic manifold defined

**26.6.1. Definition.** A *symplectic form* (or a *symplectic structure*) on a manifold  $M$  is a nondegenerate, closed 2-form  $\omega$  on  $M$ . A *symplectic manifold*  $(M, \omega)$  is a manifold  $M$  together with a symplectic form  $\omega$  on  $M$ . We define the map  $b : \mathcal{X}(M) \rightarrow \mathcal{X}^*(M)$  by

$$b(\mathbf{X}) \equiv \mathbf{X}^b = i_{\mathbf{X}}\omega = \omega^b(\mathbf{X})$$

<sup>10</sup>Recall from Chapter 16 that a subset of  $\mathbb{R}^n$  is compact iff it is closed and bounded. It is a good idea to keep this in mind as a paradigm of compact spaces.

and the map  $\sharp : \mathcal{X}^*(M) \rightarrow \mathcal{X}(M)$  as the inverse of  $\flat$ .

Chapter 25 identified some special basis, the canonical basis, in which the symplectic form of a symplectic vector space took on a simple expression. The analogue of such a basis exists in a symplectic manifold. The reader should keep in mind that this existence is not automatic, because although one can find such bases at every point of the manifold, the smooth patching up of all such bases to cover the entire manifold is not trivial and is the content of the following important theorem, which we state without proof (see [Abra 85, p. 175]):

**26.6.2. Theorem. (Darboux)** Suppose  $\omega$  is a 2-form on a  $2n$ -dimensional manifold  $M$ . Then  $d\omega = 0$  if and only if there is a chart  $(U, \varphi)$  at each  $P \in M$  such that  $\varphi(P) = \mathbf{0}$  and

$$\omega = \sum_{i=1}^n dx^i \wedge dy^i,$$

where  $x^1, \dots, x^n, y^1, \dots, y^n$  are coordinates on  $U$ . Furthermore, on such a chart, the volume element  $\mu_\omega$  is

$$\mu_\omega = dx^1 \wedge \dots \wedge dx^n \wedge dy^1 \wedge \dots \wedge dy^n.$$

symplectic charts,  
canonical  
coordinates, and  
canonical  
transformations

**26.6.3. Definition.** The charts guaranteed by Darboux's theorem are called **symplectic charts**, and the coordinates  $x^i, y^i$  are called **canonical coordinates**. If  $(M, \omega)$  and  $(N, \rho)$  are symplectic manifolds, then a  $C^\infty$  map  $f : M \rightarrow N$  is called **symplectic**, or a **canonical transformation**, if  $f^* \rho = \omega$ .

Coordinate  
representation of  
sharp and flat maps

**26.6.4. Example.** In this example, we derive a formula that gives the action of  $\omega^\flat$  and  $\omega^\sharp$  in terms of components of vectors and 1-forms in canonical coordinates. Let

$$\mathbf{Z} \equiv X^i \frac{\partial}{\partial x^i} + Y^i \frac{\partial}{\partial y^i}$$

be a vector field. When  $\omega^\flat$  acts on  $\mathbf{Z}$ , it gives a 1-form, which we write as  $\omega^\flat(\mathbf{Z}) \equiv U_k dx^k + W_k dy^k$ . To find the unknowns  $U_k$  and  $W_k$ , we let both sides act on coordinate basis vectors. For the RHS, we get

$$(U_k dx^k + W_k dy^k) \left( \frac{\partial}{\partial x^j} \right) = \underbrace{U_k dx^k \left( \frac{\partial}{\partial x^j} \right)}_{= \delta_j^k} + \underbrace{W_k dy^k \left( \frac{\partial}{\partial x^j} \right)}_{= 0} = U_j$$

and

$$(U_k dx^k + W_k dy^k) \left( \frac{\partial}{\partial y^j} \right) = W_j.$$

For the LHS, we obtain

$$\begin{aligned} \left[ \omega^\flat \left( X^i \frac{\partial}{\partial x^i} + Y^i \frac{\partial}{\partial y^i} \right) \right] \left( \frac{\partial}{\partial x^j} \right) &= X^i \left[ \omega^\flat \left( \frac{\partial}{\partial x^i} \right) \right] \left( \frac{\partial}{\partial x^j} \right) \\ &\quad + Y^i \left[ \omega^\flat \left( \frac{\partial}{\partial y^i} \right) \right] \left( \frac{\partial}{\partial x^j} \right) \\ &= X^i \omega \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) + Y^i \omega \left( \frac{\partial}{\partial y^i}, \frac{\partial}{\partial x^j} \right). \end{aligned}$$

But

$$\begin{aligned} \omega \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) &= \left( \sum_{k=1}^n dx^k \wedge dy^k \right) \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \\ &= \sum_{k=1}^n dx^k \underbrace{\left( \frac{\partial}{\partial x^i} \right) dy^k}_{=0} \left( \frac{\partial}{\partial x^j} \right) - \sum_{k=1}^n dx^k \left( \frac{\partial}{\partial x^j} \right) \underbrace{dy^k}_{=0} \left( \frac{\partial}{\partial x^i} \right) = 0 \end{aligned}$$

and

$$\begin{aligned} \omega \left( \frac{\partial}{\partial y^i}, \frac{\partial}{\partial x^j} \right) &= \left( \sum_{k=1}^n dx^k \wedge dy^k \right) \left( \frac{\partial}{\partial y^i}, \frac{\partial}{\partial x^j} \right) \\ &= \sum_{k=1}^n dx^k \underbrace{\left( \frac{\partial}{\partial y^i} \right) dy^k}_{=0} \left( \frac{\partial}{\partial x^j} \right) - \sum_{k=1}^n dx^k \left( \frac{\partial}{\partial x^j} \right) \underbrace{dy^k}_{=\delta_j^k} \left( \frac{\partial}{\partial y^i} \right) = -\delta_j^i. \end{aligned}$$

It follows that

$$\left[ \omega^\flat \left( X^i \frac{\partial}{\partial x^i} + Y^i \frac{\partial}{\partial y^i} \right) \right] \left( \frac{\partial}{\partial x^j} \right) = -Y^j.$$

Similarly,

$$\left[ \omega^\flat \left( X^i \frac{\partial}{\partial x^i} + Y^i \frac{\partial}{\partial y^i} \right) \right] \left( \frac{\partial}{\partial y^j} \right) = X^j.$$

Therefore,

$$\omega^\flat \left( X^i \frac{\partial}{\partial x^i} + Y^i \frac{\partial}{\partial y^i} \right) = -Y^j dx^j + X^j dy^j, \quad (26.47)$$

where a summation over repeated indices is understood.

If we multiply both sides of this equation by  $\omega^\sharp$  on the left and recall that  $\omega^\sharp \omega^\flat = 1$ , we obtain the following equation for the action of  $\omega^\sharp$ :

$$\omega^\sharp (X^j dx^j + Y^j dy^j) = Y^i \frac{\partial}{\partial x^i} - X^i \frac{\partial}{\partial y^i}. \quad (26.48)$$

Equations (26.47) and (26.48) are very useful in Hamiltonian mechanics. ■

Our discussion of symplectic transformations of symplectic vector spaces showed that such maps are necessarily isomorphisms. Applied to the present situation, this means that if  $f : M \rightarrow N$  is symplectic, then  $f_* : \mathcal{T}_P(M) \rightarrow \mathcal{T}_{f(P)}(N)$  is an isomorphism. Theorem 26.3.2, the inverse mapping theorem, now gives the following theorem.

**26.6.5. Theorem.** *If  $f : M \rightarrow N$  is symplectic, then it is a local diffeomorphism.*

**26.6.6. Example.** Hamiltonian mechanics takes place in the phase space of a system. The phase space is derived from the configuration space as follows. Let  $(q_1, \dots, q_n)$  be the generalized coordinates of a mechanical system. They describe an  $n$ -dimensional manifold  $N$ . The dynamics of the system is described by the (time-independent) Lagrangian  $L$ , which is a function of  $(q^i, \dot{q}^i)$ . But  $\dot{q}^i$  are the components of a vector at  $(q_1, \dots, q_n)$  [see Equation (26.13) and replace  $\gamma^i$  with  $x^i$ ]. Thus, in the language of manifold theory, a Lagrangian is a function on the tangent bundle,  $L : T(N) \rightarrow \mathbb{R}$ .

from Lagrangian to Hamiltonian in the language of differential forms

The Hamiltonian is obtained from the Lagrangian by a Legendre transformation:  $H = \sum_{i=1}^n p_i \dot{q}^i - L$ . The first term can be thought of as a pairing of an element of the tangent space with its dual. In fact, if  $P$  has coordinates  $(q_1, \dots, q_n)$ , then  $\dot{q} \equiv \dot{q}^i \partial_i \in \mathcal{T}_P(N)$  (with the Einstein summation convention enforced), and if we pair this with the dual vector  $p_j dx^j \in \mathcal{T}_P^*(N)$ , we obtain the first term in the definition of the Hamiltonian. The effect of the Legendre transformation is to replace  $\dot{q}^i$  by  $p_i$  as the second set of independent variables. This has the effect of replacing  $T(N)$  with  $T^*(N)$ . Thus

**26.6.7. Box.** *The manifold of Hamiltonian dynamics, or the phase space, is  $T^*(N)$ , with coordinates  $(q^i, p_i)$  on which the Hamiltonian  $H : T^*(N) \rightarrow \mathbb{R}$  is defined.*

symplectic 2-form of  $T^*(M)$

$T^*(N)$  is  $2n$ -dimensional; so it has the potential of becoming a symplectic manifold. In fact, it can be shown that<sup>11</sup> the 2-form suggested by Darboux's theorem,

$$\omega \equiv \sum_{i=1}^n dq^i \wedge dp_i, \tag{26.49}$$

is nondegenerate, and therefore a symplectic form for  $T^*(N)$ . ■

The phase space, equipped with a symplectic form, turns into a geometric arena in which Hamiltonian mechanics unfolds. We saw in the above example that a Hamiltonian is a function on the phase space. More generally, if  $(M, \omega)$  is a symplectic manifold, a Hamiltonian  $H$  is a real-valued function,  $H : M \rightarrow \mathbb{R}$ . Given a Hamiltonian, one can define a vector field as follows. Consider  $dH \in T^*(M)$ . For a symplectic manifold, there is a natural isomorphism between  $T^*(M)$

<sup>11</sup>Here, we are assuming that the mechanical system in question is nonsingular, by which is meant that there are precisely  $n$  independent  $p_i$ 's. There are systems of considerable importance that happen to be singular. Such systems, among which are included all gauge theories such as the general theory of relativity, are called **constrained systems** and are characterized by the fact that  $\omega$  is degenerate. Although of great interest and currently under intense study, we shall not discuss constrained systems in this book.

Hamiltonian vector field and Hamiltonian systems defined

and  $T(M)$ , namely,  $\omega^\sharp$ . The unique vector field  $X_H$  associated with  $dH$  is the vector field we are after.

**26.6.8. Definition.** Let  $(M, \omega)$  be a symplectic manifold and  $H : M \rightarrow \mathbb{R}$  a real-valued function. The vector field

$$X_H \equiv \omega^\sharp(dH) \equiv (dH)^\sharp$$

is called the **Hamiltonian vector field with energy function  $H$** . The triplet  $(M, \omega, X_H)$  is called a **Hamiltonian system**.

The significance of the Hamiltonian vector field lies in its integral curve which turns out to be the path of evolution of the system in the phase space. This is shown in the following proposition.

**26.6.9. Proposition.** If  $(q^1, \dots, q^n, p_1, \dots, p_n)$  are canonical coordinates for  $\omega$ —so  $\omega = \sum dq^i \wedge dp_i$ —then, in these coordinates

$$X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial x^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i} \equiv \left( \frac{\partial H}{\partial p_i}, -\frac{\partial H}{\partial q^i} \right). \tag{26.50}$$

Therefore,  $(q(t), p(t))$  is an integral curve of  $X_H$  iff Hamilton's equations hold:

$$\frac{\partial q^i}{\partial t} = \frac{\partial H}{\partial p_i}, \quad \frac{\partial p_i}{\partial t} = -\frac{\partial H}{\partial q^i}, \quad i = 1, \dots, n. \tag{26.51}$$

*Proof.* The first part of the proposition follows from

$$dH = \frac{\partial H}{\partial q^i} dq^i + \frac{\partial H}{\partial p_i} dp_i,$$

from the definition of  $X_H$  in terms of  $dH$ , and from Equation (26.48). The second part follows from the definition of integral curve and Equation (26.19).  $\square$

We called  $H$  the energy function; this is for good reason:

conservation of energy in the language of symplectic geometry

**26.6.10. Theorem.** Let  $(M, \omega, X_H)$  be a Hamiltonian system and  $\gamma(t)$  an integral curve of  $X_H$ . Then  $H(\gamma(t))$  is constant in  $t$ .

*Proof.* We show that the time-derivative of  $H(\gamma(t))$  is zero:

$$\begin{aligned} \frac{d}{dt} H(\gamma(t)) &= \gamma_{*t}(H) && \text{by Proposition 26.2.4} \\ &= dH(\gamma_{*t}) && \text{by Equation (26.14)} \\ &= dH(X_H(\gamma(t))) && \text{by definition of integral curve} \\ &= [\omega^\flat(X_H(\gamma(t)))](X_H(\gamma(t))) && \text{by definition of } X_H(\gamma(t)) \\ &= \omega(X_H(\gamma(t)), X_H(\gamma(t))) && \text{by the definition of } \omega^\flat \\ &= 0 && \text{because } \omega \text{ is skew-symmetric} \end{aligned}$$

$\square$

Theorem 26.6.10 is the statement of the conservation of energy.

In the theoretical development of mechanics, canonical transformations play a central role. The following proposition shows that the flows of a Hamiltonian system are such transformations:

flow of Hamiltonian  
vector field is  
canonical  
transformation of  
mechanics

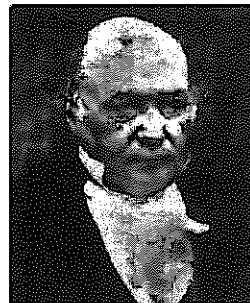
**26.6.11. Proposition.** *Let  $(M, \omega, X_H)$  be a Hamiltonian system, and  $F_t$  the flow of  $X_H$ . Then for each  $t$ ,  $F_t^* \omega = \omega$ , i.e.,  $F_t$  is symplectic.*

*Proof.* We have

$$\begin{aligned} \frac{d}{dt} F_t^* \omega &= F_t^* L_{X_H} \omega && \text{by Equation (26.38)} \\ &= F_t^* (i_{X_H} d\omega + di_{X_H} \omega) && \text{by Theorem 26.5.10} \\ &= F_t^* (0 + ddH) && \text{because } d\omega = 0 \text{ and } i_{X_H} \omega = \omega^b(X) \\ &= 0 && \text{because } d^2 = 0 \end{aligned}$$

Thus,  $F_t^* \omega$  is constant in  $t$ . But  $F_0^* = \text{id}$ . Therefore,  $F_t^* \omega = \omega$ . □

**Sir William Rowan Hamilton** (1805–1865), the fourth of nine children, was mostly raised by an uncle, who quickly realized the extraordinary nature of his young nephew. By the age of five, Hamilton spoke Latin, Greek, and Hebrew, and by the age of nine had added more than a half dozen languages to that list. He was also quite famous for his skill at rapid calculation. Hamilton's introduction to mathematics came at the age of 13, when he studied Clairaut's *Algebra*, a task made somewhat easier as Hamilton was fluent in French by this time. At age 15 he started studying Newton, whose *Principia* spawned an interest in astronomy that would provide a great influence in Hamilton's early career.



In 1822, at the age of 18, Hamilton entered Trinity College, Dublin, and in his first year he obtained the top mark in classics. He divided his studies equally between classics and mathematics and in his second year he received the top award in mathematical physics. Hamilton discovered an error in Laplace's *Mécanique céleste*, and as a result, he came to the attention of John Brinkley, the Astronomer Royal of Ireland, who said: "This young man, I do not say will be, but is, the first mathematician of his age." While in his final year as an undergraduate, he presented a memoir entitled *Theory of Systems of Rays* to the Royal Irish Academy in which he planted the seeds of **symplectic geometry**.

Hamilton's personal life was marked at first by despondency. Rejected by a college friend's sister, he became ill and nearly suicidal. He was rejected a few years later by another friend's sister and wound up marrying a very timid woman prone to ill health. Hamilton's own personality was much more energetic and humorous, and he easily acquired friends among the literati. His own attempts at poetry, which he himself fancied, were generally considered quite poor. No less an authority than Wordsworth attempted to convince him



that his true calling was mathematics, not poetry. Nevertheless, Hamilton maintained close connection with the worlds of literature and philosophy, insisting that the ideas to be gleaned from them were integral parts of his life's work. While Hamilton is best known in physics for his work in dynamics, more of his time was spent on studies in optics and the theory of quaternions. In optics, he derived a function of the initial and final coordinates of a ray and termed it the "characteristic function," claiming that it contained "the whole of mathematical optics." Interestingly, his approach shed no new light on the wave/corpuscular debate (being independent of which view was taken), another appearance of Hamilton's quest for ultimate generality.

In 1833 Hamilton published a study of *vectors* as ordered pairs. He used algebra to study dynamics in *On a General Method in Dynamics* in 1834. The theory of quaternions, on which he spent most of his time, grew from his dissatisfaction with the current state of the theoretical foundation of algebra. He was aware of the description of complex numbers as points in a plane and wondered if any other geometrical representation was possible or if there existed some hypercomplex number that could be represented by three-dimensional points in space. If the latter supposition were true, it would entail a natural algebraic representation of ordinary space. To his surprise, Hamilton found that in order to create a hypercomplex number algebra for which the modulus of a product equaled the product of the two moduli, *four components* were required—hence, quaternions.

Hamilton felt that this discovery would revolutionize mathematical physics, and he spent the rest of his life working on quaternions, including publication of a book entitled *Elements of Quaternions*, which he estimated would be 400 pages long and take two years to write. The title suggests that Hamilton modeled his work on Euclid's *Elements* and indeed this was the case. The book ended up double its intended length and took seven years to write. In fact, the final chapter was incomplete when Hamilton died, and the book was finally published with a preface by his son, William Edwin Hamilton. While quaternions themselves turned out to be of no such monumental importance, their appearance as the first noncommutative algebra opened the door for much research in this field, including much of vector and matrix analysis. (As a side note, the "del" operator, named later by Gibbs, was introduced by Hamilton in his papers on quaternions.)

In dynamics, Hamilton extended his characteristic function from optics to the classical action for a system moving between two points in configuration space. A simple transformation of this function gives the quantity (the time integral of the Lagrangian) whose variation equals zero in what we now call Hamilton's principle. Jacobi later simplified the application of Hamilton's idea to mechanics, and it is the Hamilton–Jacobi equation that is most often used in such problems. Hamiltonian dynamics was rescued from what could have become historical obscurity with the advent of quantum mechanics, in which its close association with ideas in optics found fertile application in the wave mechanics of de Broglie and Schrödinger. Hamilton's later life was unhappy, and he became addicted to alcohol. He died from a severe attack of gout shortly after receiving the news that he had been elected the first foreign member of the National Academy of Sciences of the USA.

The celebrated Liouville's theorem of mechanics, concerning the preservation of volume of the phase space, is a consequence of the proposition above:

**26.6.12. Corollary.** (Liouville's Theorem)  $F_t$  preserves the phase volume  $\mu_\omega$ .

Poisson brackets in the language of symplectic geometry

**26.6.13. Definition.** Let  $(M, \omega)$  be a symplectic manifold. Let  $f, g : M \rightarrow \mathbb{R}$  with  $\mathbf{X}_f = (df)^\sharp$  and  $\mathbf{X}_g = (dg)^\sharp$  their corresponding Hamiltonian vector fields. The Poisson bracket of  $f$  and  $g$  is the function

$$\{f, g\} \equiv \omega(\mathbf{X}_f, \mathbf{X}_g) = i_{\mathbf{X}_g} i_{\mathbf{X}_f} \omega = -i_{\mathbf{X}_f} i_{\mathbf{X}_g} \omega.$$

We can immediately obtain the familiar expression for the Poisson bracket of two functions.

**26.6.14. Proposition.** In canonical coordinates  $(q^1, \dots, q^n, p_1, \dots, p_n)$ , we have

$$\{f, g\} = \sum_{i=1}^n \left( \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} \right).$$

In particular,

$$\{q^i, q^j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{q^i, p_j\} = \delta^i_j.$$

*Proof.* From Equation (26.50), we have

$$\begin{aligned} \omega(\mathbf{X}_f, \mathbf{X}_g) &= \omega \left( \frac{\partial f}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i} \frac{\partial}{\partial p_i}, \frac{\partial g}{\partial p_j} \frac{\partial}{\partial q^j} - \frac{\partial g}{\partial q^j} \frac{\partial}{\partial p_j} \right) \\ &= \sum_{i,j=1}^n \left[ \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial p_j} \underbrace{\omega \left( \frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j} \right)}_{=0} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^j} \underbrace{\omega \left( \frac{\partial}{\partial q^i}, \frac{\partial}{\partial p_j} \right)}_{=\delta^i_j} \right. \\ &\quad \left. - \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_j} \underbrace{\omega \left( \frac{\partial}{\partial p_i}, \frac{\partial}{\partial q^j} \right)}_{=-\delta^i_j} + \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial q^j} \underbrace{\omega \left( \frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j} \right)}_{=0} \right] \\ &= \sum_{i=1}^n \left( \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} \right), \end{aligned}$$

where we have assumed that  $\omega = \sum_{k=1}^n dq^k \wedge dp_k$ . The other formulas follow immediately once we substitute  $p_i$  or  $q^i$  for  $f$  or  $g$ . □

## 26.7 Problems

**26.1.** Provide the details of the fact that a finite-dimensional vector space  $\mathcal{V}$  is a manifold of dimension  $\dim \mathcal{V}$ .

**26.2.** Choose a different curve  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$  whose tangent at  $u = 0$  is still  $(a_x, a_y)$  of Example 26.2.2. For instance, you may choose

$$\gamma(u) = \left(\frac{1}{2}(u+1)^2, \frac{1}{3}(u-1)^3\right).$$

Show that this curve gives the same relation between partials and unit vectors as obtained in that example. Can you find another curve doing the same job?

**26.3.** For every  $t \in \mathcal{T}_P(M)$  and every constant function  $c \in F^\infty(P)$ , show that  $t(c) = 0$ . Hint: Use both parts of Definition 26.2.3 on the two functions  $f = c$  and  $g = 1$ .

**26.4.** Find the coordinate vector field  $\partial_1$  of Example 26.2.10.

**26.5.** Use the procedure of Example 26.2.10 to find a coordinate frame for  $S^2$  corresponding to the stereographic projection charts (See Example 26.1.12).

**26.6.** Let  $(x^i)$  and  $(y^j)$  be coordinate systems on a subset  $U$  of a manifold  $M$ . Let  $X^i$  and  $Y^i$  be the components of a vector field with respect to the two coordinate systems. Show that  $Y^i = X^j \partial y^i / \partial x^j$

**26.7.** Show that if  $\psi : M \rightarrow N$  is a local diffeomorphism at  $P \in M$ , then  $\psi_{*P} : \mathcal{T}_P(M) \rightarrow \mathcal{T}_{\psi(P)}(N)$  is a vector space isomorphism.

**26.8.** Let  $\mathbf{X}$  be a vector field on  $M$  and  $\psi : M \rightarrow N$  a differentiable map. Then for any function  $f$  on  $N$ ,  $[\psi_*\mathbf{X}](f)$  is a function on  $N$ . Show that

$$\mathbf{X}(f \circ \psi) = \{[\psi_*\mathbf{X}](f)\} \circ \psi.$$

**26.9.** Verify that the vector field  $\mathbf{X} = -y\partial_x + x\partial_y$  has an integral curve through  $(x_0, y_0)$  given by

$$\begin{aligned}x &= x_0 \cos t - y_0 \sin t, \\y &= x_0 \sin t + y_0 \cos t.\end{aligned}$$

**26.10.** Show that the vector field  $\mathbf{X} = x^2\partial_x + xy\partial_y$  has an integral curve through  $(x_0, y_0)$  given by

$$x(t) = \frac{x_0}{1 - x_0 t}, \quad y(t) = \frac{y_0}{1 - x_0 t}.$$

**26.11.** Let  $\mathbf{X}$  and  $\mathbf{Y}$  be vector fields. Show that  $\mathbf{X} \circ \mathbf{Y} - \mathbf{X} \circ \mathbf{Y}$  is also a vector field, i.e., it satisfies the derivation property.

**26.12.** Prove the remaining parts of Proposition 26.4.13.

**26.13.** Suppose that  $x^i$  are coordinate functions on a subset of  $M$  and  $\omega$  and  $\mathbf{X}$  are a 1-form and a vector field there. Express  $\omega(\mathbf{X})$  in terms of component functions of  $\omega$  and  $\mathbf{X}$ .

**26.14.** Show that  $d \circ L_X = L_X \circ d$ . Hint: Use the definition of the Lie derivative for  $p$ -forms and the fact that  $d$  commutes with the pullback.

**26.15.** Let  $M = \mathbb{R}^3$  and let  $f$  be a real-valued function. Let  $\omega = a_i dx^i$  be a one-form and  $\eta = b_1 dx^2 \wedge dx^3 + b_2 dx^3 \wedge dx^1 + b_3 dx^1 \wedge dx^2$  be a two-form on  $\mathbb{R}^3$ . Show that

- (a)  $df$  gives the gradient of  $f$ ,  
 (b)  $d\eta$  gives the divergence of the vector  $\mathbf{B} = (b_1, b_2, b_3)$ , and that  
 (c)  $\nabla \times (\nabla f) = 0$  and  $\nabla \cdot (\nabla \times \mathbf{A}) = 0$  are consequences of  $d^2 = 0$ .

**26.16.** Show that  $i_X$  is an antiderivation with respect to the wedge product.

**26.17.** Given that  $\mathbf{F} = \frac{1}{2} F_{\alpha\beta} dx^\alpha \wedge dx^\beta$ , show that  $\mathbf{F} \wedge (*\mathbf{F}) = |\mathbf{B}|^2 - |\mathbf{E}|^2$ .

**26.18.** Use Equation (26.40) to show that the zeroth component of the relativistic Lorentz force law gives the rate of change of energy due to the electric field, and that the magnetic field does not change the energy.

**26.19.** Derive Equation (26.43).

**26.20.** Given that  $\mathbf{F} = \frac{1}{2} F_{\alpha\beta} dx^\alpha \wedge dx^\beta$ , write the two homogeneous Maxwell's equations,  $\nabla \cdot \mathbf{B} = 0$  and  $\nabla \times \mathbf{E} + \partial \mathbf{B} / \partial t = 0$ , in terms of  $F_{\alpha\beta}$ .

**26.21.** Write the equation

$$F_{\alpha\beta} = A_{\beta,\alpha} - A_{\alpha,\beta} = \frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta}$$

in terms of  $\mathbf{E}$ ,  $\mathbf{B}$ , and vector and scalar potentials.

**26.22.** With  $\mathbf{F} = \frac{1}{2} F_{\alpha\beta} dx^\alpha \wedge dx^\beta$  and  $\mathbf{J} = J_\gamma dx^\gamma$ , show that  $d * \mathbf{F} = 4\pi(*\mathbf{J})$  takes the following form in components:

$$\frac{\partial F^{\alpha\beta}}{\partial x^\beta} = 4\pi J^\alpha,$$

where indices are raised and lowered by  $\text{diag}(-1, -1, -1, 1)$ .

**26.23.** Interpret Theorem 26.5.14 for  $p = 1$  and  $p = 2$  on  $\mathbb{R}^3$ .

**26.24.** Let  $f$  be a function on  $\mathbb{R}^3$ . Calculate  $d * df$ .

**26.25.** Show that current conservation is an automatic consequence of Maxwell's inhomogeneous equation  $d * \mathbf{F} = 4\pi(*\mathbf{J})$ .

## Additional Reading

1. Abraham, R. and Marsden, J. *Foundations of Mechanics*, 2nd ed., Addison-Wesley, 1985. The definitive textbook on classical mechanics presented in the language of differentiable manifolds. Contains a thorough treatment of exterior calculus and symplectic geometry.

2. Bishop, R. and Goldberg, S. *Tensor Analysis on Manifolds*, Dover, 1980.
3. Bott, R. and Tu, L. *Differential Forms in Algebraic Topology*, Springer-Verlag, 1982. Despite its frightening title, the first chapter of this book is actually a very good introduction to tensors.
4. Choquet-Bruhat, Y., DeWitt-Morette, C., and Dillard-Bleick, M. *Analysis, Manifolds, and Physics*, 2nd ed., North-Holland, 1982. A two-volume reference written by mathematical physicists. Excellent for readers already familiar with the subject and seeking detailed applications in physics.
5. Warner, F. *Foundations of Differentiable Manifolds and Lie Groups*, Springer-Verlag, 1983. A formal but readable introduction to differentiable manifolds.



Part VIII

---

# Lie Groups and Their Applications





# Lie Groups and Lie Algebras

The theory of differential equations had flourished to such a level by the 1860s that a systematic study of their solutions became possible. Sophus Lie, a Norwegian mathematician, undertook such a study using the same tool that was developed by Galois and others to study algebraic equations: group theory. The groups associated with the study of differential equations, now called Lie groups, unlike their algebraic counterparts, are uncountably infinite, and, as such, are both intricate and full of far-reaching structures. It was beyond the wildest dream of any 19th-century mathematician to imagine that a concept as abstract as Lie groups would someday find application in the study of the heart of matter. Yet, three of the four fundamental interactions are described by Lie groups, and the fourth one, gravity, is described in a language very akin to the other three.

## 27.1 Lie Groups and Their Algebras

Lie groups are infinite groups that have the extra property that their multiplication law is differentiable. We have seen that the natural setting for differentiation is the structure of a manifold. Thus, Lie groups must have manifold properties as well as group properties.

Lie groups defined **27.1.1. Definition.** *A Lie group  $G$  is a differentiable manifold endowed with a group structure such that the group operation  $G \times G \rightarrow G$  and the map  $G \rightarrow G$  given by  $g \mapsto g^{-1}$  are differentiable. If the dimension of the underlying manifold is  $r$ , we say that  $G$  is an  $r$ -parameter Lie group.*

Because of the dual nature of Lie groups, most of their mapping properties combine those of groups and manifolds. For instance, a Lie group homomor-

phism is a group homomorphism that is also  $\mathcal{C}^\infty$ , and a Lie group isomorphism is a group isomorphism that is also a diffeomorphism.

**27.1.2. Example.**  $GL(\mathcal{V})$  IS A LIE GROUP

$GL(\mathcal{V})$  is a Lie group

As the paradigm of Lie groups, we consider  $GL(\mathcal{V})$ , the set of invertible operators on an  $n$ -dimensional real vector space  $\mathcal{V}$ , and show that it is indeed a Lie group. The set  $\mathcal{L}(\mathcal{V})$  is a vector space of dimension  $n^2$  (Proposition 25.1.1), and therefore, by Example 26.1.7, a manifold of the same dimension. The map  $\det : \mathcal{L}(\mathcal{V}) \rightarrow \mathbb{R}$  is a  $\mathcal{C}^\infty$  map because the determinant, when expressed in terms of a matrix, is a polynomial. In particular, it is continuous. Now note that

$$GL(\mathcal{V}) = \det^{-1}(\mathbb{R} - \{0\})$$

and that  $\mathbb{R} - \{0\}$  is open. It follows that  $GL(\mathcal{V})$  is an open submanifold of  $\mathcal{L}(\mathcal{V})$ . Thus,  $GL(\mathcal{V})$  is an  $n^2$ -dimensional manifold. Choosing a basis  $B$  for  $\mathcal{V}$  and representing operators (points)  $\mathbf{A}$  of  $GL(\mathcal{V})$  as matrices  $(a_{ij})$  in that basis provides a coordinate patch for  $GL(\mathcal{V})$ . We denote this coordinate patch by  $\{x^{ij}\}$ , where  $x^{ij}(\mathbf{A}) = a_{ij}$ .

To show that  $GL(\mathcal{V})$  is a Lie group, we need to prove that if  $\mathbf{A}, \mathbf{B} \in GL(\mathcal{V})$ , then

$$\mathbf{AB} : GL(\mathcal{V}) \times GL(\mathcal{V}) \rightarrow GL(\mathcal{V}) \quad \text{and} \quad \mathbf{A}^{-1} : GL(\mathcal{V}) \rightarrow GL(\mathcal{V})$$

are  $\mathcal{C}^\infty$  maps of manifolds. This is done by showing that the coordinate representations of these maps are  $\mathcal{C}^\infty$ . These representations are simply the matrix representations of operators. Since  $\mathbf{AB}$  is a linear function of elements of the two matrices, it has derivatives of all orders. It follows that  $\mathbf{AB}$  is  $\mathcal{C}^\infty$ . The case of  $\mathbf{A}^{-1}$  is only slightly more complicated. We note that

$$\mathbf{A}^{-1} = \frac{P(a_{ij})}{\det \mathbf{A}}, \quad P(a_{ij}) = \text{a polynomial in } a_{ij}.$$

Thus, since  $\det \mathbf{A}$  is also a polynomial in  $a_{ij}$ , the  $k$ th derivative of  $\mathbf{A}^{-1}$  is of the form  $Q(a_{ij})/(\det \mathbf{A})^k$ , where  $Q$  is another polynomial. The fact that  $\det \mathbf{A} \neq 0$  establishes the  $\mathcal{C}^\infty$  property of  $\mathbf{A}^{-1}$ .

One can similarly show that if  $\mathcal{V}$  is a complex vector space, then  $GL(\mathcal{V})$  is a manifold of dimension  $2n^2$ . ■

**27.1.3. Example.**  $SL(\mathcal{V})$  IS A LIE GROUP

$SL(\mathcal{V})$  is a Lie group

Recall that  $SL(\mathcal{V})$  is the subgroup of  $GL(\mathcal{V})$  whose elements have unit determinant. Since  $\det : GL(\mathcal{V}) \rightarrow \mathbb{R}$  is  $\mathcal{C}^\infty$ , Theorem 26.3.7 and the example after it show that  $SL(\mathcal{V}) = \det^{-1}(1)$  is a submanifold of  $GL(\mathcal{V})$  of dimension  $\dim GL(\mathcal{V}) - \dim \mathbb{R} = n^2 - 1$ . Since it is already a subgroup, we conclude that  $SL(\mathcal{V})$  is also a Lie group (Problem 27.5). Similarly, when  $\mathcal{V}$  is a complex vector space, one can show that  $\dim SL(\mathcal{V}) = 2n^2 - 2$ . ■

**27.1.4. Example.** OTHER EXAMPLES OF LIE GROUPS

The reader may verify the following:

- (a) Any finite-dimensional vector space is a Lie group under vector addition.
- (b) The unit circle  $S^1$ , as a subset of nonzero multiplicative complex numbers is a Lie group under multiplication.
- (c) The product  $G \times H$  of two Lie groups is itself a Lie group with the product manifold structure and the direct product group structure.
- (d)  $GL(n, \mathbb{R})$ , the set of invertible  $n \times n$  matrices, is a Lie group under matrix multiplication.

group of affine motions of  $\mathbb{R}^n$

(e) Let  $G = GL(n, \mathbb{R}) \times \mathbb{R}^n$  be the product manifold. Define the group operation by  $(A, \mathbf{u})(B, \mathbf{v}) \equiv (AB, A\mathbf{v} + \mathbf{u})$ . The reader may verify that this operation indeed defines a group structure on  $G$ . In fact,  $G$  becomes a Lie group, called the **group of affine motions of  $\mathbb{R}^n$** , for if we identify  $(A, \mathbf{u})$  with the affine motion<sup>1</sup>  $\mathbf{x} \mapsto A\mathbf{x} + \mathbf{u}$  of  $\mathbb{R}^n$ , then the group operation in  $G$  is composition of affine motions. We shall study in some detail the Poincaré group, a subgroup of the group of affine motions, in which the matrices are (pseudo) orthogonal. ■

local Lie groups

In calculations, one translates all group operations to the corresponding operations of charts. This is particularly useful when the group multiplication can be defined only locally. One then speaks of an  $r$ -parameter **local Lie group**. To be precise, one considers a neighborhood  $U$  of the origin of  $\mathbb{R}^r$  and defines an *associative* “multiplication”  $m : U \times U \rightarrow \mathbb{R}^r$  and an inversion  $i : U_0 \rightarrow U$  where  $U_0$  is a subset of  $U$ . We therefore write the multiplication as

$$m(\mathbf{a}, \mathbf{b}) = \mathbf{c}, \quad \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^r,$$

where  $\mathbf{a} = (a^1, a^2, \dots, a^r)$ , etc. are coordinates of elements of  $G$ . The coordinates of the identity element of  $G$  are taken to be all zero. Thus,  $m(\mathbf{a}, \mathbf{0}) = \mathbf{a}$  and  $m(\mathbf{a}, i(\mathbf{a})) = \mathbf{0}$ . In component forms,

$$c^k = m^k(\mathbf{a}, \mathbf{b}), \quad m^k(\mathbf{a}, i(\mathbf{a})) = 0, \quad k = 1, 2, \dots, r. \quad (27.1)$$

The fact that  $G$  is a manifold implies that all functions in Equation (27.1) are infinitely differentiable.

**27.1.5. Example.** As an example of a local 1-parameter Lie group, consider the multiplication rule  $m : U \times U \rightarrow \mathbb{R}$  where  $U = \{x \in \mathbb{R} \mid |x| < 1\}$  and

$$m(x, y) = \frac{2xy - x - y}{xy - 1}, \quad x, y \in U.$$

The reader can check that  $m(x, (y, z)) = m((x, y), z)$ , so that the multiplication is associative. Moreover,  $m(0, x) = m(x, 0) = x$  for all  $x \in U$ , and  $i(x) = x/(2x - 1)$ , defined for  $U_0 = \{x \in \mathbb{R} \mid |x| < \frac{1}{2}\}$ . ■

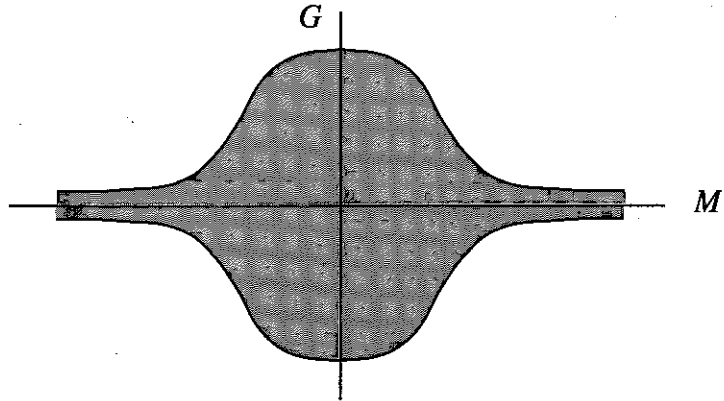
### 27.1.1 Group Action

As mentioned in our discussion of finite groups, the action of a group on a set is more easily conceived than abstract groups. In the case of Lie groups, the natural action is not on an arbitrary set, but on a manifold.

**27.1.6. Definition.** Let  $M$  be a manifold. A **local group of transformations acting on  $M$**  is a (local) Lie group  $G$ , an (open) subset  $U$  with the property  $\{e\} \times M \subset U \subset G \times M$ , and a map  $\Psi : U \rightarrow M$  satisfying the following conditions:

local group of transformations

<sup>1</sup>These consist of a linear transformation followed by a translation.



**Figure 27.1** For small regions of  $M$ , we may be able to include a large portion of  $G$ . However, if we want to include all of  $M$ , as we should, then only a small neighborhood of the identity may be available.

1. If  $(g, P) \in U$ ,  $(h, \Psi(g, P)) \in U$ , and  $(hg, P) \in U$ , then

$$\Psi(h, \Psi(g, P)) = \Psi(hg, P).$$

2.  $\Psi(e, P) = P$  for all  $P \in M$ .

3. If  $(g, P) \in U$ , then  $(g^{-1}, \Psi(g, P)) \in U$  and  $\Psi(g^{-1}, \Psi(g, P)) = P$ .

Normally, we shall denote  $\Psi(g, P)$  by  $g \cdot P$ , or  $gP$ . Then the conditions of the definition above take the simple form

$$\begin{aligned} g \cdot (h \cdot P) &= (gh) \cdot P, & g, h \in G, & P \in M, \\ e \cdot P &= P & \text{for all } P \in M, & \\ g^{-1} \cdot (g \cdot P) &= P, & g \in G, & P \in M, \end{aligned} \tag{27.2}$$

whenever  $g \cdot P$  is defined. Note that the word “local” refers to  $G$  and not  $M$ , i.e., we may have to choose a very small neighborhood of the identity before all the elements of that neighborhood can act on *all* points of  $M$  (see Figure 27.1).

orbit; transitive  
action; stabilizer

All the properties of a group action described in Chapter 23 can be applied here as well. So, one talks about the **orbit** of  $G$  as the collection of points in  $M$  obtained from one another by the action of  $G$ ; a **transitive** action of  $G$  on  $M$  when there is only one orbit; the **stabilizer** of a point of  $M$ , etc. The only extra condition one has to be aware of is that the group action is not defined for all elements of  $G$ , and that a sufficiently small neighborhood of the identity needs to be chosen.

In the old literature, the group action is described in terms of coordinates. Although for calculations this is desirable, it can be very clumsy for formal discussions, as we shall see later. Let  $\mathbf{a} = (a^1, \dots, a^r)$  be a coordinate system

on  $G$  and  $\mathbf{x} = (x^1, \dots, x^n)$  a coordinate system on  $M$ . Then the group action  $\Psi : G \times M \rightarrow M$  becomes a set of  $n$  functions described by

$$\mathbf{x}' = \Psi(\mathbf{a}, \mathbf{x}), \quad \mathbf{x}'' = \Psi(\mathbf{b}, \mathbf{x}') = \Psi(m(\mathbf{b}, \mathbf{a}), \mathbf{x}), \tag{27.3}$$

where  $m$  is the multiplication law of the Lie group written in terms of coordinates as given in Equation (27.1). It is assumed that  $\Psi$  is infinitely differentiable.

**27.1.7. Box.** Equation (27.3) can be used to unravel the multiplication law for the Lie group when the latter is given in terms of transformations.

**27.1.8. Example.** EXAMPLES OF GROUPS OF TRANSFORMATION

(a) The two-dimensional rotation group acts on the  $xy$ -plane as

$$\Phi(\theta, \mathbf{r}) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta).$$

If we write  $\mathbf{r}' = \Phi(\theta_1, \mathbf{r})$  and  $\mathbf{r}'' = \Phi(\theta_2, \mathbf{r}')$ , then a simple calculation shows that

$$\mathbf{r}'' = (x \cos(\theta_1 + \theta_2) - y \sin(\theta_1 + \theta_2), x \sin(\theta_1 + \theta_2) + y \cos(\theta_1 + \theta_2)).$$

With  $\mathbf{r}'' = (m(\theta_1, \theta_2); \mathbf{r})$ , we recognize the “multiplication” law as  $m(\theta_1, \theta_2) = \theta_1 + \theta_2$ . The orbits are circles centered at the origin.

(b) Let  $M = \mathbb{R}^n$ ,  $\mathbf{a}$  a fixed vector in  $\mathbb{R}^n$ , and  $G = \mathbb{R}$ . Define  $\Psi : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  by

$$\Psi(t, \mathbf{x}) = \mathbf{x} + t\mathbf{a}, \quad \mathbf{x} \in \mathbb{R}^n, t \in \mathbb{R}.$$

This group action is globally defined. The orbits are straight lines parallel to  $\mathbf{a}$ . The group translations is the set of **translations** in the direction  $\mathbf{a}$  in  $\mathbb{R}^n$ . The reader may verify that the “multiplication” law is addition of  $t$ 's.

(c) Let  $G = \mathbb{R}^+$  be the multiplicative group of nonzero positive real numbers. Fix real numbers  $\alpha_1, \alpha_2, \dots, \alpha_n$ , not all zero. Define the action of  $G$  on  $\mathbb{R}^n$  by

$$\Psi(\lambda, \mathbf{x}) \equiv \lambda \cdot \mathbf{x} = (\lambda^{\alpha_1} x_1, \dots, \lambda^{\alpha_n} x_n), \quad \lambda \in \mathbb{R}^+, \quad \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

The orbits are obtained by choosing a point in  $\mathbb{R}^n$  and applying  $G$  to it for different  $\lambda$ 's. The result is a curve in  $\mathbb{R}^n$ . For example, if  $n = 2, \alpha_1 = 1$ , and  $\alpha_2 = 2$ , we get, as the orbit containing  $\mathbf{x}_0$  the curve

$$\lambda \cdot \mathbf{x}_0 = (\lambda x_0, \lambda^2 y_0) \Rightarrow y = \frac{y_0}{x_0^2} x^2,$$

which is a parabola going through the origin and the point  $(x_0, y_0)$ . Note that the orbit containing the origin has only one point. This group is called the group of **scale transformations**. The multiplication law is ordinary multiplication of (positive) real numbers.

(d) Let  $G = \mathbb{R}^4$  act on  $M = \mathbb{R}$  by

$$\Phi(\mathbf{a}, x) \equiv \frac{a_1 x + a_2}{a_3 x + a_4}, \quad \mathbf{a} = (a_1, a_2, a_3, a_4), \quad a_1 a_4 - a_2 a_3 \neq 0.$$

scale transformations  
one-dimensional projective group

The reader may verify that this is indeed the action of a group (catch where the condition  $a_1a_4 - a_2a_3 \neq 0$  is used!), and if  $x' = \Phi(\mathbf{b}, x)$  and  $x'' = \Phi(\mathbf{a}, x')$ , then

$$x'' = \frac{(a_1b_1 + a_2b_3)x + a_1b_2 + a_2b_4}{(a_3b_1 + a_4b_3)x + a_3b_2 + a_4b_4},$$

so that the multiplication rule is

$$m(\mathbf{a}, \mathbf{b}) = (a_1b_1 + a_2b_3, a_1b_2 + a_2b_4, a_3b_1 + a_4b_3, a_3b_2 + a_4b_4).$$

This group is called the **one-dimensional projective group**. ■

### 27.1.2 Lie Algebra of a Lie Group

The group property of a Lie group  $G$  provides a natural diffeomorphism on  $G$  that determines a substantial part of its structure.

**27.1.9. Definition.** Let  $G$  be a Lie group and  $g \in G$ . The **left translation by  $g$**  is a diffeomorphism  $L_g : G \rightarrow G$  defined by

$$L_g(h) = gh \quad \forall h \in G.$$

A vector field  $\xi$  on  $G$  is called **left-invariant** if for each  $g \in G$ ,  $\xi$  is  $L_g$ -related to itself; i.e.,<sup>2</sup>

$$L_{g*} \circ \xi = \xi \circ L_g, \quad \text{or} \quad L_{g*}(\xi(h)) = \xi(gh) \quad \forall g, h \in G.$$

The set of left-invariant vector fields on  $G$  is denoted by  $\mathfrak{g}$ . A 1-form whose pairing with a left-invariant vector field gives a constant function on  $G$  is called a **left-invariant 1-form**.

The **right translation by  $g$** ,  $R_g : G \rightarrow G$ , and **right-invariant vector fields and 1-forms** are defined similarly.

The reader may easily check that right and left translations commute:

$$R_g \circ L_h = L_h \circ R_g \quad \forall g, h \in G \tag{27.4}$$

Moreover, if  $\omega|_e$  is a 1-form on  $\mathcal{T}_e(G)$ , then  $\omega \in \Lambda^1(G)$ , given by  $\omega|_g = L_{g^{-1}}^* \omega|_e$ , is a left-invariant 1-form:

$$\omega|_g(\mathbf{X}|_g) = L_{g^{-1}}^* \omega|_e(\mathbf{X}|_g) = \omega|_e(L_{g^{-1}*} \mathbf{X}|_g) = \omega|_e(\mathbf{X}|_{g^{-1}g}) = \omega|_e(\mathbf{X}|_e)$$

independent of  $g$ . It is convenient to have a coordinate representation of  $L_{g*}$ . The coordinate representation of  $L_g$  is simply the multiplication law  $L_g(h) = m(\mathbf{g}, \mathbf{h})$ ,

<sup>2</sup>When there is no danger of confusion, we shall use  $\xi(h)$  for  $\xi|_h$ .

left translation, left-invariant vector fields, left-invariant forms, and their "right" counterparts

where we have used the same symbol for coordinates as for group elements. Equation (26.11) can now be used to write the coordinate representation of  $L_{g^*}$ :

$$L_{g^*} \rightarrow \begin{pmatrix} \partial m^1 / \partial h^1 & \partial m^1 / \partial h^2 & \dots & \partial m^1 / \partial h^r \\ \partial m^2 / \partial h^1 & \partial m^2 / \partial h^2 & \dots & \partial m^2 / \partial h^r \\ \vdots & \vdots & \dots & \vdots \\ \partial m^r / \partial h^1 & \partial m^r / \partial h^2 & \dots & \partial m^r / \partial h^r \end{pmatrix}, \quad (27.5)$$

where all the derivatives in the matrix are evaluated at  $(\mathbf{g}, \mathbf{h})$ .

We have already mentioned in Chapter 26 that  $\mathcal{X}(G)$  is an infinite-dimensional Lie algebra under the Lie bracket "multiplication." In general, no finite-dimensional subspace can be found that—in and of itself—is also an algebra.<sup>3</sup> However, Lie groups are an exception:

**27.1.10. Proposition.** *Let  $G$  be a Lie group and  $\mathfrak{g}$  the set of its left-invariant vector fields. Then  $\mathfrak{g}$  is a real vector space, and the map  $\phi : \mathfrak{g} \rightarrow \mathcal{T}_e(G)$ , defined by  $\phi(\xi) = \xi(e)$ , is a linear isomorphism. Therefore,  $\dim \mathfrak{g} = \dim \mathcal{T}_e(G) = \dim G$ . Furthermore,  $\mathfrak{g}$  is closed under Lie brackets; i.e.,  $\mathfrak{g}$  is a Lie algebra.*

*Proof.* It is clear that  $\mathfrak{g}$  is a real vector space. If  $\phi(\xi) = \phi(\eta)$  for  $\xi, \eta \in \mathfrak{g}$ , then

$$\xi(g) = L_{g^*}(\xi(e)) = L_{g^*}(\eta(e)) = \eta(g) \quad \forall g \in G \Rightarrow \xi = \eta.$$

This shows that  $\phi$  is injective. To show that  $\phi$  is surjective, suppose that  $\mathbf{v} \in \mathcal{T}_e(G)$  and define the vector field  $\xi$  on  $G$  by  $\xi(g) = L_{g^*}(\mathbf{v})$  for all  $g \in G$ . Then  $\phi(\xi) = \mathbf{v}$  and  $\xi \in \mathfrak{g}$ , because

$$L_{g^*} \circ \xi(h) = L_{g^*} \circ L_{h^*}(\mathbf{v}) = L_{gh^*}(\mathbf{v}) = \xi(gh) \equiv \xi(L_g h) = \xi \circ L_g(h).$$

This proves the first part of the proposition. The second part follows immediately from the definition of a left-invariant vector field and Theorem 26.4.4.  $\square$

The flow of  $\xi$  at  $g \in G$  can be shown to be

$$F_t = g \exp(t\xi) = R_{\exp(t\xi)}g. \quad (27.6)$$

Indeed, let  $\mathbf{X}_\xi$  be the vector field associated with this flow. The action of this vector field on a function  $f$  is

$$\mathbf{X}_\xi|_g(f) = \left. \frac{d}{dt} (f(g \exp t\xi)) \right|_{t=0}.$$

Therefore,

$$\begin{aligned} (L_{h^*} \mathbf{X}_\xi|_g)(f) &= \mathbf{X}_\xi|_g(f \circ L_h) = \left. \frac{d}{dt} (f \circ L_h(g \exp t\xi)) \right|_{t=0} \\ &= \left. \frac{d}{dt} (f(hg \exp t\xi)) \right|_{t=0} = \mathbf{X}_\xi(hg)(f) = [\mathbf{X}_\xi \circ L_h(g)](f). \end{aligned}$$

<sup>3</sup>Such a subspace is called a **subalgebra**.

Since this is true for all  $f$  and  $g$ , and  $X_\xi|_e = \xi(e)$ , we conclude that  $X_\xi$  is the unique left-invariant vector field corresponding to  $\xi(e)$ .

The Lie algebra of a Lie group

**27.1.11. Definition.** The Lie algebra of the Lie group  $G$  is the Lie algebra  $\mathfrak{g}$  of left-invariant vector fields on  $G$ . Sometimes we think of  $\xi$  as a vector in  $\mathcal{T}_e(G)$ . In that case, we denote by  $X_\xi$  the left-invariant vector field whose value at the identity is  $\xi$ .

The isomorphism of  $\mathfrak{g}$  with  $\mathcal{T}_e(G)$  induces a Lie bracket on  $\mathcal{T}_e(G)$  and turns it into a Lie algebra. In many cases of physical interest, it is this interpretation of the Lie algebra of  $G$  that is most useful. If two groups stand in some algebraic relation to one another, their Lie algebras will inherit such relations. More precisely, let  $G$  and  $H$  be Lie groups with Lie algebras  $\mathfrak{g}$  and  $\mathfrak{h}$ , respectively. Suppose  $\phi : G \rightarrow H$  is a Lie group homomorphism. Then identifying  $\mathfrak{g}$  with  $\mathcal{T}_e(G)$  and  $\mathfrak{h}$  with  $\mathcal{T}_e(H)$ , and using Theorem 26.4.4, we conclude that  $\phi_* : \mathfrak{g} \rightarrow \mathfrak{h}$  is a Lie algebra homomorphism, i.e., it preserves the Lie brackets:

$$\phi_*[\xi, \eta] = [\phi_*\xi, \phi_*\eta] \quad \forall \xi, \eta \in \mathfrak{g}. \quad (27.7)$$

In particular, if  $\phi$  is a Lie group isomorphism, then  $\phi_*$  is a Lie algebra isomorphism.

**27.1.12. Example.** Let  $\mathcal{V}$  be a complex vector space with its general linear group  $GL(\mathcal{V})$ , a  $2n^2$ -dimensional Lie group. Recall that  $GL(\mathcal{V})$  is an open submanifold of  $\mathcal{L}(\mathcal{V})$ . By Equation (26.4),  $\mathcal{T}_e(GL(\mathcal{V})) = \mathcal{T}_e(\mathcal{L}(\mathcal{V}))$ , where  $e$  is the unit operator. If we identify  $\mathcal{T}_e(\mathcal{L}(\mathcal{V}))$  with  $\mathcal{L}(\mathcal{V})$  [see the box after Equation (26.4)], we may conclude that  $\mathcal{L}(\mathcal{V})$ , which we denote by  $\mathfrak{gl}(\mathcal{V})$  in the present context, is the Lie algebra of  $GL(\mathcal{V})$  in which the Lie bracket is the commutator. We use the notation  $\mathbf{A}(t)$  for a curve in  $GL(\mathcal{V})$  and  $\dot{\mathbf{A}}$  for the vector tangent to the curve. ■

It is instructive to construct the coordinate representation of vector fields on  $GL(\mathcal{V})$ . Let  $f : GL(\mathcal{V}) \rightarrow \mathbb{R}$  be a function and  $\dot{\mathbf{A}}$  a vector field. Then, we have

$$\dot{\mathbf{A}}(f) = \frac{d}{dt}(f(\mathbf{A}(t))) = \frac{da_{ij}}{dt} \frac{\partial f}{\partial x^{ij}},$$

or, since  $f$  is arbitrary,

$$\dot{\mathbf{A}} = \frac{da_{ij}}{dt} \frac{\partial}{\partial x^{ij}} \equiv \dot{a}_{ij} \frac{\partial}{\partial x^{ij}} \equiv \frac{d\mathbf{A}}{dt}(t),$$

where summation over repeated indices is understood and we introduced  $d\mathbf{A}/dt$  as an abbreviation for  $\dot{a}_{ij}(\partial/\partial x^{ij})$ . However, the one-to-one correspondence between matrices and operators makes this more than just an abbreviation. Indeed, we can interpret  $d\mathbf{A}/dt$  as the derivative of  $\mathbf{A}$  and perform such differentiation whenever it is possible. The equation above states that

**27.1.13. Box.** To obtain the matrix elements (coordinates) of the operator  $\dot{\mathbf{A}}$ , one differentiates the  $t$ -dependent elements of the (matrix representation of the) operator  $\mathbf{A}(t)$ .



Of particular interest are the left invariant vector fields, or equivalently, the vectors belonging to  $\mathcal{T}_e(GL(\mathcal{V}))$ . This amounts to substituting  $t = 0$  in the formulas above. Thus, if  $\dot{\mathbf{A}} \in \mathcal{T}_e(GL(\mathcal{V}))$ ,

$$\dot{\mathbf{A}} = \dot{a}_{ij}(0) \frac{\partial}{\partial x^{ij}} \equiv \frac{d\mathbf{A}}{dt}(0). \tag{27.8}$$

For the product of two operators, we get

$$\begin{aligned} \widehat{\mathbf{AB}} &= \left. \frac{d}{dt} f(\mathbf{A}(t)\mathbf{B}(t)) \right|_{t=0} = \left. \frac{d}{dt} (a_{ik}b_{kj}) \right|_{t=0} \frac{\partial}{\partial x^{ij}} \\ &= (\dot{a}_{ik}(0) \underbrace{b_{kj}(0)}_{\delta_{kj}} + \underbrace{a_{ik}(0)}_{\delta_{ik}} \dot{b}_{kj}(0)) \frac{\partial}{\partial x^{ij}} \\ &= \dot{a}_{ij}(0) \frac{\partial}{\partial x^{ij}} + \dot{b}_{ij}(0) \frac{\partial}{\partial x^{ij}} = \frac{d\mathbf{A}}{dt}(0) + \frac{d\mathbf{B}}{dt}(0). \end{aligned} \tag{27.9}$$

Many of the Lie groups used in physics are subgroups of  $GL(\mathcal{V})$ . A characterization of the Lie algebras of these subgroups is essential for understanding the subgroups themselves and applying them to physical situations. These subgroups are typically defined in terms of maps  $\phi : GL(\mathcal{V}) \rightarrow M$  for which  $M$  is a manifold and  $\phi_*$  is surjective. To construct the Lie algebra of subgroups of  $GL(\mathcal{V})$ , we need to concentrate on the map  $\phi_{*e}$  as defined on  $\mathcal{T}_e(GL(\mathcal{V}))$ .

An important map is  $\det : GL(\mathcal{V}) \rightarrow \mathbb{C}$  for a complex vector space  $\mathcal{V}$ . We are interested in evaluating the map  $\det_* : \mathcal{T}_e(GL(\mathcal{V})) \rightarrow \mathcal{T}_1(\mathbb{C})$  in which we consider  $\mathbb{C} \cong \mathbb{R}^2$  to be a manifold. For an operator  $\dot{\mathbf{A}} \in \mathcal{T}_e(GL(\mathcal{V})) \cong \mathfrak{gl}(\mathcal{V})$  and a complex-valued function, we have

Differential of the determinant map is the trace:  $\det_* = \text{tr}$

$$\begin{aligned} \det_*(\dot{\mathbf{A}})f &\equiv \left. \frac{d}{dt} f(\det \mathbf{A}(t)) \right|_{t=0} \equiv \frac{dx}{dt} \frac{\partial f}{\partial x} + \frac{dy}{dt} \frac{\partial f}{\partial y} \\ &= \left. \frac{d}{dt} \text{Re} \det \mathbf{A}(t) \right|_{t=0} \frac{\partial f}{\partial x} + \left. \frac{d}{dt} \text{Im} \det \mathbf{A}(t) \right|_{t=0} \frac{\partial f}{\partial y} \\ &= \text{Re} \text{tr} \dot{\mathbf{A}} \frac{\partial f}{\partial x} + \text{Im} \text{tr} \dot{\mathbf{A}} \frac{\partial f}{\partial y}, \end{aligned}$$

where we used Equation (3.28). Since  $f$  is arbitrary and  $\{\partial/\partial x, \partial/\partial y\}$  can be identified with  $\{1, i\}$ , we have

$$\det_*(\dot{\mathbf{A}}) = \text{tr} \dot{\mathbf{A}}. \tag{27.10}$$

**27.1.14. Example.** LIE ALGEBRA OF  $SL(\mathcal{V})$

The special linear group  $SL(\mathcal{V})$  is characterized by the fact that all its elements have unit determinant.

**27.1.15. Box.** The Lie algebra  $\mathfrak{sl}(\mathcal{V})$  of the special linear group is the set of all traceless operators.

This is because if we use (27.10) and (26.12) and the fact that  $SL(\mathcal{V}) = \det^{-1}(1)$ , we can conclude that  $\det_*(\dot{\mathbf{A}}) = \text{tr } \dot{\mathbf{A}} = 0$  for all  $\dot{\mathbf{A}} \in \mathfrak{sl}(\mathcal{V})$ . ■

**27.1.16. Example.** LIE ALGEBRAS OF UNITARY AND RELATED GROUPS

Let us first show that the set of unitary operators on  $\mathcal{V}$ , denoted by  $U(\mathcal{V})$ , is a Lie subgroup of  $GL(\mathcal{V})$ , called the **unitary group** of  $\mathcal{V}$ . Consider the map  $\psi : GL(\mathcal{V}) \rightarrow \mathbb{H}$ , where  $\mathbb{H}$  is the set of Hermitian operators considered as a vector space (therefore, a manifold) over the reals, defined by  $\psi(\mathbf{A}) = \mathbf{A}\mathbf{A}^\dagger$ . Using Equation (27.9), the reader may verify that  $\psi_*$  is surjective and

$$\psi_*(\dot{\mathbf{A}}) = \dot{\mathbf{A}} + \dot{\mathbf{A}}^\dagger. \tag{27.11}$$

It follows from Theorem 26.3.7 that  $U(\mathcal{V}) \equiv \psi^{-1}(\mathbf{1})$  is a subgroup of  $GL(\mathcal{V})$ . Using Equation (26.12), we conclude that  $\psi_*(\dot{\mathbf{A}}) = \dot{\mathbf{A}} + \dot{\mathbf{A}}^\dagger = 0$  for all  $\dot{\mathbf{A}} \in \mathfrak{u}(\mathcal{V})$ , i.e.,

**27.1.17. Box.** *The Lie algebra  $\mathfrak{u}(\mathcal{V})$  of the unitary group is the set of all anti-hermitian operators.*

By counting the number of independent *real* parameters of a matrix representing a hermitian operator, we can conclude that  $\dim \mathbb{H} = n^2$ . It follows from Theorem 26.3.7 that  $\dim U(\mathcal{V}) = n^2$ .

The intersection of  $SL(\mathcal{V})$  and  $U(\mathcal{V})$ , denoted by  $SU(\mathcal{V})$ , is called the **special unitary group**. Its Lie algebra  $\mathfrak{su}(\mathcal{V})$  consists of anti-hermitian traceless operators. The reader may check that  $\dim SU(\mathcal{V}) = n^2 - 1$ . When the vector space is  $\mathbb{C}^n$ , we write  $U(n)$  and  $SU(n)$  instead of  $U(\mathbb{C}^n)$  and  $SU(\mathbb{C}^n)$ .

If we restrict ourselves to real vector spaces, then unitary and special unitary groups become the **orthogonal group**  $O(\mathcal{V})$  and **special orthogonal group**  $SO(\mathcal{V})$ , respectively. Their algebras consist of antisymmetric and traceless antisymmetric operators, respectively. When  $\mathcal{V} = \mathbb{R}^n$ , we use the notation  $O(n)$  and  $SO(n)$ . ■

Let  $\mathbf{X}$  be a vector field on  $G$ . We know from our discussion of flows that  $\mathbf{X}$  has a flow  $F_t \equiv \exp(t\mathbf{X})$  at every point  $g$  of  $G$  with  $-\epsilon < t < \epsilon$ . Now, since  $F_t(g) \neq g$  is in  $G$ , it follows from the group property of  $G$  that  $(F_t)^n(g) = F_{nt}(g) \in G$  for all  $n$ . This shows that the flow of every vector field on a Lie group is defined for all  $t \in \mathbb{R}$ , i.e., all vector fields on a Lie group are complete. Now consider  $\mathfrak{g}$  as a vector space *and manifold* and define a map  $\exp : \mathfrak{g} \rightarrow G$  that is simply the flow evaluated at  $t = 1$ . It can be shown that the following result holds ([Warn 83, pp. 103–104]):

**27.1.18. Theorem.**  $\exp : \mathfrak{g} \rightarrow G$ , called the **exponential map**, is a diffeomorphism of a neighborhood of the origin of  $\mathfrak{g}$  with a neighborhood of the identity element of  $G$ .

This theorem states that in a neighborhood of the identity element, a Lie group, as a manifold, “looks like” its tangent space there. In particular,

**27.1.19. Box.** *Two Lie groups that have identical Lie algebras are locally diffeomorphic.*

**27.1.20. Example.** WHY  $\exp$  IS CALLED THE EXPONENTIAL MAP

Let  $\mathcal{V}$  be a finite-dimensional vector space and  $\mathbf{A} \in \mathfrak{gl}(\mathcal{V})$ . Define, as in Chapter 2,

$$e^{t\mathbf{A}} = \sum_{k=0}^{\infty} \frac{t^k \mathbf{A}^k}{k!} = \mathbf{1} + t\mathbf{A} + \dots$$

and note that

$$\frac{d}{dt} e^{t\mathbf{A}} = \mathbf{A} e^{t\mathbf{A}} \Rightarrow \left. \frac{d}{dt} e^{t\mathbf{A}} \right|_{t=0} = \mathbf{A}.$$

Furthermore,

$$\begin{aligned} e^{t\mathbf{A}} e^{s\mathbf{A}} &= \sum_{k=0}^{\infty} \frac{t^k \mathbf{A}^k}{k!} \sum_{n=0}^{\infty} \frac{s^n \mathbf{A}^n}{n!} = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{t^k s^n}{k! n!} \mathbf{A}^{k+n} \\ &= \sum_{m=0}^{\infty} \underbrace{\left( \sum_{n=0}^m \frac{t^{m-n} s^n}{(m-n)! n!} \right)}_{=(t+s)^m / m!} \mathbf{A}^m = e^{(t+s)\mathbf{A}}. \end{aligned}$$

It follows that  $e^{t\mathbf{A}}$  has all the properties expected of the flow of the vector field  $\mathbf{A}$ . ■

The exponential map has some important properties that we shall have occasion to use later. The first of these properties is the content of the following proposition, whose proof is left as an exercise for the reader.

**27.1.21. Proposition.** *Let  $\phi : H \rightarrow G$  be a Lie group homomorphism. Then, for all  $\eta \in \mathfrak{h}$ , we have  $\phi(\exp_H \eta) = \exp_G(\phi_* \eta)$ .*

For every  $g \in G$ , let  $I_g \equiv R_g^{-1} \circ L_g$ . The reader may readily verify that  $I_g$ , which takes  $x \in G$  to  $gxg^{-1} \in G$ , is an isomorphism of  $G$ , i.e.,  $I_g(xy) = I_g(x)I_g(y)$  and  $I_g$  is bijective. It is called the **inner automorphism** associated with  $g$ .

**27.1.22. Definition.** *The Lie algebra isomorphism  $I_{g*} = R_{g*}^{-1} \circ L_{g*} : \mathfrak{g} \rightarrow \mathfrak{g}$  is denoted by  $Ad_g$  and is called the **adjoint map** associated with  $g$ .*

Using Proposition 27.1.21, we have the following corollary.

**27.1.23. Corollary.**  $\exp(Ad_g \xi) = I_g \exp \xi = g \exp \xi g^{-1}$  for all  $\xi \in \mathfrak{g}$  and  $g \in G$ .

Let  $\{\xi_i\}$  be a basis for the (finite-dimensional) Lie algebra of the Lie group  $G$ . The Lie bracket of two basis vectors, being itself a left-invariant vector field, can be written as a linear combination of  $\{\xi_i\}$ :

$$[\xi_i, \xi_j] = \sum_{k=1}^n c_{ij}^k \xi_k.$$

On a general manifold,  $c_{ij}^k$  will depend on the point at which the fields are being evaluated. However, on Lie groups, they are independent of the point, as the following manipulation shows:

$$\begin{aligned} [\xi_i(g), \xi_j(g)] &= [L_{g*}\xi_i(e), L_{g*}\xi_j(e)] = L_{g*}[\xi_i(e), \xi_j(e)] \\ &= L_{g*} \sum_{k=1}^n c_{ij}^k(e) \xi_k(e) = \sum_{k=1}^n c_{ij}^k(e) L_{g*}\xi_k(e) = \sum_{k=1}^n c_{ij}^k(e) \xi_k(g). \end{aligned}$$

Lie's second theorem Therefore, the value of  $c_{ij}^k$  at any point  $g \in G$  is the same as its value at the identity, i.e.,  $c_{ij}^k$  is a constant. This statement is called **Lie's second theorem**.

structure constants of a Lie algebra **27.1.24. Definition.** Let  $\{\xi_i\}_{i=1}^n$  be a basis for the Lie algebra  $\mathfrak{g}$  of the Lie group  $G$ . Then

$$[\xi_i(g), \xi_j(g)] = \sum_{k=1}^n c_{ij}^k \xi_k(g), \tag{27.12}$$

where  $c_{ij}^k$ , which are independent of  $g$ , are called the **structure constants** of  $G$ .

Lie's third theorem The structure constants satisfy certain relations that are immediate consequences of the commutation relations. The antisymmetry of the Lie bracket and the Jacobi identity lead directly to

$$\begin{aligned} c_{\rho\sigma}^{\kappa} &= -c_{\sigma\rho}^{\kappa}, \\ c_{\rho\sigma}^{\kappa} c_{\kappa\mu}^{\nu} + c_{\sigma\mu}^{\kappa} c_{\kappa\rho}^{\nu} + c_{\mu\rho}^{\kappa} c_{\kappa\sigma}^{\nu} &= 0. \end{aligned} \tag{27.13}$$

The fact that  $\{c_{\sigma\rho}^{\kappa}\}$  obey Equation (27.13) is the content of **Lie's third theorem**.

### 27.1.3 Infinitesimal Action

The action  $\Phi : G \times M \rightarrow M$  of a Lie group on a manifold  $M$  induces a homomorphism of its algebra with  $\mathcal{X}(M)$ . If  $\xi \in \mathfrak{g}$ , then  $\exp(t\xi) \in G$  can act on  $M$  at a point  $P$  to produce a curve  $\gamma(t) = \exp(t\xi) \cdot P$  going through  $P$ . The tangent to this curve at  $P$  is defined to be the image of this homomorphism.

infinitesimal generators of an action **27.1.25. Definition.** Let  $\Phi : G \times M \rightarrow M$  be an action. If  $\xi \in \mathfrak{g}$ , then

$\Phi(\exp t\xi, P)$  is a flow on  $M$ . The corresponding vector field on  $M$  given by

$$\xi_M|_P \equiv \xi_M(P) \equiv \left. \frac{d}{dt} \Phi(\exp t\xi, P) \right|_{t=0}$$

is called the *infinitesimal generator* of the action induced by  $\xi$ .

Infinitesimal generators of representations of  $G$  form a representation of  $\mathfrak{g}$ .

In particular,

**27.1.26. Box.** If  $M$  happens to be a vector space, and the action a representation as given in Box 24.1.3, then the infinitesimal generators constitute a representation of the Lie algebra of the group.

**27.1.27. Example.** One can think of left translation on a Lie group  $G$  as an action of  $G$  on itself. Let  $\Phi : G \times G \rightarrow G$  be given by  $\Phi(g, h) = L_g(h)$ . Then Definition 27.1.25 gives

$$\xi_G(g) = \left. \frac{d}{dt} \Phi(\exp t\xi, g) \right|_{t=0} = \left. \frac{d}{dt} \exp t\xi g \right|_{t=0} = \left. \frac{d}{dt} R_g(\exp t\xi) \right|_{t=0} = R_{g*}\xi$$

by the first equation in (26.24). It follows that  $\xi_G$  is *right-invariant*. Indeed,

$$\xi_G \circ R_h(g) = \xi_G(gh) = (R_{gh})_*\xi = (R_h \circ R_g)_*\xi = R_{h*} \circ R_{g*}\xi = R_{h*} \circ \xi_G(g).$$

Since this holds for all  $g \in G$ , it follows that  $\xi_G \circ R_h = R_{h*} \circ \xi_G$ , demonstrating that  $\xi_G$  is right-invariant. ■

The adjoint map of Definition 27.1.22 induces a natural action on the Lie algebra  $\mathfrak{g}$  with some important properties that we now explore. Define the **adjoint action**  $\Phi : G \times \mathfrak{g} \rightarrow \mathfrak{g}$  of  $G$  on  $\mathfrak{g} \cong \mathcal{T}_e(G)$  by  $\Phi(g, \xi) = Ad_g(\xi)$ . We claim that the infinitesimal generator  $\xi_{\mathfrak{g}}$  is  $\text{ad}\xi$ , where  $\text{ad}\xi(\eta) \equiv [\xi, \eta]$ . In fact,

$$\begin{aligned} \xi_{\mathfrak{g}}(\eta) &= \left. \frac{d}{dt} \Phi(\exp t\xi, \eta) \right|_{t=0} = \left. \frac{d}{dt} Ad_{\exp t\xi}(\eta) \right|_{t=0} \\ &= \left. \frac{d}{dt} R_{\exp t\xi}^{-1} \circ L_{\exp t\xi}(\eta) \right|_{t=0} = \left. \frac{d}{dt} F_{t*}^{-1} \eta(\exp t\xi) \right|_{t=0} \\ &= L_{\xi}(\eta) = [\xi, \eta] \equiv \text{ad}\xi(\eta), \end{aligned} \tag{27.14}$$

where we used Equation (27.6) as well as the definition of Lie derivative, Equation (26.31). If  $\Phi : G \times M \rightarrow M$  is an action, then  $\Phi_g : M \rightarrow M$ , defined by  $\Phi_g(P) = \Phi(g, P)$ , is a diffeomorphism of  $M$ . Consequently,  $\Phi_{g*} : \mathcal{T}_P(M) \rightarrow \mathcal{T}_{g.P}(M)$  is an isomorphism for every  $P \in M$  whose inverse is  $\Phi_{g*}^{-1} = \Phi_{g^{-1}*}$ .

**27.1.28. Proposition.** Let  $\Phi : G \times M \rightarrow M$  be an action. Then for every  $g \in G$  and  $\xi, \eta \in \mathfrak{g}$ , we have

$$(Ad_g \xi)_M = \Phi_{g*}^{-1} \xi_M \quad \text{and} \quad [\xi_M, \eta_M] = -[\xi, \eta]_M.$$

*Proof.* Let  $P$  be any point in  $M$ . Then,

$$\begin{aligned}
 (Ad_g \xi)_M(P) &= \left. \frac{d}{dt} \Phi(\exp t Ad_g \xi, P) \right|_{t=0} && \text{(by Definition 27.1.25)} \\
 &= \left. \frac{d}{dt} \Phi(g(\exp t \xi) g^{-1}, P) \right|_{t=0} && \text{(by Corollary 27.1.23)} \\
 &= \left. \frac{d}{dt} \Phi(g(\exp t \xi), \Phi_{g^{-1}}(P)) \right|_{t=0} && \text{(by definition of action)} \\
 &= \left. \frac{d}{dt} \Phi_g \circ \Phi(\exp t \xi, \Phi_{g^{-1}}(P)) \right|_{t=0} && \text{(by definition of } \Phi_g) \\
 &= \Phi_{g*} |_{\Phi_{g^{-1}}(P)} \left. \frac{d}{dt} \Phi(\exp t \xi, \Phi_{g^{-1}}(P)) \right|_{t=0} && \text{[by (26.24)]} \\
 &= \Phi_{g*} |_{\Phi_{g^{-1}}(P)} \xi_M(\Phi_{g^{-1}}(P)) = (\Phi_{g*} \xi_M)(P). && \text{[by (26.28)]}
 \end{aligned}$$

The second part of the proposition follows by replacing  $g$  with  $\exp t \eta$ , so that

$$(Ad_{\exp t \eta} \xi)_M = \Phi_{\exp t \eta*} \xi_M = \Phi_{\exp t(-\eta)*}^{-1} \xi_M.$$

Differentiate both sides with respect to  $t$  and note that the LHS gives  $[\eta, \xi]_M$ . The derivative of the RHS is the Lie derivative of  $\xi_M$  with respect to  $-\eta_M$ , which is  $-[\eta_M, \xi_M]$ . □

As mentioned earlier, a Lie group action is usually described in terms of the parameters of the group, which are simply coordinate functions on the group  $G$ , as well as coordinate functions on the manifold  $M$ . The infinitesimal generators, being vector fields on  $M$ , will then be expressed as a linear combination of coordinate frames.

comparison of  
coordinate  
manipulations with  
geometric  
(coordinate-free)  
analysis

In the older literature, no mention of the manifold structure is made. A Lie group is *defined* in terms of multiplication functions and other functions that represent the action of the group on the manifold. Thus, an  $r$ -parameter Lie group  $G$  is a collection of two sets of functions,  $m^\rho : \mathbb{R}^r \rightarrow \mathbb{R}$ ,  $\rho = 1, 2, \dots, r$ , representing the group multiplication, and  $\phi^i : \mathbb{R}^r \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, 2, \dots, n$ , representing the action of  $G$  on the  $n$ -dimensional manifold  $M$ . We sketch the procedure below, leaving most of the calculations as exercises for the reader. As we develop the theory, the reader is urged to compare this “coordinate-dependent” procedure with the “geometric” procedure—which does not use coordinates—described so far.

The action of the group is described by the coordinate transformations<sup>4</sup>

$$\begin{aligned}
 x'_i &= \phi_i(a_1, \dots, a_r; x_1, \dots, x_n), & i &= 1, \dots, n, \\
 x_i &= \phi_i(0, \dots, 0; x_1, \dots, x_n), & &
 \end{aligned} \tag{27.15}$$

<sup>4</sup>We use subscripts for coordinate functions here for typographical convenience.

as well as the group multiplication properties

$$\begin{aligned}c_\rho &= m_\rho(a_1, \dots, a_r; b_1, \dots, b_r), \quad \rho = 1, \dots, r, \\a_\rho &= m_\rho(0, \dots, 0; a_1, \dots, a_r) = m_\rho(a_1, \dots, a_r; 0, \dots, 0), \\m_\rho(\mathbf{a}; m(\mathbf{b}; \mathbf{c})) &= m_\rho(m(\mathbf{a}; \mathbf{b}); \mathbf{c}).\end{aligned}\tag{27.16}$$

Equation (27.15) is to be interpreted as a rule that takes the second set of arguments and transforms them via the first set into the LHS. Now suppose that we translate from  $x'_i$  to a neighboring point  $x'_i + dx'_i$  via a set of group parameters  $\{\delta a_\rho\}_{\rho=1}^r$ . We can also get to  $x'_i + dx'_i$  from  $x_i$  via a new set of parameters,<sup>5</sup> which have to be slightly different from  $\{a_\rho\}_{\rho=1}^r$ , say  $\{a_\rho + da_\rho\}_{\rho=1}^r$ . We then have

$$\begin{aligned}x'_i + dx'_i &= \phi_i(\delta a_1, \dots, \delta a_r; x'_1, \dots, x'_n), \\x'_i + dx'_i &= \phi_i(a_1 + da_1, \dots, a_r + da_r; x_1, \dots, x_n), \\a_\rho + da_\rho &= m_\rho(a_1, \dots, a_r; \delta a_1, \dots, \delta a_r),\end{aligned}\tag{27.17}$$

and, with summation over repeated indices understood,

$$\begin{aligned}dx'_i &= \left. \frac{\partial \phi_i(\mathbf{a}; \mathbf{x}')}{\partial a_\kappa} \right|_{\mathbf{a}=0} \delta a_\kappa \equiv u_{i\kappa}(\mathbf{x}') \delta a_\kappa, \\da_\lambda &= \left. \frac{\partial m_\lambda(\mathbf{a}; \mathbf{b})}{\partial a_\kappa} \right|_{\mathbf{b}=0} \delta a_\kappa \equiv \theta_{\lambda\kappa}(\mathbf{a}) \delta a_\kappa.\end{aligned}\tag{27.18}$$

Inverting the second equation and substituting the resulting  $\delta a$ 's in the first equation yields

$$dx'_i = u_{i\kappa}(\mathbf{x}') \theta_{\kappa\lambda}^{-1}(\mathbf{a}) da_\lambda, \quad \text{or} \quad dx_i = u_{i\kappa}(\mathbf{x}) \theta_{\kappa\lambda}^{-1}(\mathbf{a}) da_\lambda,$$

where in the last equation, we changed the free coordinate variable on both sides. It then follows that

$$\frac{\partial x_i}{\partial a_\lambda} = \sum_{\kappa=1}^r u_{i\kappa}(\mathbf{x}) \theta_{\kappa\lambda}^{-1}(\mathbf{a}).\tag{27.19}$$

Lie's first theorem Equation (27.19) and establishing that  $u_{i\kappa}$  is  $C^\infty$  is the content of **Lie's first theorem**.

The change of an arbitrary function  $f(\mathbf{x})$  due to an infinitesimal transformation is

$$df = \frac{\partial f}{\partial x_i} dx_i = \frac{\partial f}{\partial x_i} u_{i\kappa}(\mathbf{x}) \delta a_\kappa = \delta a_\kappa \left( u_{i\kappa}(\mathbf{x}) \frac{\partial}{\partial x_i} \right) f.$$

<sup>5</sup>Here we are assuming that the action of the group is **transitive**, i.e., that every point of the manifold can be connected to any other point via a transformation.

This suggests calling

$$\mathbf{X}_\kappa = \sum_{i=1}^n u_{i\kappa}(\mathbf{x}) \frac{\partial}{\partial x_i} \tag{27.20}$$

infinitesimal generators as vector fields on  $M$

the **infinitesimal generators** of the Lie group. The commutator of two of these generators is

$$[\mathbf{X}_\rho, \mathbf{X}_\sigma] = \left[ u_{i\rho} \frac{\partial u_{j\sigma}}{\partial x_i} - u_{i\sigma} \frac{\partial u_{j\rho}}{\partial x_i} \right] \frac{\partial}{\partial x_j}. \tag{27.21}$$

This commutator does not appear to be similar to the one in Definition 27.1.24, which is necessary if the generators are to form a Lie algebra. However, through a long and tortuous manipulation, outlined in Problem 27.9, one can show that

$$[\mathbf{X}_\rho, \mathbf{X}_\sigma] = c_{\rho\sigma}^\kappa \mathbf{X}_\kappa \tag{27.22}$$

where  $c_{\rho\sigma}^\kappa$  are constants.

One can also obtain this same result by the much simpler method of applying Proposition 27.1.28 to both sides of Equation (27.12):

$$[(\xi_i)_M, (\xi_j)_M] = -[\xi, \eta]_M = - \left( \sum_{k=1}^n c_{ij}^k \xi_k \right)_M = - \sum_{k=1}^n c_{ij}^k (\xi_k)_M.$$

This equation is equivalent to (27.22) if we identify the  $\mathbf{X}_\rho$ 's with the  $(\xi_i)_M$ 's and ignore the irrelevant minus sign.

The reader has hopefully been able to appreciate the power and elegance of the geometric approach to Lie groups and Lie algebras. The above illustration (Problem 27.9) brings out the tedium and the error-prone procedure of obtaining group-theoretic results through coordinate manipulations, a procedure used in the old literature including the work of Sophus Lie himself. Although such calculations are inevitable in practice, where most Lie groups are given in terms of parameters, they are not suitable for obtaining formal results.

**27.1.29. Example.** The two-dimensional rotation group  $SO(2)$  is a 1-parameter Lie group defined by

$$\begin{aligned} x'_1 &\equiv \phi_1(x_1, x_2; \theta) = x_1 \cos \theta - x_2 \sin \theta, \\ x'_2 &\equiv \phi_2(x_1, x_2; \theta) = x_1 \sin \theta + x_2 \cos \theta. \end{aligned}$$

Using Equation (27.20), we find the (only) generator of this group:

$$\mathbf{X} = u_i \frac{\partial}{\partial x_i} \quad \text{where} \quad u_i = \left. \frac{\partial \phi_i}{\partial \theta} \right|_{\theta=0}.$$

Explicitly, we have

$$\begin{aligned} u_1 &= \left. \frac{\partial \phi_1}{\partial \theta} \right|_{\theta=0} = (-x_1 \sin \theta - x_2 \cos \theta)|_{\theta=0} = -x_2, \\ u_2 &= \left. \frac{\partial \phi_2}{\partial \theta} \right|_{\theta=0} = (x_1 \cos \theta - x_2 \sin \theta)|_{\theta=0} = x_1. \end{aligned}$$



and

$$\mathbf{X} = u_1 \frac{\partial}{\partial x_1} + u_2 \frac{\partial}{\partial x_2} = -x_2 \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2} = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}.$$

The reader recognizes this, within a factor of  $i$ , as the  $z$ -component of the angular momentum operator. In general,

**27.1.30. Box.** *Angular momentum operators are the infinitesimal generators of rotation.*

Inclusion of the other two rotations about the  $x$ -axis and the  $y$ -axis completes the set of infinitesimal generators of the rotation group in three dimensions. ■

The action of a Lie group on  $M$  can be reconstructed from its infinitesimal action. The flow of  $\mathbf{X}_\kappa$  is the solution of the DE

$$\frac{dx'_i}{dt} = u_{i\kappa}(\mathbf{x}'), \quad x'_i(0) = x_i. \quad (27.23)$$

Once the solution is obtained, one can replace  $t$  with  $a_\kappa$  for each  $\kappa$ . In some applications,  $u_{i\kappa}(\mathbf{x})$  will be given implicitly in terms of certain parameters of integration of some DEs [unrelated to (27.23)]. The solution of these DEs are typically generators of coordinate transformations that can be written linearly in terms of the parameters. To be more precise, suppose that after solving some DEs, we obtain

$$X_i = \sum_{\kappa=1}^r c_{i\kappa} f_\kappa^{(i)}(x_1, \dots, x_n), \quad (27.24)$$

where  $\{c_{i\kappa}\}$  are the parameters of integration, and  $X_i$  are components of the vector field that generates the coordinate transformation. This means that for small parameters, one can write

$$x'_i = x_i + \sum_{\kappa=1}^r c_{i\kappa} f_\kappa^{(i)}(x_1, \dots, x_n)$$

and read off  $u_{i\kappa}(\mathbf{x}) = f_\kappa^{(i)}(\mathbf{x})$ . In that case, we have

$$\frac{dx'_i}{dt} = f_\kappa^{(i)}(x'_1, \dots, x'_n), \quad x'_i(0) = x_i. \quad (27.25)$$

We shall have occasion to use this formula later.

### 27.1.4 Integration on Lie Groups

As any other manifold, one can define integration on Lie groups; i.e., one can construct nonvanishing  $n$ -forms and use Equation (26.46) to define integrals on a Lie group  $G$ . Because of the left-invariant property of objects on  $G$ , it would be helpful if the integration process were also left-invariant. For this to happen, the  $n$ -form would have to be left-invariant. It turns out that this can be accomplished more or less uniquely:

**27.1.31. Proposition.** *Let  $G$  be a Lie group of dimension  $n$ . Then there exists a left-invariant nonvanishing  $n$ -form  $\mu$  that is unique up to a nonzero multiplicative constant. If  $G$  is compact, then  $\mu$  is also right-invariant and the multiplicative constant can be chosen to be 1.*

*Proof.* Let  $\mu_e$  be any nonzero  $n$ -form on  $\mathcal{T}_e(G)$ . The desired  $n$ -form is the left translation of this form, i.e.,  $L_g^* \mu_e$ . Indeed, let  $\{X_i\}_{i=1}^n$  be left invariant. Then

$$\begin{aligned} \mu_g(X_1|_g, \dots, X_n|_g) &= L_{g^{-1}}^* \mu_e(X_1|_g, \dots, X_n|_g) \\ &= \mu_e(L_{g^{-1}*} X_1|_g, \dots, L_{g^{-1}*} X_n|_g) \\ &= \mu_e(X_1|_{g^{-1}g}, \dots, X_n|_{g^{-1}g}) = \mu_e(X_1|_e, \dots, X_n|_e). \end{aligned}$$

This shows that  $\mu$  is left-invariant. Now note that any other  $n$ -form  $\mu'_e$  on  $\mathcal{T}_e(G)$  is a constant multiple of  $\mu_e$ . Therefore, the corresponding  $n$ -form  $\mu'_g$  will be a constant multiple of  $\mu_g$ .

Let  $x \in G$  and consider  $\mu' \equiv R_x^* \mu$ . We have

$$L_g^* \mu' = L_g^* \circ R_x^* \mu = R_x^* \circ L_g^* \mu = R_x^* \mu = \mu',$$

where we used the fact that  $L_g$  and  $R_x$  commute and that  $\mu$  is left invariant. The equation above shows that  $\mu'$  is also left-invariant. Therefore,  $\mu' = c\mu$ . If  $G$  is compact, we can integrate both sides and note that  $\int_G \mu = \int_G \mu'$  because  $\mu'$  is related to  $\mu$  by a change of variable. Therefore,  $c = 1$  and  $R_x^* \mu = \mu$ .  $\square$

The left-invariant volume element (nonvanishing  $n$ -form) guaranteed by the proposition above is called **Haar measure**. Since all calculations are done using some coordinate system, we give an explicit expression of the Haar measure in terms of coordinates (parameters) of a general Lie group. Let  $\mathbf{y} = (y^1, \dots, y^r)$  be the coordinates of the translation of  $\mathbf{x} = (x^1, \dots, x^r)$  by  $g \in G$ . Then we can write  $\mathbf{y} = m(\mathbf{g}, \mathbf{x})$ , so that  $dy^j = (\partial y^j / \partial x^i) dx^i = (\partial m^j / \partial x^i) dx^i$ . Therefore,

$$dy^1 \wedge \dots \wedge dy^r = \det \left( \frac{\partial m^j(\mathbf{g}, \mathbf{x})}{\partial x^i} \right) dx^1 \wedge \dots \wedge dx^r.$$

In particular, if  $\mathbf{x} = 0$ , the coordinates of the identity, then  $\mathbf{y}$  will be the coordinates of  $g$ . So, the volume element at  $g$ , denoted by  $d^r y$ , will be given by

$$d^r y = \det \left( \frac{\partial m^j(\mathbf{g}, \mathbf{x})}{\partial x^i} \right) \Big|_{\mathbf{x}=0} d^r x.$$

Note that this is consistent with the geometric definition of the invariant measure given in Proposition 27.1.31 because  $L_{g^{-1}*} = L_{g*}^{-1}$  and the matrix of  $L_g^*$  is the inverse of the matrix of  $L_{g*}$ . The volume element at  $g$ , which is invariant on  $G$ —and therefore has the same value as at the identity—and which we denote by  $d\mu(g)$ , will be given by

$$d\mu(g) = d\mu(e) \equiv d^r x = \det^{-1} \left( \frac{\partial m^j(\mathbf{g}, \mathbf{x})}{\partial x^i} \right) \Big|_{\mathbf{x}=0} d^r g, \tag{27.26}$$

density functions  
associated with Haar  
measure

where we have replaced  $y$  with the more suggestive  $g$ . The volume element  $d^r g$  is the ordinary Euclidean volume element of  $\mathbb{R}^r$  evaluated at the parameters corresponding to  $g$ . The quantity multiplying  $d^r g$  is called the **density function**. Note that since we are interested in the derivatives of  $m^j$  at small values of  $\mathbf{x}$ , we can take the components of  $\mathbf{x}$  to be small, and retain them only up to the first order. This will sometimes simplify the calculation of the invariant Haar measure.

**27.1.32. Example.** From the multiplication rule for the one-dimensional projective group given in Example 27.1.8, we easily find

$$\det \left( \frac{\partial m_i}{\partial b_j} \right) \Big|_{\mathbf{b}=0} = \det \begin{pmatrix} a_1 & 0 & a_2 & 0 \\ 0 & a_1 & 0 & a_2 \\ a_3 & 0 & a_4 & 0 \\ 0 & a_3 & 0 & a_4 \end{pmatrix} = (a_1 a_4 - a_2 a_3)^2.$$

Thus the density function is  $(a_1 a_4 - a_2 a_3)^{-2}$ , and the invariant Haar measure is

$$d\mu(\mathbf{a}) = (a_1 a_4 - a_2 a_3)^{-2} d^4 a. \quad \blacksquare$$

## 27.2 An Outline of Lie Algebra Theory

The notion of a Lie algebra has appeared on a number of occasions both in our study of vector fields on manifolds and, more recently, in the study of Lie groups in the vicinity of their identity elements. Lie algebras play an important role in the representation theory of Lie groups as well. It is therefore worth our effort to spend some time getting acquainted with the formal structure and properties of these algebras. We shall restrict our discussion to finite-dimensional Lie algebras.

Lie algebra defined

**27.2.1. Definition.** A finite-dimensional vector space  $\mathcal{V}$  over  $\mathbb{R}$  (or  $\mathbb{C}$ ) is called a **Lie algebra over  $\mathbb{R}$  (or  $\mathbb{C}$ )** if there is a binary operation, called **Lie multiplication**,  $[\cdot, \cdot] : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$  on  $\mathcal{V}$ , satisfying

1.  $[\mathbf{X}, \mathbf{Y}] = -[\mathbf{Y}, \mathbf{X}]$  for all  $\mathbf{X}, \mathbf{Y} \in \mathcal{V}$  (antisymmetry).
2.  $[\alpha\mathbf{X} + \beta\mathbf{Y}, \mathbf{Z}] = \alpha[\mathbf{X}, \mathbf{Z}] + \beta[\mathbf{Y}, \mathbf{Z}]$  for  $\alpha, \beta \in \mathbb{R}$  (or  $\mathbb{C}$ ) (linearity).
3.  $[\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]] + [\mathbf{Z}, [\mathbf{X}, \mathbf{Y}]] + [\mathbf{Y}, [\mathbf{Z}, \mathbf{X}]] = 0$  (Jacobi identity).

The concepts of a homomorphism, its kernel, its range, etc. are the same as before.

To distinguish Lie algebras from vector spaces, we shall denote the former by lowercase German letters as we have done for the Lie algebras of Lie groups.

**27.2.2. Example.** Recall from Chapter 1 that an algebra is a vector space with a product. If this product is associative, then one can construct a Lie algebra out of the associative algebra by defining  $[a, b] \equiv ab - ba$ . In particular, the matrix algebra under commutation of matrices becomes a Lie algebra, which we denote by  $\mathfrak{gl}(n, \mathbb{R})$  [or  $\mathfrak{gl}(n, \mathbb{C})$ ]. ■

subalgebra, ideal,  
and center of a Lie  
algebra

**27.2.3. Definition.** Let  $\mathfrak{v}$  be a Lie algebra. A subspace  $\mathfrak{u}$  of  $\mathfrak{v}$  is called a *subalgebra* if  $[X, Y] \in \mathfrak{u}$  whenever  $X, Y \in \mathfrak{u}$ . The subspace  $\mathfrak{u}$  is called an *ideal* if  $[X, Y] \in \mathfrak{u}$  whenever either  $X \in \mathfrak{u}$  or  $Y \in \mathfrak{u}$ . The *center*  $\mathfrak{z}$  of  $\mathfrak{v}$  is the collection of all  $X \in \mathfrak{v}$  whose Lie multiplication with all vectors of  $\mathfrak{v}$  vanishes. A Lie algebra is *abelian*, or *commutative*, if  $\mathfrak{z} = \mathfrak{v}$ .

Knowing the  
structure constants,  
one can reconstruct  
the Lie algebra!

If we choose a basis in the Lie algebra  $\mathfrak{v}$ , and express the Lie multiplication of basis vectors as a linear combination of basis vectors, we end up with *basis-dependent* structure constants that satisfy Equation (27.13). The structure constants completely determine the Lie algebra: Given these constants, one can choose a vector space  $\mathcal{V}$  of correct dimension, a basis in that space, and impose the Lie multiplication law among the basis vectors suggested by the structure constants. Once the Lie multiplication law for basis vectors is established, the law for arbitrary vectors follows from linearity of Lie multiplication. This procedure induces a binary operation on  $\mathcal{V}$  and turns it into a Lie algebra  $\mathfrak{v}$ . Any other algebra so constructed will be isomorphic to  $\mathfrak{v}$ .

**27.2.4. Example.** We can classify all two-dimensional Lie algebras by analyzing their structure constants. Let  $X_1$  and  $X_2$  be any two linearly independent vectors of the two-dimensional Lie algebra  $\mathfrak{v}$ . Write the only nonzero Lie bracket as

$$[X_1, X_2] = c_1 X_1 + c_2 X_2.$$

There are two cases to consider: Either  $c_1 = 0 = c_2$  or one of the constants is nonzero. The first case corresponds to a 2-dimensional abelian Lie algebra:

$$[X_i, X_j] = 0 \quad \text{for } i, j = 1, 2.$$

For the second case, suppose  $c_1 \neq 0$  and define the vectors

$$X \equiv c_1 X_1 + c_2 X_2, \quad Y \equiv X_2/c_1.$$

Then the nonzero Lie bracket becomes  $[X, Y] = X$ . ■

The result of Example 27.2.4 is summarized as follows:

**27.2.5. Box.** There are only two 2-dimensional Lie algebras given by either one of the following nonzero Lie bracket relations:

$$[X_1, X_2] = 0 \quad \text{or} \quad [X_1, X_2] = X_1.$$

**27.2.6. Example.** The Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

form a Lie algebra under the commutation relation given by

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l.$$

Thus,  $c_{jk}^l = 2i\epsilon_{jkl}$ . Pauli spin matrices are a basis for  $\mathfrak{su}(2)$ . ■

**27.2.7. Example.** The Lie group  $GL(n, \mathbb{R})$  has  $\mathfrak{gl}(n, \mathbb{R})$ , the set of all real  $n \times n$  matrices, as its Lie algebra. The standard basis of this Lie algebra, also called the **Weyl basis**, consists of matrices  $e_{ij}$  that have zeros everywhere except at the  $ij$ th position. We therefore have

$$(e_{ij})_{kl} = \delta_{ik}\delta_{jl}. \tag{27.27}$$

We can readily find the Lie multiplication (commutation relations) for these matrices. We simply need to look at the elements of the matrix of the commutator:

$$\begin{aligned} (e_{ij}, e_{kl})_{mn} &= (e_{ij}e_{kl})_{mn} - (e_{kl}e_{ij})_{mn} \\ &= (e_{ij})_{mr}(e_{kl})_{rn} - (e_{kl})_{mr}(e_{ij})_{rn} \\ &= \delta_{im}\delta_{jr}\delta_{kr}\delta_{ln} - \delta_{km}\delta_{lr}\delta_{ir}\delta_{jn} = \delta_{im}\delta_{jk}\delta_{ln} - \delta_{km}\delta_{li}\delta_{jn} \\ &= (e_{il})_{mn}\delta_{jk} - (e_{kj})_{mn}\delta_{li}, \end{aligned}$$

or

$$[e_{ij}, e_{kl}] = \delta_{jk}e_{il} - \delta_{li}e_{kj}. \tag{27.28}$$

The structure constants, which are naturally double-indexed, can be read off from Equation (27.28):

$$c_{ij,kl}^{mn} = \delta_{jk}\delta_i^m\delta_l^n - \delta_{li}\delta_k^m\delta_j^n, \tag{27.29}$$

where we have used a superscript for some of the Kronecker deltas to conform to the position of the corresponding index on the LHS. ■

**27.2.8. Example.** An important datum is the dimension of the Lie group (or its associated Lie algebra, since they are the same). This datum is not apparent in most cases of interest in which the group is defined in terms of some geometric property. For example, the symplectic group is defined as all linear transformations  $A$  that leave a certain antisymmetric bilinear form invariant (Example 23.2.2). In terms of matrices, we have

$$x^t J x' = x^t J x \Rightarrow x^t A^t J A x = x^t J x \quad \forall x \in \mathbb{R}^{2n}, \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

It follows that the symplectic group consists of all matrices  $A$  such that

$$A^t J A = J. \tag{27.30}$$

If we write  $A$  in block form,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where  $A_{ij}$  are  $n \times n$  matrices, then, Equation (27.30) becomes

$$\begin{pmatrix} A_{11}^t & A_{21}^t \\ A_{12}^t & A_{22}^t \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

or

$$A_{11}^t A_{21} = A_{21}^t A_{11}, \quad A_{22}^t A_{12} = A_{12}^t A_{22}, \quad A_{11}^t A_{22} - A_{21}^t A_{12} = 1. \quad (27.31)$$

For the symplectic algebra  $\mathfrak{sp}(2n, \mathbb{R})$ , we are interested in the matrix  $A$  when it is close to the identity. This means that

$$A_{11} = 1 + \epsilon X_{11}, \quad A_{22} = 1 + \epsilon X_{22}, \quad A_{12} = \epsilon X_{12}, \quad A_{21} = \epsilon X_{21}.$$

Substituting these in Equation (27.31) and keeping terms linear in  $\epsilon$ , we obtain the following relations among  $X_{ij}$ :

$$X_{22}^t = -X_{11}, \quad X_{12}^t = X_{12}, \quad X_{21}^t = X_{21}. \quad (27.32)$$

It follows that we need  $n^2$  parameters to describe the  $n \times n$  matrices  $X_{11}$  and  $X_{22}$  simultaneously. For the symmetric matrices  $X_{12}$  and  $X_{21}$ , we need  $n(n+1)/2$  independent parameters each. Therefore, the total number of independent parameters needed for (or the dimension of) the symplectic algebra  $\mathfrak{sp}(2n, \mathbb{R})$  is

$$n^2 + 2 \frac{n(n+1)}{2} = n(2n+1). \quad \blacksquare$$

Although our attempt is to give a formal discussion of the Lie algebras and their structure in this section, we shall do this with an eye to the eventual utility of this discussion in a better understanding of the Lie algebras of Lie groups. To make the connection between the present formalism and the Lie algebras arising from Lie groups, we shall make heavy use of matrix groups, i.e.,  $GL(n, \mathbb{R})$  [or  $GL(n, \mathbb{C})$ ] and its subgroups. Equation (27.8) gives a method of finding the matrices of the algebra if those of the group are known:

**27.2.9. Box.** Differentiate the matrix with respect to a parameter at the identity (where all parameters are set equal to zero) to find the matrix "in the direction" of that parameter.

### 27.2.1 The Lie Algebras $\mathfrak{o}(p, n - p)$ and $\mathfrak{p}(p, n - p)$

Many of the Lie groups encountered in physical applications are special cases of the (pseudo) orthogonal group  $O(p, n - p)$  and its associated Poincaré group  $P(p, n - p)$ . It is therefore worthwhile to study their Lie algebras in some detail.

Introduce the diagonal matrix

$$\eta = \text{diag}(\underbrace{-1, -1, \dots, -1}_{p \text{ times}}, \underbrace{1, 1, \dots, 1}_{n - p \text{ times}})$$

and note that the (pseudo) orthogonal group  $O(p, n - p)$  consists of  $n \times n$  matrices that leave the bilinear form  $x \cdot x \equiv x^t \eta x$  invariant for  $x \in \mathbb{R}^n$ . This means that the matrices  $A$  will have to satisfy

$$A^t \eta A = \eta \Rightarrow (\det A)^2 = 1. \tag{27.33}$$

$\eta$ -orthogonal matrices

Such matrices are called  $\eta$ -orthogonal. The fact that  $O(p, n - p)$  is a group and that  $\eta^{-1} = \eta$  can be used to show that

$$A \eta A^t = \eta. \tag{27.34}$$

#### 27.2.10. Example. THE LORENTZ GROUP

The group of the special theory of relativity is the full Lorentz group  $O(3, 1)$ . This is the group of transformations that leave the invariant length<sup>6</sup>

$$\eta^{ij} x_i x_j = -x_1^2 - x_2^2 - x_3^2 + x_4^2 \equiv x_0^2 - x_1^2 - x_2^2 - x_3^2$$

of a 4-vector  $(x_1, x_2, x_3, x_0 = ct)$  invariant. The  $(0, 0)$ -components of Equations (27.33) and (27.34) yield

$$\begin{aligned} a_{00}^2 - a_{10}^2 - a_{20}^2 - a_{30}^2 &= 1, \\ a_{00}^2 - a_{01}^2 - a_{02}^2 - a_{03}^2 &= 1. \end{aligned} \tag{27.35}$$

orthochronous and proper orthochronous Lorentz transformations

Either one of these equations implies that  $a_{00} \geq 1$  or  $a_{00} \leq -1$ . Lorentz transformations for which  $a_{00} \geq 1$  are called **orthochronous**. Since  $\det \mathbf{1} = +1$  and  $1_{00} = +1$ , the identity belongs to the subset consisting of transformations with  $\det A = +1$  and  $a_{00} \geq 1$ . Such transformations form a subgroup of  $O(3, 1)$  called the **proper orthochronous** Lorentz transformations, and have the property that they can be reached continuously from the identity.

Timelike, spacelike, and null vectors;  $\mathbb{R}^4$  as the set of events

Depending on whether  $x \cdot x > 0$ ,  $x \cdot x < 0$ , or  $x \cdot x = 0$ , the vector  $x$  is called **timelike**, **spacelike**, or **null**, respectively. In the special theory of relativity  $\mathbb{R}^4$  becomes the set of **events**. At every event  $x$  the set  $\mathbb{R}^4$  is divided into 5 regions:

1. All events  $y = (y_1, y_2, y_3, y_0)$  to which one can go from  $x$  by material objects, with speed less than  $c$ , lie to the future of  $x$ , i.e.,  $y_0 - x_0 > 0$ , and are timelike:

$$(y_0 - x_0)^2 > (y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2.$$

future light cone

They form a 4-dimensional subset of  $\mathbb{R}^4$  and are said to lie *inside* the **future light cone**.

<sup>6</sup>It is common to label the time coordinate with index 0 rather than 4. We shall use this convention.

2. All events  $y = (y_1, y_2, y_3, y_0)$  to which one can go from  $x$  only by a light signal lie to the future of  $x$ , i.e.,  $y_0 - x_0 > 0$ , and

$$(y_0 - x_0)^2 - (y_1 - x_1)^2 - (y_2 - x_2)^2 - (y_3 - x_3)^2 = 0.$$

They form a 3-dimensional subset of  $\mathbb{R}^4$  and are said to lie *on* the future light cone.

3. All events  $y = (y_1, y_2, y_3, y_0)$  from which one can come to  $x$  by material objects, with speed less than  $c$ , lie in the past of  $x$ , i.e.,  $x_0 - y_0 > 0$ , and are timelike:

$$(x_0 - y_0)^2 > (x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2.$$

past light cone

They form a 4-dimensional subset of  $\mathbb{R}^4$  and are said to lie *inside* the **past light cone**.

4. All events  $y = (y_1, y_2, y_3, y_0)$  from which one can come from  $x$  only by a light signal lie to the past of  $x$ , i.e.,  $x_0 - y_0 > 0$ , and

$$(x_0 - y_0)^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2 = 0.$$

They form a 3-dimensional subset of  $\mathbb{R}^4$  and are said to lie *on* the past light cone.

elsewhere

5. All events in the remaining part of  $\mathbb{R}^4$  form a 4-dimensional subset, are spacelike, and cannot be connected to  $x$  by any means. They are said to belong to **elsewhere**.

From a physical standpoint, future and past are observer-independent. Therefore, if  $y$  lies in or on the future light cone of  $x$  with respect to one observer, it should also do so with respect to all observers. Since observers are connected by Lorentz transformations, we expect the latter to preserve this relation between  $x$  and  $y$ . Not all elements of  $O(3, 1)$  have this property. However, the proper orthochronous transformations do. The details are left as a problem for the reader (see Problem 27.14). ■

As a prototype of  $\eta$ -orthogonal matrices, consider the matrix obtained from the unit matrix by removing the  $ii$ th,  $ij$ th,  $ji$ th, and  $jj$ th elements, and replacing them by an overall  $2 \times 2$  matrix. The result, denoted by  $A^{(ij)}$ , will look like

$$A^{(ij)} = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & a_{ii} & \dots & a_{ij} & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & a_{ji} & \dots & a_{jj} & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}.$$

This matrix will transform  $(x_1, \dots, x_n) \in \mathbb{R}^n$  according to

$$x'_i = a_{ii}x_i + a_{ij}x_j, \quad (\text{no summation!})$$

$$x'_j = a_{ji}x_i + a_{jj}x_j,$$

$$x'_k = x_k \quad \text{for } k \neq i, j.$$



In order for  $A^{(ij)}$  to leave the bilinear form  $x^t \eta x$  invariant, the  $2 \times 2$  submatrix  $\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix}$  must be either a rotation (corresponding to the case where  $i, j \leq p$  or  $i, j > p$ ), or a Lorentz boost<sup>7</sup> (corresponding to the case where  $i \leq p$  and  $j > p$ ). In the first case, we have

$$\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

and in the second case

$$\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix} = \begin{pmatrix} \cosh \xi & -\sinh \xi \\ -\sinh \xi & \cosh \xi \end{pmatrix},$$

where  $\xi \equiv \tanh^{-1}(v/c)$  is the “rapidity.”

The matrices of the algebra are obtained by differentiation at  $\theta = 0$  (or  $\xi = 0$ ). Denoting these matrices by  $M_{ij}$ , we readily find that for the case of rotations,  $M_{ij}$  has  $-1$  at the  $ij$ th position,  $+1$  at the  $ji$ th position, and  $0$  everywhere else. For the case of boosts,  $M_{ij}$  has  $-1$  at the  $ij$ th and the  $ji$ th position, and  $0$  everywhere else. Both cases can be described by the single relation

$$(M_{ij})^m_l = \eta_{il} \delta_j^m - \eta_{jl} \delta_i^m, \quad M_{ij} = -M_{ji}.$$

It is convenient to have all indices in the lower position. So, we multiply both sides by  $\eta_{mk}$  to obtain

$$(M_{ij})_{kl} = \eta_{il} \eta_{jk} - \eta_{jl} \eta_{ik}, \quad M_{ij} = -M_{ji}. \tag{27.36}$$

We can use Equation (27.36) to find the Lie multiplication (in this case, matrix commutation relations) for the algebra  $\mathfrak{o}(p, n - p)$ :

$$[M_{ij}, M_{kl}] = \eta_{ik} M_{jl} - \eta_{il} M_{jk} + \eta_{jl} M_{ik} - \eta_{jk} M_{il}. \tag{27.37}$$

The Lie group  $O(p, n - p)$  includes rotations and Lorentz transformation. Another group with considerable significance in physics is the **Poincaré group**  $P(p, n - p)$ , which includes translations<sup>8</sup> in space and time as well. An element of  $P(p, n - p)$  transforms  $x \in \mathbb{R}^n$  to  $x' = Ax + u$ , where  $u$  is a column vector representing the translation part of the group. It is convenient to introduce matrices to represent these group operations. This is possible if we represent an element of  $\mathbb{R}^n$  as an  $(n + 1)$ -column whose last element is an insignificant  $1$ . then, the reader may check that a Poincaré transformation can be written as

$$\begin{pmatrix} x' \\ 1 \end{pmatrix} = \begin{pmatrix} A & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}, \tag{27.38}$$

Lie brackets for the algebra  $\mathfrak{o}(p, n - p)$

Poincaré group as  $(n + 1) \times (n + 1)$  matrices

<sup>7</sup>The elementary Lorentz transformations involving only one space dimension.  
<sup>8</sup>One can think of the Poincaré group as a subgroup of the group of affine motions in which the matrices belong to  $O(p, n - p)$  rather than  $GL(n, \mathbb{R})$ .

where  $A$  is the  $n \times n$  matrix of  $O(p, n - p)$ , and  $u$  is an  $n$ -dimensional column vector.

The Lie algebra of the Poincaré group is obtained by differentiating the  $(n + 1) \times (n + 1)$  matrix of Equation (27.38). The differentiation of the matrix  $A$  will give  $o(p, n - p)$  of Equation (27.37). The translation part will lead to matrices  $P_i$  with matrix elements given by

$$(P_i)^k_l = \delta_i^k \delta_l^{n+1} \Rightarrow (P_i)_{kl} = \eta_{ik} \delta_l^{n+1}. \tag{27.39}$$

These matrices satisfy the following Lie multiplication rules:

$$[P_i, P_j] = 0, \quad [M_{ij}, P_k] = \eta_{ik} P_j - \eta_{jk} P_i.$$

It then follows that the full Poincaré algebra  $\mathfrak{p}(p, n - p)$  is described by the following Lie brackets:

Lie brackets for the Poincaré algebra  $\mathfrak{p}(p, n - p)$

$$\begin{aligned} [M_{ij}, M_{kl}] &= \eta_{ik} M_{jl} - \eta_{il} M_{jk} + \eta_{jl} M_{ik} - \eta_{jk} M_{il}, \\ [M_{ij}, P_k] &= \eta_{ik} P_j - \eta_{jk} P_i, \\ [P_i, P_j] &= 0. \end{aligned} \tag{27.40}$$

### 27.2.2 Operations on Lie Algebras

**27.2.11. Definition.** Let  $\mathfrak{v}$  be a Lie algebra. A linear operator  $D : \mathfrak{v} \rightarrow \mathfrak{v}$  satisfying

$$D[X, Y] = [DX, Y] + [X, DY]$$

is called a *derivation* of  $\mathfrak{v}$ .

derivation algebra of a Lie algebra

Although the product of two derivations is not a derivation, their commutator is. Therefore, the set of derivations of a Lie algebra  $\mathfrak{v}$  themselves form a Lie algebra  $\mathfrak{D}_{\mathfrak{v}}$  under commutations, which is called the **derivation algebra**.

Recall that the infinitesimal generators of the adjoint action of a Lie group on its algebra were given by  $\text{ad}_{\xi}$  [Equation (27.14)]. We can apply this to a general Lie algebra  $\mathfrak{v}$  by fixing a vector  $X \in \mathfrak{v}$  and defining the map  $\text{ad}_X : \mathfrak{v} \rightarrow \mathfrak{v}$  given by  $\text{ad}_X(Y) = [X, Y]$ . The reader may verify that  $\text{ad}_X$  is a derivation of  $\mathfrak{v}$  and that  $\text{ad}_{[X, Y]} = [\text{ad}_X, \text{ad}_Y]$ . Therefore, the set  $\text{ad}_{\mathfrak{v}} \equiv \{\text{ad}_X \mid X \in \mathfrak{v}\}$  is a Lie algebra, a subalgebra of the derivation algebra  $\mathfrak{D}_{\mathfrak{v}}$  of  $\mathfrak{v}$ , and is called the **adjoint algebra** of  $\mathfrak{v}$ . There is a natural homomorphism  $\psi : \mathfrak{v} \rightarrow \text{ad}_{\mathfrak{v}}$  given by  $\psi(X) = \text{ad}_X$  whose kernel is the center of  $\mathfrak{v}$ . Furthermore,  $\text{ad}_{\mathfrak{v}}$  is an *ideal* of  $\mathfrak{D}_{\mathfrak{v}}$ .

adjoint algebra of a Lie algebra

**27.2.12. Example.** We construct the matrix representation of the operators in the adjoint algebra of  $\mathfrak{su}(2)$  with Pauli spin matrices as a basis. From

$$\text{ad}_{\sigma_1}(\sigma_1) = [\sigma_1, \sigma_1] = 0$$

we conclude that the first column of the matrix of  $\text{ad}_{\sigma_1}$  is zero. From

$$\text{ad}_{\sigma_1}(\sigma_2) = [\sigma_1, \sigma_2] = 2i\sigma_3$$

Illustration of homomorphism of  $\mathfrak{su}(2)$  and its adjoint using Pauli spin matrices

we conclude that the second column of the matrix of  $\text{ad}_{\sigma_1}$  has zeros for the first two elements and  $2i$  for the last. Similarly, from

$$\text{ad}_{\sigma_1}(\sigma_3) = [\sigma_1, \sigma_3] = -2i\sigma_2$$

we conclude that the third column of the matrix of  $\text{ad}_{\sigma_1}$  has zeros for the first and third elements and  $-2i$  for the second. Thus, the matrix representation of  $\text{ad}_{\sigma_1}$  is

$$\text{ad}_{\sigma_1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2i \\ 0 & 2i & 0 \end{pmatrix}.$$

Likewise, we can obtain the other two matrix representations; they are

$$\text{ad}_{\sigma_2} = \begin{pmatrix} 0 & 0 & 2i \\ 0 & 0 & 0 \\ -2i & 0 & 0 \end{pmatrix}, \quad \text{ad}_{\sigma_3} = \begin{pmatrix} 0 & -2i & 0 \\ 2i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The reader may readily verify that  $[\text{ad}_{\sigma_j}, \text{ad}_{\sigma_k}] = 2i\epsilon_{jkl}\text{ad}_{\sigma_l}$ . ■

If  $\psi$  is an automorphism of  $\mathfrak{v}$ , i.e., an isomorphism of  $\mathfrak{v}$  onto itself, then

$$\text{ad}_{\psi(X)} = \psi \circ \text{ad}_X \circ \psi^{-1} \quad \forall X \in \mathfrak{v}. \tag{27.41}$$

Since  $\text{ad}_X$  is an operator on the vector space  $\mathfrak{v}$ , one can define the trace of  $\text{ad}_X$ . However, the notion of trace attains a far greater significance when it is combined with the notion of composition of operators. For  $X, Y \in \mathfrak{v}$ , define

$$(X | Y) \equiv \text{tr}(\text{ad}_X \circ \text{ad}_Y). \tag{27.42}$$

Then one can show that  $(\cdot | \cdot)$  is bilinear and symmetric. It becomes an inner product if the “vectors” of the Lie algebra are hermitian operators on some vector space  $\mathcal{V}$ , or if the underlying vector space is over  $\mathbb{R}$  (see Proposition 3.6.6 in Chapter 3). Furthermore,  $(\cdot | \cdot)$  satisfies

$$([X, Y] | Z) + ([X, Z] | Y) = 0. \tag{27.43}$$

Killing form of a Lie algebra

**27.2.13. Definition.** *The symmetric bilinear form  $(\cdot | \cdot) : \mathfrak{v} \times \mathfrak{v} \rightarrow \mathbb{R}$  defined by (27.42) is called the Killing form of  $\mathfrak{v}$ .*

It is an immediate consequence of this definition and Equation (27.41) that the Killing form of  $\mathfrak{v}$  is invariant under all automorphisms of  $\mathfrak{v}$ .

Most mathematicians seem to have little or no interest in history, so that often the name attached to a key result is that of the follow-up person who exploited an idea or theorem rather than its originator (the Jordan form is due to Weierstrass, Wedderburn theory to Cartan and Molien). No one has suffered from this ahistoricism more than Killing. For example, the so-called “Cartan sub-algebra” and “Cartan matrix” were defined and exploited by Killing.

He exhibited the characteristic equation of an arbitrary element of the Weyl group when Weyl was 3 years old and listed the orders of the Coxeter transformation 19 years before Coxeter was born!

**Wilhelm Karl Joseph Killing** (1847–1923) began university study in Münster in 1865 but quickly moved to Berlin and came under the influence of Kummer and Weierstrass. From 1868 to 1882 much of Killing's energy was devoted to teaching at the gymnasium level in Berlin and Brilon (south of Münster). At one stage, when Weierstrass was urging him to write up his research on space structures, he was spending as much as 36 hours per week in the classroom or tutoring. (Now many mathematicians consider 6 hours a week an intolerable burden!) On the recommendation of Weierstrass, Killing was appointed professor of mathematics at the Lyzeum Hosianum in Braunsberg, in East Prussia (now Braniewo in the region of Olsztyn, in Poland). This was a college founded in 1565 by Bishop Stanislaus Hosius, whose treatise on the Christian faith ran to 39 editions! The main object of the college was the training of Roman Catholic clergy, so Killing had to teach a wide range of topics, including the reconciliation of faith and science. Although he was isolated mathematically during his ten years in Braunsberg, this was the most creative period in his mathematical life. Killing produced his brilliant work despite worries about the health of his wife and seven children, demanding administrative duties as rector of the college and as a member and chairman of the City Council, and his active role in the Church of St. Catherine.

What we now call Lie algebras were invented by the Norwegian mathematician Sophus Lie about 1870 and independently by Killing about 1880. Lie was seeking to develop an approach to the solution of differential equations analogous to the Galois theory of algebraic equations. Killing's consuming passion was non-Euclidean geometries and their generalizations, so he was led to the problem of classifying infinitesimal motions of a rigid body in any type of space (or *Raumformen*, as he called them).

In 1892 he was called back to his native Westphalia as professor of mathematics at the University of Münster, where he was quickly submerged in teaching, administration, and charitable activities. He was Rector Magnificus for some period and president of the St. Vincent de Paul charitable society for ten years. Killing's work was neglected partly because he was a modest man with high standards; he vastly underrated his own achievement. His interest was geometry, and for this he needed all real Lie algebras. To obtain merely the simple Lie algebras over the complex numbers did not appear to him to be very significant. Another reason was due to Lie, who was quite negative about Killing's work. At the top of page 770 of a three-volume joint work of Lie and Engel we find the following less than generous comment about Killing: "With the exception of the preceding unproved theorem . . . all the theorems that are correct are due to Lie and all the false ones are due to Killing!"



Killing was conservative in his political views and vigorously opposed the attempt to reform the examination requirements for graduate students at the University of Münster by deleting the compulsory study of philosophy. Engel comments "Killing could not see that for most candidates the test in philosophy was completely worthless." He had a profound patriotic love of his country, so that his last years (1918-1923) were deeply pained by the collapse of social cohesion in Germany after the War of 1914-18.

(Taken from A. J. Coleman, "The Greatest Mathematical Paper of All Times," *Mathematical Intelligencer* 11 (3) (1989) 29–38.)

As noted above, the Killing form is an *inner product* if the Lie algebra consists of hermitian operators. This will certainly happen if the Lie algebra is that of a group whose elements are unitary operators on some vector space  $\mathcal{V}$ . We shall see shortly that such unitary operators are not only possible, but have extremely useful properties in the representation of compact Lie groups. A unitary representation of a Lie group induces a representation of its Lie algebra whose "vectors" are hermitian operators. Then the Killing form becomes an inner product. The natural existence of such Killing forms for the representation of a *compact* Lie group motivates the following:

compact Lie algebra **27.2.14. Definition.** A Lie algebra  $\mathfrak{v}$  is *compact* if it has an inner product  $(\cdot|\cdot)$  satisfying

$$([\mathbf{X}, \mathbf{Y}] | \mathbf{Z}) + ([\mathbf{X}, \mathbf{Z}] | \mathbf{Y}) = 0.$$

Choose a basis  $\{\mathbf{X}_i\}$  for the Lie algebra  $\mathfrak{v}$  and note that  $(\text{ad}_{\mathbf{X}_i})_j^k = c_{ij}^k$ . Therefore,

$$(\mathbf{X}_i | \mathbf{X}_j) \equiv \text{tr}(\text{ad}_{\mathbf{X}_i} \circ \text{ad}_{\mathbf{X}_j}) = (\text{ad}_{\mathbf{X}_i})_l^k (\text{ad}_{\mathbf{X}_j})_k^l = c_{il}^k c_{jk}^l \equiv g_{ij}, \quad (27.44)$$

Cartan metric tensor of a Lie algebra where  $g_{ij}$  are components of the so-called **Cartan metric tensor** in the basis  $\{\mathbf{X}_i\}$ . If  $\mathbf{A}, \mathbf{B} \in \mathfrak{v}$  have components  $\{a^i\}$  and  $\{b^i\}$  in the basis  $\{\mathbf{X}_i\}$ , then it follows from Equation (27.44) that

$$(\mathbf{A} | \mathbf{B}) = a^i b^j g_{ij}, \quad (27.45)$$

as expected of a symmetric bilinear form. We can use the Cartan metric to lower the upper index of the structure constants:  $c_{ijk} \equiv c_{ij}^l g_{lk}$ . By virtue of Equation (27.44), the new constants may be written in the form

$$\begin{aligned} c_{ijk} &= c_{ij}^l c_{ls}^r c_{kr}^s = (-c_{js}^l c_{li}^r - c_{si}^l c_{lj}^r) c_{kr}^s \quad \text{by (27.13)} \\ &= c_{js}^l c_{il}^r c_{kr}^s + c_{si}^l c_{lj}^r c_{rk}^s. \end{aligned}$$

The reader may now verify that the RHS is completely antisymmetric in  $i, j$ , and  $k$ . If the Lie algebra is compact, then one can choose an orthonormal basis in which  $g_{lk} = \delta_{lk}$  (because the inner product is, by definition, positive definite) and obtain  $c_{ij}^k = c_{ijk}$ . We therefore have the following result.

**27.2.15. Proposition.** Let  $\mathfrak{v}$  be a compact Lie algebra. Then there exists a basis of  $\mathfrak{v}$  in which the structure constants are represented by a third-order completely antisymmetric covariant tensor.

**27.2.16. Example.** We can calculate explicitly the Killing form of the Lie algebras  $\mathfrak{gl}(n, \mathbb{R})$  and  $\mathfrak{sl}(n, \mathbb{R})$ . Choose the Weyl basis introduced in Example 27.2.7 and expand  $A, B \in \mathfrak{gl}(n, \mathbb{R})$  in terms of the Weyl basis vectors:  $A = a^{ij}e_{ij}$ ,  $B = b^{ij}e_{ij}$ . The Cartan metric tensor becomes

$$g_{ij,kl} = c_{ij,mn}^{rs} c_{kl,rs}^{mn} = (\delta_{jm} \delta_i^r \delta_n^s - \delta_{in} \delta_m^r \delta_j^s) (\delta_{lr} \delta_k^m \delta_s^n - \delta_{ks} \delta_r^m \delta_l^n),$$

where we have used Equation (27.29). It follows from these relations, Equation (27.45), and a simple index manipulation that

$$(A | B) \equiv a^{ij} b^{kl} g_{ij,kl} = 2n \operatorname{tr}(AB) - 2 \operatorname{tr} A \operatorname{tr} B \tag{27.46}$$

for  $A, B \in \mathfrak{gl}(n, \mathbb{R})$ , and

$$(A | B) = 2n \operatorname{tr}(AB) \tag{27.47}$$

for  $A, B \in \mathfrak{sl}(n, \mathbb{R})$ , because all matrices in  $\mathfrak{sl}(n, \mathbb{R})$  are traceless. ■

A Lie algebra  $\mathfrak{v}$ , as a vector space, may be written as a direct sum of its subspaces. We express this as

$$\mathfrak{v} = u_1 \dot{+} u_2 \dot{+} \cdots \dot{+} u_r = \sum_{k=1}^r \dot{+} u_k.$$

If in addition  $\{u_k\}$  are Lie subalgebras every one of which commutes with the rest, we write

$$\mathfrak{v} = u_1 \oplus u_2 \oplus \cdots \oplus u_r = \sum_{k=1}^r \oplus u_k \tag{27.48}$$

and say that  $\mathfrak{v}$  has been **decomposed** into a direct sum of Lie algebras. In this case, each  $u_k$  is not only a subalgebra, but also an ideal of  $\mathfrak{v}$  as the reader may verify.

The study of the structure of Lie algebras boils down to the study of the “simplest” kind of Lie algebras in terms of which other Lie algebras can be decomposed. Intuitively, one would want to call a Lie algebra “simple” if it has no proper subalgebras. However, in terms of decomposition, such subalgebras are required to be ideals. So the natural definition of a simple Lie algebra would be the following:

semisimple Lie algebras

**27.2.17. Definition.** A Lie algebra that has no proper ideal is called a **simple Lie algebra**. A Lie algebra is **semisimple** if it has no (nonzero) commutative ideal.

For example, the pseudo-orthogonal algebra  $\mathfrak{o}(p, n - p)$  is semisimple, but the Poincaré algebra  $\mathfrak{p}(p, n - p)$  is not because the translation generators  $P_j$  form a commutative ideal.

A useful criterion for semisimplicity is given by the following theorem due to Cartan, which we state without proof (for a proof, see [Baru 86, pp. 15–16]):

**27.2.18. Theorem.** (Cartan) A Lie algebra  $\mathfrak{v}$  is semisimple if and only if  $\det(g_{ij}) \neq 0$ .

The importance of semisimple Lie algebras is embodied [Baru 86, pp. 19–20].

**27.2.19. Theorem.** (Cartan) *A semisimple complex or real Lie algebra can be decomposed into a direct sum of pairwise orthogonal simple subalgebras. This decomposition is unique up to ordering.*

The orthogonality is with respect to the Killing form. Theorem 27.2.19 reduces the study of semisimple Lie algebras to that of simple Lie algebras. What about a general Lie algebra  $\mathfrak{v}$ ? If  $\mathfrak{v}$  is compact, then it turns out that it can be written as  $\mathfrak{v} = \mathfrak{z} \oplus \mathfrak{s}$  where  $\mathfrak{z}$  is the center of  $\mathfrak{v}$  and  $\mathfrak{s}$  is semisimple. If  $\mathfrak{v}$  is not compact, then the decomposition will not be in terms of a direct sum, but in terms of what is called a *semidirect* sum one of whose factors is semisimple. For details, the reader is referred to the fairly accessible treatment of Barut and Raczka, Chapter 1. From now on we shall restrict our discussion to semisimple Lie algebras. These algebras are completely known, because simple algebras have been completely classified. We shall not pursue the classification of Lie algebras. However, we simply state a definition that is used in such a classification, because we shall have an occasion to use it in the representation theory of Lie algebras.

Cartan subalgebra  
and the rank of a Lie  
algebra

**27.2.20. Definition.** *Let  $\mathfrak{v}$  be a Lie algebra. A subalgebra  $\mathfrak{h}$  of  $\mathfrak{v}$  is called a **Cartan subalgebra** if  $\mathfrak{h}$  is the largest commutative subalgebra of  $\mathfrak{v}$ , and for all  $X \in \mathfrak{h}$ , if  $\text{ad}_X$  leaves a subspace of  $\mathfrak{v}$  invariant, then it leaves the complement of  $\mathfrak{v}$  invariant as well. The dimension of  $\mathfrak{h}$  is called the **rank** of  $\mathfrak{v}$ .*

## 27.3 Representation of Compact Lie Groups

Representation of general Lie groups is closely related to representation of their Lie algebras, and we shall discuss them in the next two sections. In this section, however, we shall consider the representation of compact Lie groups, because for such groups, many of the ideas developed for finite groups hold. Before discussing compact groups, let us state a definition and a proposition that hold for all Lie groups.

**27.3.1. Definition.** *A **representation** of a Lie group  $G$  on a Hilbert space  $\mathcal{H}$  is a Lie group homomorphism  $T : G \rightarrow GL(\mathcal{H})$ . Similarly, a representation of the Lie algebra  $\mathfrak{g}$  is a Lie algebra homomorphism  $\mathcal{T} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{H})$ .*

The proposition we have in mind is the important **Schur's lemma** which we state without proof (for a proof see [Baru 86, pp. 143–144]).

Schur's lemma

**27.3.2. Proposition.** (Schur's lemma) *A unitary representation  $T : G \rightarrow GL(\mathcal{H})$  of a Lie group  $G$  is irreducible if and only if the only operators commuting with all the  $T_g$  are scalar multiples of the unit operator.*

**27.3.3. Example.** COMPACTNESS OF  $U(n)$ ,  $O(n)$ ,  $SU(n)$ , AND  $SO(n)$ 

Identify  $GL(n, \mathbb{C})$  with  $\mathbb{R}^{2n^2}$  via components. The map

$$f : GL(n, \mathbb{C}) \rightarrow GL(n, \mathbb{C}) \quad \text{given by} \quad f(A) = AA^\dagger$$

is continuous because it is simply the products of elements of matrices. It follows that  $f^{-1}(1)$  is closed, because the matrix 1 is a single point in  $\mathbb{R}^{2n^2}$ , which is therefore closed.  $f^{-1}(1)$  is also bounded, because

$$AA^\dagger = 1 \Rightarrow \sum_{j=1}^n a_{ij}a_{kj}^* = \delta_{ik} \Rightarrow \sum_{i,j=1}^n |a_{ij}|^2 = n.$$

Thus,  $f^{-1}(1)$  is a  $(2n^2 - 1)$ -dimensional sphere of radius  $\sqrt{n}$  in  $\mathbb{R}^{2n^2}$ , which is clearly bounded. The BWHB theorem (of Chapter 16) now implies that  $f^{-1}(1)$  is compact. Now note that  $f^{-1}(1)$  consists of all matrices that have their hermitian adjoints for an inverse; but these are precisely the set  $U(n)$  of unitary matrices.

Now consider the map  $\det : U(n) \rightarrow \mathbb{C}$ . This map is also continuous, implying that  $\det^{-1}(1)$  is a closed subset of  $U(n)$ . The boundedness of  $U(n)$  implies that  $\det^{-1}(1)$  is also bounded. Invoking the BWHB theorem again, we conclude that  $\det^{-1}(1) = SU(n)$ , being closed and bounded, is compact.

If instead of complex numbers, we restrict ourselves to the reals,  $O(n)$  and  $SO(n)$  will replace  $U(n)$  and  $SU(n)$ , respectively. ■

The result of the example above can be summarized:

**27.3.4. Box.** *The unitary  $U(n)$ , orthogonal  $O(n)$ , special unitary  $SU(n)$ , and special orthogonal  $SO(n)$  groups are all compact.*

We now start our study of the representations of compact Lie groups. We first show that we can always assume that the representation is unitary.

All representations of compact groups can be made unitary.

**27.3.5. Theorem.** *Let  $T : G \rightarrow GL(\mathcal{H})$  be any representation of the compact group  $G$ . There exists a new inner product in  $\mathcal{H}$  relative to which  $T$  is unitary.*

*Proof.* Let  $(| \rangle)$  be the initial inner product. Define a new inner product  $(| \rangle)$  by

$$(u|v) \equiv \int_G \langle T_g u | T_g v \rangle d\mu_g$$



where  $d\mu_g$  is the Haar measure, which is both left- and right-invariant. The reader may check that this is indeed an inner product. For every  $h \in G$ , we have

$$\begin{aligned} (\mathbf{T}_h u | \mathbf{T}_h v) &= \int_G \langle \mathbf{T}_g \mathbf{T}_h u | \mathbf{T}_g \mathbf{T}_h v \rangle d\mu_g \\ &= \int_G \langle \mathbf{T}_{gh} u | \mathbf{T}_{gh} v \rangle d\mu_g \quad (\text{because } T \text{ is a representation}) \\ &= \int_G \langle \mathbf{T}_{gh} u | \mathbf{T}_{gh} v \rangle d\mu_{gh} \quad (\text{because } \mu_g \text{ is right invariant}) \\ &= (u | v). \end{aligned}$$

This shows that  $\mathbf{T}_h$  is unitary for all  $h \in G$ . □

From now on, we shall restrict our discussion to unitary representations of compact groups.

The study of representations of compact groups is facilitated by the following construction:

**27.3.6. Definition.** Let  $T : G \rightarrow GL(\mathcal{H})$  be a unitary representation of the compact group  $G$  and  $|u\rangle \in \mathcal{H}$  a fixed vector. The **Weyl operator**  $\mathbf{K}_u$  associated with  $|u\rangle$  is defined as

Weyl operator for a compact Lie group

$$\mathbf{K}_u = \int_G |\mathbf{T}_g u\rangle \langle \mathbf{T}_g u| d\mu_g. \quad (27.49)$$

The essential properties of the Weyl operator are summarized in the following:

**27.3.7. Proposition.** Let  $T : G \rightarrow GL(\mathcal{H})$  be a unitary representation of the compact group  $G$ . Then the Weyl operator has the following properties

1.  $\mathbf{K}_u$  is hermitian.
2.  $\mathbf{K}_u \mathbf{T}_g = \mathbf{T}_g \mathbf{K}_u$  for all  $g \in G$ . Therefore, any eigenspace of  $\mathbf{K}_u$  is an invariant subspace of all  $\mathbf{T}_g$ 's.
3.  $\mathbf{K}_u$  is a Hilbert-Schmidt operator.

*Proof.* Statement (1), in the form  $\langle w | \mathbf{K}_u | v \rangle^* = \langle v | \mathbf{K}_u | w \rangle$ , follows directly from the definition.

(2) From  $\mathbf{T}_g \int_G |\mathbf{T}_x u\rangle \langle \mathbf{T}_x u| d\mu_x = \int_G |\mathbf{T}_g \mathbf{T}_x u\rangle \langle \mathbf{T}_x u| d\mu_x$ , the fact that  $T$  is a representation (therefore,  $\mathbf{T}_g \mathbf{T}_x = \mathbf{T}_{gx}$ ), and redefining the integration variable to  $y = gx$ , we get

$$\mathbf{T}_g \mathbf{K}_u = \int_G |\mathbf{T}_y u\rangle \langle \mathbf{T}_{g^{-1}y} u| \underbrace{d\mu_{g^{-1}y}}_{=d\mu_y} = \int_G |\mathbf{T}_y u\rangle \langle \mathbf{T}_{g^{-1}y} u| d\mu_y,$$

where we used the left invariance of  $\mu$  and the fact that  $T$  is a representation. Unitarity of  $T$  now gives

$$T_g K_u = \int_G |T_y u\rangle \langle T_g^\dagger T_y u| d\mu_y = \int_G |T_y u\rangle \langle T_y u| T_g d\mu_y = K_u T_g.$$

(3) Recall that an operator  $\mathbf{A} \in \mathcal{L}(\mathcal{H})$  is Hilbert–Schmidt if  $\sum_{i=1}^\infty \|\mathbf{A}|e_i\rangle\|^2$  is finite for any orthonormal basis  $\{|e_i\rangle\}$  of  $\mathcal{H}$ . In the present case, we have

$$K_u |e_i\rangle = \int_G |T_x u\rangle \langle T_x u| e_i\rangle d\mu_x.$$

Therefore,

$$\begin{aligned} \sum_{i=1}^\infty \|K_u |e_i\rangle\|^2 &= \sum_{i=1}^\infty \left( \int_G \langle e_i| T_y u\rangle \langle T_y u| d\mu_y \right) \left( \int_G |T_x u\rangle \langle T_x u| e_i\rangle d\mu_x \right) \\ &= \sum_{i=1}^\infty \int_G \int_G \langle e_i| T_y u\rangle \langle T_y u| T_x u\rangle \langle T_x u| e_i\rangle d\mu_x d\mu_y. \end{aligned}$$

If we switch the order of summation and integration and use

$$\sum_{i=1}^\infty \langle T_x u| e_i\rangle \langle e_i| T_y u\rangle = \langle T_x u| T_y u\rangle,$$

we obtain

$$\sum_{i=1}^\infty \|K_u |e_i\rangle\|^2 = \int_G \int_G |\langle T_y u| T_x u\rangle|^2 d\mu_x d\mu_y,$$

and using the Schwarz inequality in the integral yields

$$\begin{aligned} \sum_{i=1}^\infty \|K_u |e_i\rangle\|^2 &\leq \int_G \int_G \langle T_x u| T_x u\rangle \langle T_y u| T_y u\rangle d\mu_x d\mu_y \\ &= \int_G \int_G \langle u| u\rangle \langle u| u\rangle d\mu_x d\mu_y \quad (\text{because rep. is unitary}) \\ &= \|u\|^4 \int_G d\mu_x \int_G d\mu_y = \|u\|^4 V_G^2 < \infty, \end{aligned}$$

where  $V_G$  is the *finite* volume of  $G$ . □

**Hermann Klaus Hugo Weyl** (1885–1955) attended the gymnasium at Altona and, on the recommendation of the headmaster of his gymnasium, who was a cousin of Hilbert, decided at the age of eighteen to enter the University of Göttingen. Except for one year at Munich he

remained at Göttingen, as a student and later as Privatdozent, until 1913, when he became professor at the University of Zurich. After Klein's retirement in 1913, Weyl declined an offer to be his successor at Göttingen but accepted a second offer in 1930, after Hilbert had retired. In 1933 he decided he could no longer remain in Nazi Germany and accepted a position at the Institute for Advanced Study at Princeton, where he worked until his retirement in 1951. In the last years of his life he divided his time between Zurich and Princeton.

Weyl undoubtedly was the most gifted of Hilbert's students. Hilbert's thought dominated the first part of his mathematical career; and although later he sharply diverged from his master, particularly on questions related to foundations of mathematics, Weyl always shared his convictions that the value of abstract theories lies in their success in solving classical problems and that the proper way to approach a question is through a deep analysis of the concepts it involves rather than by blind computations.

Weyl arrived at Göttingen during the period when Hilbert was creating the spectral theory of self-adjoint operators, and spectral theory and harmonic analysis were central in Weyl's mathematical research throughout his life. Very soon, however, he considerably broadened the range of his interests, including areas of mathematics into which Hilbert had never penetrated, such as the theory of Lie groups and the analytic theory of numbers, thereby becoming one of the most universal mathematicians of his generation. He also had an important role in the development of mathematical physics, the field to which his most famous books, *Raum, Zeit und Materie* (1918), on the theory of relativity, and *Gruppentheorie und Quantenmechanik* (1928), are devoted.



Weyl's versatility is illustrated in a particularly striking way by the fact that immediately after some original advances in number theory (which he obtained in 1914), he spent more than ten years as a geometer—a geometer in the most modern sense of the word, uniting in his methods topology, algebra, analysis, and geometry in a display of dazzling virtuosity and uncommon depth reminiscent of Riemann. Drawn by war mobilization into the German army, Weyl did not resume his interrupted work when he was allowed to return to civilian life in 1916. At Zurich he had worked with Einstein for one year, and he became keenly interested in the general theory of relativity, which had just been published; with his characteristic enthusiasm he devoted most of the next five years to exploring the mathematical framework of the theory. In these investigations Weyl introduced the concept of what is now called a *linear connection*, linked not to the Lorentz group of orthogonal transformations, but to the enlarged group of conformal transformations; he even thought for a time that this would give him a unified theory of gravitation and electromagnetism, the forerunner of what is now called *gauge theories*.

Weyl's use of tensor calculus in his work on relativity led him to reexamine the basic methods of that calculus and, more generally, of classical invariant theory that had been its forerunner but had fallen into near oblivion after Hilbert's work of 1890. On the other hand, his semiphilosophical, semimathematical ideas on the general concept of "space" in connection with Einstein's theory had directed his investigations to generalizations of Helmholtz's problem of characterizing Euclidean geometry by properties of "free mobility." From these two directions Weyl was brought into contact with the theory of linear representations of

Lie groups; his papers on the subject (1925–1927) certainly represent his masterpiece and must be counted among the most influential works in twentieth-century mathematics.

Based on the early 1900s works of Frobenius, I. Schur, and A. Young, Weyl inaugurated a new approach for the representation of continuous groups by focusing his attention on Lie groups, rather than Lie algebras.

Very few of Weyl's 150 published books and papers—even those chiefly of an expository character—lack an original idea or a fresh viewpoint. The influence of his works and of his teaching was considerable: He proved by his example that an “abstract” approach to mathematics is perfectly compatible with “hard” analysis and, in fact, can be one of the most powerful tools when properly applied.

Weyl was one of that rare breed of modern mathematician whose contribution to physics was also substantial. In an interview with a reporter in 1929, Dirac is asked the following question: “. . . I want to ask you something more: They tell me that you and Einstein are the only two real sure-enough high-brows and the only ones who can really understand each other. I won't ask you if this is straight stuff, for I know you are too modest to admit it. But I want to know this—Do you ever run across a fellow that even you can't understand?” To this Dirac replies one word: “Weyl.”

Weyl had a lifelong interest in philosophy and metaphysics, and his mathematical activity was seldom free from philosophical undertones or afterthoughts. At the height of the controversy over the foundations of mathematics, between the formalist school of Hilbert and the intuitionist school of Brouwer, he actively fought on Brouwer's side. His own comment, stated somewhat jokingly, sums up his personality: “My work always tried to unite the truth with the beautiful, but when I had to choose one or the other, I usually chose the beautiful.”

We now come to the most fundamental theorem of representation theory of compact Lie groups. Before stating and proving this theorem, we need the following lemma:

**27.3.8. Lemma.** *Let  $T : G \rightarrow GL(\mathcal{H})$  be an irreducible unitary representation of a compact Lie group  $G$ . For any nonzero  $|u\rangle, |v\rangle \in \mathcal{H}$ , we have*

$$\frac{1}{\|u\|^2\|v\|^2} \int_G |\langle v | \mathbf{T}_x | u \rangle|^2 d\mu_x = c, \quad (27.50)$$

where  $c > 0$  is a constant independent of  $|u\rangle$  and  $|v\rangle$ .

*Proof.* By Schur's lemma and (2) of Proposition 27.3.7,  $\mathbf{K}_u = \lambda(u)\mathbf{1}$ . Therefore, on the one hand,

$$\langle v | \mathbf{K}_u | v \rangle = \lambda(u) \|v\|^2. \quad (27.51)$$

On the other hand,

$$\langle v | \mathbf{K}_u | v \rangle = \int_G \langle v | \mathbf{T}_x u \rangle \langle \mathbf{T}_x u | v \rangle d\mu_x = \int_G |\langle v | \mathbf{T}_x | u \rangle|^2 d\mu_x. \quad (27.52)$$

Moreover, if we use  $d\mu_g = d\mu_{g^{-1}}$  (see Problem 27.11), then

$$\begin{aligned} \langle v | \mathbf{K}_u | v \rangle &= \int_G \langle v | \mathbf{T}_x | u \rangle \langle u | \mathbf{T}_x^\dagger | v \rangle d\mu_x = \int_G \langle v | \mathbf{T}_x | u \rangle \langle u | \mathbf{T}_x^{-1} | v \rangle d\mu_x \\ &= \int_G \langle u | \mathbf{T}_{x^{-1}} | v \rangle \langle v | \mathbf{T}_x | u \rangle d\mu_x = \int_G \langle u | \mathbf{T}_y | v \rangle \langle v | \mathbf{T}_{y^{-1}} | u \rangle d\mu_{y^{-1}} \\ &= \int_G \langle u | \mathbf{T}_y | v \rangle \underbrace{\langle v | \mathbf{T}^\dagger | u \rangle}_{= \langle \mathbf{T}_y v | u \rangle} \underbrace{d\mu_{y^{-1}}}_{d\mu_y} = \langle u | \mathbf{K}_v | u \rangle. \end{aligned}$$

This equality plus Equation (27.51) gives

$$\lambda(u) \|v\|^2 = \lambda(v) \|u\|^2 \Rightarrow \frac{\lambda(v)}{\|v\|^2} = \frac{\lambda(u)}{\|u\|^2}.$$

Since  $|u\rangle$  and  $|v\rangle$  are arbitrary, we conclude that  $\lambda(u) = c\|u\|^2$  for all  $|u\rangle \in \mathcal{H}$ , where  $c$  is a constant. Equations (27.51) and (27.52) now yield Equation (27.50). If we let  $|u\rangle = |v\rangle$  in Equation (27.52) and use (27.51), we obtain

$$\int_G |\langle u | \mathbf{T}_x | u \rangle|^2 d\mu_x = \lambda(u) \|u\|^2 = c \|u\|^4.$$

That  $c > 0$  follows from the fact that the LHS is a nonnegative continuous function that has at least one strictly positive value in its integration range, namely at  $x = e$ , the identity.  $\square$

**27.3.9. Theorem.** *Every irreducible unitary representation of a compact Lie group is finite-dimensional.*

*Proof.* Let  $\{|e_i\rangle\}_{i=1}^n$  be any set of orthonormal vectors in  $\mathcal{H}$ . Then, unitarity of  $\mathbf{T}_g$  implies that  $\{\mathbf{T}_g |e_i\rangle\}_{i=1}^n$  is also an orthonormal set. Applying Lemma 27.3.8 to  $|e_j\rangle$  and  $|e_1\rangle$ , we obtain

$$\int_G |\langle e_1 | \mathbf{T}_x | e_j \rangle|^2 d\mu_x = c.$$

Now sum over  $j$  to get

$$\begin{aligned} nc &= \sum_{j=1}^n \int_G |\langle e_1 | \mathbf{T}_x | e_j \rangle|^2 d\mu_x = \int_G \sum_{j=1}^n |\langle e_1 | \mathbf{T}_x e_j \rangle|^2 d\mu_x \\ &\leq \int_G \langle e_1 | e_1 \rangle d\mu_x = V_G, \end{aligned}$$

where we used the Parseval inequality [Equation (5.3)] as applied to the vector  $|e_1\rangle$  and the orthonormal set  $\{\mathbf{T}_g |e_i\rangle\}_{i=1}^n$ . Since both  $V_G$  and  $c$  are finite,  $n$  must be finite as well. Thus,  $\mathcal{H}$  cannot have an infinite set of orthonormal vectors.  $\square$

So far, we have discussed irreducible representations. What can we say about arbitrary representations? We recall that in the case of finite groups, every representation can be written as a direct sum of irreducible representations. Is this also true for compact Lie groups?

Firstly, we note that the Weyl operator, being Hilbert–Schmidt, is necessarily compact. It is also hermitian. Therefore, by the spectral theorem, its eigenspaces span the carrier space  $\mathcal{H}$ . Specifically, we can write  $\mathcal{H} = \mathcal{M}_0 \oplus \sum_{j=1}^N \mathcal{M}_j$ , where  $\mathcal{M}_0$  is the eigenspace corresponding to the zero eigenvalue of  $\mathbf{K}_u$ , and  $N$  could be infinity.

Secondly, from the relation  $\langle v | \mathbf{K}_u | v \rangle = c \|u\|^2 \|v\|^2$  and the fact that  $c \neq 0$  and  $|u\rangle \neq 0$ , we conclude that  $\mathbf{K}_u$  cannot have any nonzero eigenvector for its zero eigenvalue. It follows that  $\mathcal{M}_0$  contains only the zero vector. Therefore, if  $\mathcal{H}$  is infinite-dimensional, then  $N = \infty$ .

Thirdly, consider any representation  $T$  of  $G$ . Because  $\mathbf{K}_u$  commutes with all  $\mathbf{T}_g$ , each eigenspace of  $\mathbf{K}_u$  is an invariant subspace under  $T$ . If a subspace  $\mathcal{U}$  is invariant under  $T$ , then  $\mathcal{U} \cap \mathcal{M}_j$ , a subspace of  $\mathcal{M}_j$ , is also invariant (reader, please verify!). Thus, all invariant subspaces of  $G$  are reducible to invariant subspaces of eigenspaces of  $\mathbf{K}_u$ . In particular, all irreducible invariant subspaces of  $T$  are subspaces of eigenspaces of  $\mathbf{K}_u$ .

Lastly, since all  $\mathcal{M}_j$  are finite-dimensional, we can use the procedure used in the case of finite groups and decompose  $\mathcal{M}_j$  into irreducible invariant subspaces of  $T$ . We have just shown the following result:

**27.3.10. Theorem.** *Every unitary representation  $T$  of a compact Lie group  $G$  is a direct sum of irreducible finite-dimensional unitary representations.*

By choosing a basis for the finite-dimensional invariant subspaces of  $T$ , we can represent each  $\mathbf{T}_g$  by a matrix. Therefore,

**27.3.11. Box.** *Compact Lie groups can be represented by matrices.*

As in the case of finite groups, one can work with matrix elements and characters of representations. The only difference is that summations are replaced with integration and order of the group  $|G|$  is replaced with  $V_G$ , which we take to be unitary.<sup>9</sup> For example, Equation (24.6) becomes

$$\int_G T^{(\alpha)}(g) \mathbf{X} T^{(\beta)}(g^{-1}) d\mu_g = \lambda_X \delta_{\alpha\beta} \mathbf{1}, \tag{27.53}$$

and the analogue of Equation (24.8) is

$$\int_G T_{il}^{(\alpha)}(g) T_{jm}^{(\beta)*}(g) d\mu_g = \frac{1}{n_\alpha} \delta_{mi} \delta_{\alpha\beta} \delta_{ij}. \tag{27.54}$$

<sup>9</sup>This can always be done by rescaling the volume element.

Characters satisfy similar relations: Equation (24.11) becomes

$$\int_G \chi^{(\alpha)}(g) \chi^{(\beta)*}(g) d\mu_g = \delta_{\alpha\beta}, \tag{27.55}$$

and the useful Equation (24.16) turns into

$$\int_G |\chi(g)|^2 d\mu_g = \sum_{\alpha} m_{\alpha}^2. \tag{27.56}$$

This formula can be used to test for irreducibility of a representation: If the integral is unity, the representation is irreducible; otherwise, it is reducible.

Finally, we state the celebrated Peter–Weyl theorem (for a proof, see [Baru 86, pp. 172–173])

Peter–Weyl theorem

**27.3.12. Theorem.** (Peter–Weyl theorem) *The functions*

$$\sqrt{n_{\alpha}} T_{ij}^{(\alpha)}(g), \quad \forall \alpha \text{ and } 1 \leq i, j \leq n_{\alpha},$$

*form a complete set of functions in  $\mathcal{L}^2(G)$ , the Hilbert space of square-integrable functions on  $G$ .*

If  $u \in \mathcal{L}^2(G)$ , we can write

$$u(g) = \sum_{\alpha} \sum_{i,j} b_{ij}^{\alpha} T_{ij}^{(\alpha)}(g) \quad \text{where} \quad b_{ij}^{\alpha} = n_{\alpha} \int_G u(g) T_{ij}^{(\alpha)*}(g) d\mu_g. \tag{27.57}$$

**27.3.13. Example.** Equation (27.57) is the generalization of the Fourier series expansion of functions. The connection with Fourier series becomes more transparent if we consider a particular compact group. The unit circle  $S^1$  is a one-dimensional abelian compact 1-parameter Lie group. In fact, fixing an “origin” on the circle, any other point can be described by the parameter  $\theta$ , the angular distance from the point to the origin.  $S^1$  is obviously abelian; it is also compact, because it is a bounded closed region of  $\mathbb{R}^2$  (BWHB theorem). By Theorem 24.2.3, which holds for all Lie groups, all irreducible representations of  $S^1$  are 1-dimensional. So  $T_{ij}^{(\alpha)}(g) \rightarrow T^{(\alpha)}(\theta)$ . Furthermore,  $T^{(\alpha)}(\theta)T^{(\alpha)}(\theta') = T^{(\alpha)}(\theta + \theta')$ . Differentiating both sides with respect to  $\theta'$  at  $\theta' = 0$  yields the differential equation

The Peter–Weyl theorem is the generalization of the Fourier series expansion of periodic functions.

$$T^{(\alpha)}(\theta) \underbrace{\frac{dT^{(\alpha)}}{d\theta'} \Big|_{\theta'=0}}_{\equiv a} = \frac{dT^{(\alpha)}}{dy} \Big|_{y=\theta} \equiv \frac{dT^{(\alpha)}}{d\theta}, \quad y \equiv \theta + \theta'.$$

The solution to this DE is  $Ae^{a\theta}$ . Since  $T^{(\alpha)}$  are unitary, and since a 1-dimensional unitary matrix must look like  $e^{i\varphi}$ , we must have  $A = 1$ . Furthermore,  $\theta$  and  $\theta + 2\pi$  are identified on the unit circle; therefore, we must conclude that  $a$  is  $i$  times an integer  $n$ , which determines the irreducible representation. We label the irreducible representation by  $n$  and write

$$T^{(n)}(\theta) = e^{in\theta}, \quad n = 0 \pm 1 \pm 2 \dots$$

The Peter–Weyl theorem now becomes the rule of Fourier series expansion of *periodic functions*. This last property follows from the fact that any function  $u : S^1 \rightarrow \mathbb{R}$  is necessarily periodic. ■

There are many occasions in physics where the state functions describing physical quantities transform irreducibly under the action of a Lie group (which we assume to be compact). Often this Lie group also acts on the underlying space-time manifold. So we have a situation in which a Lie group  $G$  acts on a Euclidean space  $\mathbb{R}^n$  as well as on the space of (square-integrable) functions  $\mathcal{L}(\mathbb{R}^n)$ . Therefore, the functions  $\{\phi_i^{(\alpha)}(\mathbf{x})\}$ , belonging to the  $\alpha$ th irreducible representation transform among themselves not only because of the index  $i$ , but also because of the argument  $\mathbf{x} \in \mathbb{R}^n$ .

To see the connection between physics and representation theory, consider the transformation of the simplest case, a scalar function. As a concrete example, choose temperature. To observer  $O$  at the corner of a room 8 meters long, 6 meters wide, and 3 meters high, the temperature of the center of the room is given by  $\theta(4, 3, 1.5)$  where  $\theta(x, y, z)$  is a function that gives  $O$  the temperature of various points of the room. Observer  $O'$  is sitting in the middle of the floor, so that the center of the room has coordinates  $(0, 0, 1.5)$ .  $O'$  also has a function that gives her the the temperature at various points. But this function must necessarily be different from  $\theta$  because of the different coordinates the same points have for  $O$  and  $O'$ . Calling this function  $\theta'$ , we have  $\theta'(0, 0, 1.5) = \theta(4, 3, 1.5)$ , and in general,

$$\theta'(x', y', z') = \theta(x, y, z),$$

where  $(x', y', z')$  describes the same point for  $O'$  that  $(x, y, z)$  describes for  $O$ .

In the context of representation theory, we can think of  $(x', y', z')$  as the transformed coordinates obtained as a result of the action of some group:  $(x', y', z') = g \cdot (x, y, z)$ , or  $\mathbf{x}' = g \cdot \mathbf{x}$ . So, the equation above can be written as

$$\theta'(\mathbf{x}') = \theta(\mathbf{x}) = \theta(g^{-1} \cdot \mathbf{x}') \quad \text{or} \quad \theta'(\mathbf{x}) = \theta(g^{-1} \cdot \mathbf{x}).$$

It is natural to call  $\theta'$  the transform of  $\theta$  under the action of  $g$  and write  $\theta' = \mathbf{T}_g \theta$ . This is one way of constructing a representation [see the comments after Equation (24.1)]. Instead of  $g^{-1}$  on the left, one could act with  $g$  on the right.

When the physical quantity is not a scalar, it is natural to group together the smallest set of functions that transform into one another. This leads to the set of functions that transform according to a row of an irreducible representation of the group. In some sense, this situation is a combination of (24.1) and (24.37). The reader may verify that

$$\mathbf{T}_g \phi_i^{(\alpha)}(\mathbf{x}) = \sum_{j=1}^{n_\alpha} T_{ji}^{(\alpha)}(g) \phi_j^{(\alpha)}(\mathbf{x} \cdot g^{-1}) \quad (27.58)$$

defines a representation of  $G$ .

We now use Box 27.1.26 to construct an irreducible representation of the Lie algebra of  $G$  from Equation (27.58). By the definition of the infinitesimal action,



we let  $g = \exp(\xi t)$  and differentiate both sides with respect to  $t$  at  $t = 0$ . This yields

$$\begin{aligned} \underbrace{\frac{d}{dt} \mathbf{T}_{\exp(\xi t)} \phi_i^{(\alpha)}(\mathbf{x})}_{\equiv \sum_j \mathfrak{D}_{ji}(\xi) \phi_j^{(\alpha)}(\mathbf{x})} \Big|_{t=0} &= \sum_{j=1}^{n_\alpha} \frac{d}{dt} \left\{ T_{ji}^{(\alpha)}(\exp(\xi t)) \phi_j^{(\alpha)}(\mathbf{x} \cdot \exp(-\xi t)) \right\} \Big|_{t=0} \\ &= \sum_{j=1}^{n_\alpha} \frac{d}{dt} T_{ji}^{(\alpha)}(\exp(\xi t)) \Big|_{t=0} \underbrace{\phi_j^{(\alpha)}(\mathbf{x} \cdot \exp(-\xi 0))}_{=\mathbf{x}} \\ &\quad + \sum_{j=1}^{n_\alpha} \underbrace{T_{ji}^{(\alpha)}(\exp(\xi 0))}_{=\delta_{ji}} \frac{d}{dt} \phi_j^{(\alpha)}(\mathbf{x} \cdot \exp(-\xi t)) \Big|_{t=0}, \end{aligned}$$

where we have defined the matrices  $\mathfrak{D}_{ji}(\xi)$  for the LHS. The derivative in the first sum is simply  $\mathfrak{T}_{ji}^{(\alpha)}(\xi)$  the representation of the generator  $\xi$  of the 1-parameter group of transformations in the space of functions  $\{\phi_i^{(\alpha)}\}$ . The derivative in the second sum can be found by writing  $\mathbf{x}'(t) = \mathbf{x} \cdot \exp(-\xi t)$  and differentiating as follows:

$$\begin{aligned} \frac{d}{dt} \phi_j^{(\alpha)}(\mathbf{x}'(t)) \Big|_{t=0} &\equiv \frac{d}{dt} \phi_j^{(\alpha)}(x^1(t), \dots, x^m(t)) \Big|_{t=0} = \partial_k \phi_j^{(\alpha)} \frac{d}{dt} (x^k(t)) \Big|_{t=0} \\ &= \partial_k \phi_j^{(\alpha)} u_\mu^k(\mathbf{x}) \frac{da^\mu}{dt} \equiv \partial_k \phi_j^{(\alpha)} X^k(\mathbf{x}; \xi) \equiv \partial_\nu \phi_j^{(\alpha)} X^\nu(\mathbf{x}; \xi), \end{aligned}$$

where we used Equation (27.18) and defined  $X^k(\mathbf{x}; \xi)$  by the last equality. We also changed the coordinate index to Greek to avoid confusing it with the index of the functions. Collecting everything together, we obtain

$$\sum_{j=1}^{n_\alpha} \mathfrak{D}_{ij}(\xi) \phi_j^{(\alpha)}(\mathbf{x}) = \sum_{j=1}^{n_\alpha} \mathfrak{T}_{ij}^{(\alpha)}(\xi) \phi_j^{(\alpha)}(\mathbf{x}) + \sum_{j=1}^{n_\alpha} \delta_{ij} X^\nu(\mathbf{x}; \xi) \frac{\partial \phi_j^{(\alpha)}}{\partial x^\nu},$$

or, since  $\phi_j^{(\alpha)} = \sum_k \phi_k^{(\alpha)} (\partial/\partial \phi_k^{(\alpha)}) \phi_j^{(\alpha)}$ ,

$$\mathfrak{D}_{ij}(\xi) = \mathfrak{T}_{ij}^{(\alpha)}(\xi) \phi_k^{(\alpha)}(\mathbf{x}) \frac{\partial}{\partial \phi_k^{(\alpha)}} + \delta_{ij} X^\nu(\mathbf{x}; \xi) \frac{\partial}{\partial x^\nu}, \quad (27.59)$$

where  $X^\nu(\mathbf{x}; \xi)$  is the  $\nu$ th component of the infinitesimal generator of the action induced by  $\xi \in \mathfrak{g}$ . We shall put Equation (27.59) to good use when we discuss symmetries and conservation laws in Chapter 30. The derivative with respect to the functions, although meaningless at this point, will be necessary when we discuss conservation laws.

### 27.4 Representation of the General Linear Group

$GL(\mathcal{V})$  is not a compact group, but we can use the experience we gained in the analysis of the symmetric group to find the irreducible representations of  $GL(\mathcal{V})$ . The key is to construct tensor product spaces of  $\mathcal{V}$ —which, as the reader may verify, is a carrier space of  $GL(\mathcal{V})$ —and look for its irreducible subspaces. In fact, if  $r$  is an arbitrary positive integer,  $T : G \rightarrow GL(\mathcal{V})$  is a representation, and

$$\mathcal{V}^{\otimes r} \equiv \underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{r \text{ times}}.$$

Then  $T^{\otimes r} : G \rightarrow GL(\mathcal{V}^{\otimes r})$ , given by

$$[T^{\otimes r}(g)](\mathbf{v}_1, \dots, \mathbf{v}_r) \equiv \mathbf{T}_g^{\otimes r}(\mathbf{v}_1, \dots, \mathbf{v}_r) = \mathbf{T}_g(\mathbf{v}_1) \otimes \cdots \otimes \mathbf{T}_g(\mathbf{v}_r),$$

is also a representation. In particular, considering  $\mathcal{V}$  as the (natural) carrier space for  $GL(\mathcal{V})$ , we conclude that  $T^{\otimes r} : GL(\mathcal{V}) \rightarrow GL(\mathcal{V}^{\otimes r})$  is a representation.

This tensor product representation is reducible, because as is evident from its definition,  $\mathbf{T}_g^{\otimes r}$  preserves any symmetry of the tensor it acts on. For example, the subspace of the full  $n^r$ -dimensional tensor product space consisting of the completely symmetric tensors of the type

$$\mathbf{t}_s \equiv \sum_{\pi \in S_r} \mathbf{v}_{\pi(1)} \otimes \mathbf{v}_{\pi(2)} \otimes \cdots \otimes \mathbf{v}_{\pi(r)}$$

is invariant. Similarly, the subspace consisting of the completely antisymmetric tensor products—the  $r$ -fold wedge products—is invariant.

To reduce  $\mathcal{V}^{\otimes r}$ , we choose a basis  $\{\mathbf{e}_k\}_{k=1}^n$  for  $\mathcal{V}$ . Then the collection of  $n^r$  tensor products  $\{\mathbf{e}_{k_1} \otimes \cdots \otimes \mathbf{e}_{k_r}\}$ , where each  $k_i$  runs from 1 to  $n$ , is a basis for  $\mathcal{V}^{\otimes r}$ . An invariant subspace of  $\mathcal{V}^{\otimes r}$  is a span of linear combinations of certain of these basis vectors. Since the only thing that distinguishes among  $\{\mathbf{e}_{k_1} \otimes \cdots \otimes \mathbf{e}_{k_r}\}$  is a permutation of the  $r$  labels, we start to see the connection between the reduction of  $\mathcal{V}^{\otimes r}$  and  $S_r$ . This connection becomes more evident if we recall that the left multiplication of the group algebra of  $S_r$  by its elements provides the regular representation, which is reducible. The irreducible representations are the minimal ideals of the algebra generated by the Young operators.

The same idea works here as well: Certain linear combination of the basis vectors of  $\mathcal{V}^{\otimes r}$  obtained by permutations can serve as the basis vectors for irreducible representations of  $GL(\mathcal{V})$ . Let us elaborate on this. Recall that a Young operator of  $S_r$  is written in the form  $Y = QP$  where  $Q$  and  $P$  are linear combinations of permutations in  $S_r$ .  $Y$  has the property that if one operates on it (via left multiplication) with all permutations of  $S_r$ , one generates a minimal ideal, i.e., an *irreducible representation* of  $S_r$ . Now let  $\mathbf{Y}$  be a Young operator that acts on the indices  $(k_1, \dots, k_r)$ , giving linear combinations of the basis vectors of  $\mathcal{V}^{\otimes r}$ . From the minimality of the ideal generated by  $Y$  and the fact that operators in  $GL(\mathcal{V})$

permute the factors in  $\mathbf{e}_{k_1} \otimes \cdots \otimes \mathbf{e}_{k_r}$  in all possible ways, it should now be clear that if we choose any *single* basis vector  $\mathbf{e}_{k_1} \otimes \cdots \otimes \mathbf{e}_{k_r}$ , then  $\mathbf{Y}(\mathbf{e}_{k_1} \otimes \cdots \otimes \mathbf{e}_{k_r})$  generates an irreducible representation of  $GL(\mathcal{V})$ . We therefore have the following:

connection between the Young tableaux and irreducible representations of  $GL(\mathcal{V})$

**27.4.1. Theorem.** *Let  $\{\mathbf{e}_k\}_{k=1}^n$  be any basis for  $\mathcal{V}$ . Let  $\mathbf{Y} = \mathbf{QP}$  be the Young operator of  $S_r$  that permutes (and takes linear combinations of) the basis vectors  $\{\mathbf{e}_{k_1} \otimes \cdots \otimes \mathbf{e}_{k_r}\}$ . Then for any given such basis vector, the vectors*

$$\{\mathbf{T}_g^{\otimes r} \mathbf{Y}(\mathbf{e}_{k_1} \otimes \cdots \otimes \mathbf{e}_{k_r}) \mid g \in GL(\mathcal{V})\}$$

*span an irreducible subspace of  $\mathcal{V}^{\otimes r}$ .*

A basis of such an irreducible representation can be obtained by taking into account all the Young tableaux associated with the irreducible representation. But which of the symmetry types will be realized for given values of  $n$  and  $r$ ? Clearly, the Young tableau should not contain more than  $n$  rows, because then one of the symbols will be repeated in a column, and the Young operator will vanish due to the antisymmetry in its column indices. We can therefore restrict the partition  $(\lambda)$  to

$$(\lambda) = (\lambda_1, \lambda_2, \dots, \lambda_n), \quad \lambda_1 + \cdots + \lambda_n = r, \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

Let us consider an example for clarification.

**27.4.2. Example.** First, let  $n = r = 2$ . The tensor product space has  $2^2 = 4$  dimensions. To reduce it, we consider the Young operators, which correspond to  $e + (k_1, k_2)$  and  $e - (k_1, k_2)$ . Let us denote these operators by  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$ , respectively. By applying each one to a generic basis vector  $\mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2}$ , we can generate all the irreducible representations. The first operator gives

$$\mathbf{Y}_1(\mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2}) = \mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2} + \mathbf{e}_{k_2} \otimes \mathbf{e}_{k_1},$$

where  $k_1$  and  $k_2$  can be 1 or 2. For  $k_1 = k_2 = 1$ , we get  $2\mathbf{e}_1 \otimes \mathbf{e}_1$ . For  $k_1 = 1, k_2 = 2$ , or  $k_1 = 2, k_2 = 1$ , we get  $\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1$ . Finally, for  $k_1 = k_2 = 2$ , we get  $2\mathbf{e}_2 \otimes \mathbf{e}_2$ . Altogether, we obtain 3 linearly independent vectors that are completely symmetric.

When the second operator acts on a generic basis vector, it gives

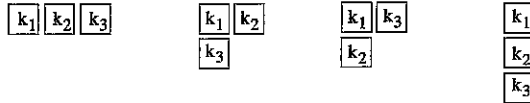
$$\mathbf{Y}_2(\mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2}) = \mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2} - \mathbf{e}_{k_2} \otimes \mathbf{e}_{k_1}.$$

The only time that this is not zero is when  $k_1$  and  $k_2$  are different. In either case, we get  $\pm(\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1)$ . This subspace is therefore one-dimensional.

The reduction of the tensor product space can therefore be written as

$$\mathcal{V}^{\otimes 2} = \underbrace{\text{Span}\{\mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1, \mathbf{e}_2 \otimes \mathbf{e}_2\}}_{\text{3D symmetric subspace}} \oplus \underbrace{\text{Span}\{\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1\}}_{\text{1D antisymmetric subspace}}.$$

Next, let us consider the case of  $n = 2, r = 3$ . The tensor product space has  $2^3 = 8$  dimensions. To reduce it, we need to consider all Young operators of  $S_3$ . There are four of these, corresponding to the following tableaux:



Let us denote these operators by  $Y_1, Y_2, Y_3,$  and  $Y_4,$  respectively. By applying each one to a generic basis vector  $e_{k_1} \otimes e_{k_2} \otimes e_{k_3},$  we can generate all the irreducible representations. The first operator gives

$$Y_1(e_{k_1} \otimes e_{k_2} \otimes e_{k_3}) = e_{k_1} \otimes e_{k_2} \otimes e_{k_3} + e_{k_1} \otimes e_{k_3} \otimes e_{k_2} + e_{k_2} \otimes e_{k_1} \otimes e_{k_3} + e_{k_2} \otimes e_{k_3} \otimes e_{k_1} + e_{k_3} \otimes e_{k_1} \otimes e_{k_2} + e_{k_3} \otimes e_{k_2} \otimes e_{k_1},$$

where  $k_1, k_2,$  and  $k_3$  can be 1 or 2. For  $k_1 = k_2 = k_3 = 1,$  we get  $6e_1 \otimes e_1 \otimes e_1.$  For the case where two of the  $k_i$ 's are 1 and the third is 2, we get

$$2(e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1).$$

For the case where two of the  $k_i$ 's are 2 and the third is 1, we get

$$2(e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_1).$$

Finally, for  $k_1 = k_2 = k_3 = 2,$  we get  $6e_2 \otimes e_2 \otimes e_2.$  Altogether, we obtain 4 linearly independent vectors that are completely symmetric.

When the second operator acts on a generic basis vector, it gives<sup>10</sup>

$$Y_2(e_{k_1} \otimes e_{k_2} \otimes e_{k_3}) = [e - (k_1, k_3)][e + (k_1, k_2)](e_{k_1} \otimes e_{k_2} \otimes e_{k_3}) = e_{k_1} \otimes e_{k_2} \otimes e_{k_3} + e_{k_2} \otimes e_{k_1} \otimes e_{k_3} - e_{k_3} \otimes e_{k_2} \otimes e_{k_1} - e_{k_2} \otimes e_{k_3} \otimes e_{k_1}.$$

If all three indices are the same, we get zero. Suppose  $k_1 = 1.$  Then  $k_2$  can be 1 or 2. For  $k_2 = 1,$  we must set  $k_3 = 2$  to get  $e_2 \otimes e_1 \otimes e_1 - e_1 \otimes e_2 \otimes e_1.$  For  $k_2 = 2,$  we must set  $k_3 = 1$  to obtain  $e_1 \otimes e_2 \otimes e_2 - e_2 \otimes e_1 \otimes e_2.$  If we start with  $k_1 = 2,$  we will not produce any new vectors, as the reader is urged to verify. Therefore, the dimension of the irreducible subspace spanned by the second Young tableau is 2.

The action of the third operator on a generic basis vector yields

$$Y_3(e_{k_1} \otimes e_{k_2} \otimes e_{k_3}) = [e - (k_1, k_2)][e + (k_1, k_3)](e_{k_1} \otimes e_{k_2} \otimes e_{k_3}) = e_{k_1} \otimes e_{k_2} \otimes e_{k_3} + e_{k_2} \otimes e_{k_1} \otimes e_{k_3} - e_{k_3} \otimes e_{k_2} \otimes e_{k_1} - e_{k_3} \otimes e_{k_1} \otimes e_{k_2}.$$

The reader may check that we obtain a two-dimensional irreducible representation spanned by  $e_1 \otimes e_1 \otimes e_2 - e_2 \otimes e_1 \otimes e_1$  and  $e_1 \otimes e_2 \otimes e_2 - e_2 \otimes e_2 \otimes e_1.$

The fourth Young operator gives zero because it is completely antisymmetric in three slots and we have only two indices. The reduction of the tensor product space can therefore

<sup>10</sup>When a symmetric group is considered as an abstract group—as opposed to a group of transformations—we may multiply permutations (keep track of how each number is repeatedly transformed) from left to right. However, since the permutations here act on vectors on their right, it is more natural to calculate their products from right to left.

be written as

$$\begin{aligned} \mathcal{V}^{\otimes 3} = & \underbrace{\text{Span}\{\mathbf{Y}_1(\mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2} \otimes \mathbf{e}_{k_3})\}}_{\dim=4} \oplus \underbrace{\text{Span}\{\mathbf{Y}_2(\mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2} \otimes \mathbf{e}_{k_3})\}}_{\dim=2} \\ & \oplus \underbrace{\text{Span}\{\mathbf{Y}_3(\mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2} \otimes \mathbf{e}_{k_3})\}}_{\dim=2}. \end{aligned}$$

We note that the total dimensions on both sides match. ■

There is a remarkable formula that gives the dimension of all irreducible representations of  $GL(\mathcal{V})$  (see [Hame 89, pp. 384–387] for a derivation):

**27.4.3. Theorem.** *Let  $\mathcal{V}$  be an  $n$ -dimensional vector space, and  $\mathcal{V}^{(\lambda)}$ , the irreducible subspace of tensors with symmetry associated with the partition  $(\lambda) = (\lambda_1, \dots, \lambda_n)$ . Then*

$$\dim \mathcal{V}^{(\lambda)} = \frac{D(l_1, \dots, l_n)}{D(n-1, n-2, \dots, 0)},$$

where  $l_j \equiv \lambda_j + n - j$  and  $D(x_1, \dots, x_n)$  is as given in Equation (24.56).

## 27.5 Representation of Lie Algebras

The diffeomorphism established by the exponential map (Theorem 27.1.18) reduces the *local* study of a Lie group to that of its Lie algebra.<sup>11</sup> In this book, we are exclusively interested in the local properties of Lie groups, and we shall therefore confine ourselves to Lie algebras to study the structure of Lie groups. Recall that any Lie group homomorphism leads to a corresponding Lie algebra homomorphism [Equation (27.7)]. Conversely, a homomorphism of Lie algebras can, through the identification of the neighborhoods of their identities with their Lie algebras, be “exponentiated” to a (local) homomorphism of their Lie groups. This leads to the following theorem (see [Fult 91, pp. 108 and 119] for a proof).

**27.5.1. Theorem.** *Let  $G$  be a Lie group with algebra  $\mathfrak{g}$ . A representation  $T : G \rightarrow GL(\mathcal{H})$  determines a Lie algebra representation  $T_* : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{H})$ . Conversely, a Lie algebra representation  $\mathfrak{T} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{H})$  determines a Lie group representation.*

It follows from this theorem that all (local) Lie group representations result from corresponding Lie algebra representations. Therefore, we shall restrict ourselves to the representations of Lie algebras.

<sup>11</sup>We use the word “local” to mean the collection of all points that can be connected to the identity by a curve in the Lie group  $G$ . If this collection exhausts  $G$ , then we say that  $G$  is **connected**. If, furthermore, all closed curves (loops) in  $G$  can be shrunk to a point, we say that  $G$  is **simply connected**. The word “local” can be replaced by “simply connected” in what follows.

### 27.5.1 Representation of Subgroups of $GL(\mathcal{V})$

Let  $\mathfrak{g}$  be any Lie algebra with basis vectors  $\{X_i\}$ . Let a representation  $T$  map these vectors to  $\{T_i\} \in \mathfrak{gl}(\mathcal{H})$  for some carrier space  $\mathcal{H}$ . Then, a general element  $X = \sum_i \alpha_i X_i$  of  $\mathfrak{g}$  will be mapped to  $T = \sum_i \alpha_i T_i$ . Now suppose that  $\mathfrak{h}$  is a subalgebra of  $\mathfrak{g}$ . Then the restriction of  $T$  to  $\mathfrak{h}$  provides a representation of  $\mathfrak{h}$ . This restriction may be reducible. If it is, then there is an invariant subspace  $\mathcal{H}_1$  of  $\mathcal{H}$ . It follows that

$$\langle b | T_X | a \rangle = 0 \quad \forall X \in \mathfrak{h} \quad \text{whenever } |a\rangle \in \mathcal{H}_1 \text{ and } |b\rangle \in \mathcal{H}_1^\perp,$$

where  $T_X \equiv T(X)$ . If we write  $T_X = \sum_i \alpha_i^{(X)} T_i$ , then in terms of  $T_i$ , the equation above can be written as

$$\sum_{i=1}^{\dim \mathfrak{g}} \alpha_i^{(X)} \langle b | T_i | a \rangle \equiv \sum_{i=1}^{\dim \mathfrak{g}} \alpha_i^{(X)} \tau_i^{(ba)} = 0 \quad \forall X \in \mathfrak{h}, \tag{27.60}$$

where  $\tau_i^{(ba)} \equiv \langle b | T_i | a \rangle$  are complex numbers. Equation (27.60) states that

**27.5.2. Box.** *If  $T$ , as a representation of  $\mathfrak{h}$  (a Lie subalgebra of  $\mathfrak{g}$ ), is reducible, then there exist a number of equations that  $\alpha_i^{(X)}$  must satisfy whenever  $X \in \mathfrak{h}$ . If  $T$ , as a representation of  $\mathfrak{g}$ , is irreducible, then no relation such as given in (27.60) will exist when  $X$  runs over all of  $\mathfrak{g}$ .*

This last statement will be used to analyze certain subgroups of  $GL(\mathcal{V})$ .

Let us first identify  $GL(\mathcal{V})$  with  $GL(n, \mathbb{C})$ . Next, consider  $GL(n, \mathbb{R})$ , which is a subgroup of  $GL(n, \mathbb{C})$ , and transfer the discussion to their respective algebras. If  $\{X_i\}$  is a basis of  $\mathfrak{gl}(n, \mathbb{C})$ , then an arbitrary element can be written as  $\sum_i \alpha_i X_i$ . The difference between  $\mathfrak{gl}(n, \mathbb{C})$  and  $\mathfrak{gl}(n, \mathbb{R})$  is that the  $\alpha_i$ 's are *real* in the latter case; i.e., for *all* real values of  $\{\alpha_i\}$ , the sum belongs to  $\mathfrak{gl}(n, \mathbb{R})$ . Now suppose that  $T$  is an irreducible representation of  $\mathfrak{gl}(n, \mathbb{C})$  that is reducible when restricted to  $\mathfrak{gl}(n, \mathbb{R})$ . Equation (27.60) states that the function

$$f(z_1, \dots, z_{n^2}) \equiv \sum_{i=1}^{n^2} z_i \tau_i^{(ba)}$$

vanishes for *all* real values of the  $z_i$ 's. Since this function is obviously entire, it must vanish for all *complex* values of  $z_i$ 's by analytic continuation (see Theorem 11.3.1). But this is impossible because  $T$  is irreducible for  $\mathfrak{gl}(n, \mathbb{C})$ . We have to conclude that  $T$  is irreducible as a representation of  $\mathfrak{gl}(n, \mathbb{R})$ .

The next subalgebra of  $\mathfrak{gl}(n, \mathbb{C})$  we consider is the Lie algebra  $\mathfrak{sl}(n, \mathbb{C})$  of the special linear group. The only restriction on the elements of  $\mathfrak{sl}(n, \mathbb{C})$  is for them

to have a vanishing trace. Denoting  $\text{tr } \mathbf{X}_i$  by  $t_i$ , we conclude that  $\mathbf{X} = \sum_i \alpha_i \mathbf{X}_i$  belongs to  $\mathfrak{sl}(n, \mathbb{C})$  if and only if  $\sum_i \alpha_i t_i = 0$ . Let  $(t_1^*, \dots, t_{n^2}^*) \equiv |t\rangle \in \mathbb{C}^{n^2}$ . Then  $\mathfrak{sl}(n, \mathbb{C})$  can be characterized as the subspace consisting of vectors  $|a\rangle \in \mathbb{C}^{n^2}$  such that  $\langle a | t \rangle = 0$ . Such a subspace has  $n^2 - 1$  dimensions. If any irreducible representation of  $\mathfrak{gl}(n, \mathbb{C})$  is reducible for  $\mathfrak{sl}(n, \mathbb{C})$ , then the set of complex numbers  $\{\alpha_i\}$  must, in addition, satisfy Equation (27.60). This amounts to the condition that  $|a\rangle$  be orthogonal to  $|\tau^{(ba)}\rangle$  as well. But this is impossible, because then the set  $\{|a\rangle, |t\rangle, |\tau^{(ba)}\rangle\}$  would constitute a subspace of  $\mathbb{C}^{n^2}$  whose dimension is at least  $n^2 + 1$ : There are  $n^2 - 1$  of  $|a\rangle$ 's, one  $|t\rangle$ , and at least one  $|\tau^{(ba)}\rangle$ . Therefore, all irreducible representations of  $\mathfrak{gl}(n, \mathbb{C})$  are also irreducible representations of  $\mathfrak{sl}(n, \mathbb{C})$ . The last subalgebra of  $\mathfrak{gl}(n, \mathbb{C})$  we consider is the Lie algebra  $\mathfrak{u}(n)$  of the unitary group. To study this algebra, we start with the Weyl basis of Equation (27.27) for  $\mathfrak{gl}(n, \mathbb{C})$ , and construct a new *hermitian* basis  $\{\mathbf{X}_{kj}\}$  defined as

$$\begin{aligned} \mathbf{X}_{jj} &\equiv \mathbf{e}_{jj} && \text{for all } j = 1, 2, \dots, n, \\ \mathbf{X}_{kj} &\equiv i(\mathbf{e}_{kj} - \mathbf{e}_{kj}^\dagger) && \text{if } k \neq j. \end{aligned}$$

A typical element of  $\mathfrak{gl}(n, \mathbb{C})$  is of the form  $\sum_{kj} \alpha_{kj} \mathbf{X}_{kj}$ , where  $\alpha_{kj}$  are complex numbers. If we restrict ourselves to *real* values of  $\alpha_{kj}$ , then we obtain the subalgebra of hermitian matrices whose Lie group is the unitary group  $U(n)$ . The fact that the irreducible representations of  $\mathfrak{gl}(n, \mathbb{C})$  will not reduce under  $\mathfrak{u}(n)$  follows immediately from our discussion concerning  $\mathfrak{gl}(n, \mathbb{R})$ . We summarize our findings in the following:

**27.5.3. Theorem.** *The irreducible representations of  $GL(n, \mathbb{C})$  are also irreducible representations of  $GL(n, \mathbb{R})$ ,  $SL(n, \mathbb{C})$ ,  $U(n)$ , and  $SU(n)$ .*

The case of  $SU(n)$  follows from the same argument given earlier that connected  $GL(n, \mathbb{C})$  to  $SL(n, \mathbb{C})$ .

### 27.5.2 Casimir Operators

In the general representation theory of Lie algebras, it is desirable to label each irreducible representation with a quantity made out of the basis vectors of the Lie algebra. An example is the labeling of the energy states of a quantum mechanical system with angular momentum. Each value of the total angular momentum labels an irreducible subspace whose vectors are further labeled by the third component of angular momentum (see Chapter 12). This subsection is devoted to the generalization of this concept to an arbitrary Lie algebra.

**27.5.4. Definition.** *Let  $\mathfrak{T} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{H})$  be a representation of the Lie algebra  $\mathfrak{g}$ . A **Casimir operator**  $\mathbf{C}$  for this representation is an operator that commutes with all  $\mathbf{T}_{\mathbf{X}}$  of the representation.*

Casimir operator  
defined

If the representation is irreducible, then by Schur's lemma,  $\mathbf{C}$  is a multiple of the unit operator. Therefore, all vectors of an irreducible invariant subspace of

Chevalley's theorem the carrier space  $\mathcal{H}$  are eigenvectors of  $\mathbf{C}$  corresponding to the same eigenvalue. That Casimir operators actually determine the irreducible representations of a semisimple Lie algebra is the content of the following theorem (for a proof, see [Vara 84, pp. 333–337]).

**27.5.5. Theorem.** (Chevalley) *For every semisimple Lie algebra  $\mathfrak{g}$  of rank<sup>12</sup>  $r$  with a basis  $\{\mathbf{X}_i\}$ , there exists a set of  $r$  Casimir operators in the form of polynomials in  $\mathbf{T}_{\mathbf{X}_i}$  whose eigenvalues characterize the irreducible representations of  $\mathfrak{g}$ .*

From now on, we shall use the notation  $\mathbf{X}_i$  for  $\mathbf{T}_{\mathbf{X}_i}$ . It follows from Theorem 27.5.5 that all irreducible invariant vector subspaces of the carrier space can be labeled by the eigenvalues of the  $r$  Casimir operators. This means that each invariant irreducible subspace has a basis all of whose vectors carry a set of  $r$  labels corresponding to the eigenvalues of the  $r$  Casimir operators.

One Casimir operator—in the form of a polynomial of degree two—which works only for semisimple Lie algebras, is obtained easily:

$$\mathbf{c} = \sum_{i,j} g^{ij} \mathbf{X}_i \mathbf{X}_j, \quad (27.61)$$

where  $g^{ij}$  is the inverse of the Cartan metric tensor. In fact, with the summation convention in place, we have

$$\begin{aligned} [\mathbf{C}, \mathbf{X}_k] &= g^{ij} [\mathbf{X}_i \mathbf{X}_j, \mathbf{X}_k] = g^{ij} \{ \mathbf{X}_i [\mathbf{X}_j, \mathbf{X}_k] + [\mathbf{X}_i, \mathbf{X}_k] \mathbf{X}_j \} \\ &= g^{ij} \{ c_{jk}^r \mathbf{X}_i \mathbf{X}_r + c_{ik}^r \mathbf{X}_r \mathbf{X}_j \} \\ &= g^{ij} c_{ik}^r (\mathbf{X}_j \mathbf{X}_r + \mathbf{X}_r \mathbf{X}_j) \quad (\text{because } g^{ij} \text{ is symmetric}) \\ &= g^{ij} g^{sr} c_{iks} (\mathbf{X}_j \mathbf{X}_r + \mathbf{X}_r \mathbf{X}_j) \\ &= 0. \quad (\text{because } g^{ij} g^{sr} c_{iks} \text{ is antisymmetric in } j, r) \end{aligned}$$

The last equality follows from the fact that  $g^{ij}$  and  $g^{sr}$  are symmetric,  $c_{iks}$  is completely antisymmetric [see the discussion following Equation (27.45)], and there is a sum over the dummy index  $s$ .

**27.5.6. Example.** The rotation group  $SO(3)$  in  $\mathbb{R}^3$  is a compact 3-parameter Lie group. The infinitesimal generators are the three components of the angular momentum operator (see Example 27.1.29). From the commutation relations of the angular momentum operators developed in Chapter 12, we conclude that  $c_{ij}^k = i\epsilon_{ijk}$ . It follows that the Cartan metric tensor is

$$g_{ij} = c_{is}^r c_{jr}^s = (i\epsilon_{isr})(i\epsilon_{jrs}) = +\epsilon_{isr}\epsilon_{jrs} = 2\delta_{ij}.$$

Ignoring the factor of 2 and denoting the angular momentum operators by  $\mathbf{L}_i$ , we conclude that

$$\mathbf{L}^2 \equiv \mathbf{L}_1^2 + \mathbf{L}_2^2 + \mathbf{L}_3^2$$

<sup>12</sup>Recall that the rank of  $\mathfrak{g}$  is the dimension of the Cartan subalgebra of  $\mathfrak{g}$ .



irreducible representations of the rotation group and spherical harmonics

is a Casimir operator. But this is precisely the operator discussed in detail in Chapter 12. We found there that the eigenvalues of  $L^2$  were labeled by  $j$ , where  $j$  was either an integer or a half odd integer. In the context of our present discussion, we note that the Lie algebra  $\mathfrak{so}(3)$  has rank one, because there is no higher dimensional subalgebra of  $\mathfrak{so}(3)$  all of whose vectors commute with one another. It follows from Theorem 27.5.5 that  $L^2$  is the only Casimir operator, and that all irreducible representations  $T^{(j)}$  of  $\mathfrak{so}(3)$  are distinguished by their label  $j$ . Furthermore, the construction of Chapter 12 showed explicitly that the dimension of  $T^{(j)}$  is  $2j + 1$ .

The connection between the representation of Lie algebras and Lie groups permits us to conclude that the irreducible representations of the rotation group are labeled by the (half) integers  $j$ , and the  $j$ th irreducible representation has dimension  $2j + 1$ . When  $j$  is an integer  $l$  and the carrier space is  $\mathcal{L}^2(S^2)$ , the square-integrable functions on the unit sphere, then  $L^2$  becomes a *differential operator*, and the spherical harmonics  $Y_{lm}(\theta, \varphi)$ , with a fixed value of  $l$ , provide a basis for the  $l$ th irreducible invariant subspace. ■

Connection between Casimir operators and the PDEs of mathematical physics

The last sentence of Example 27.5.6 is at the heart of the connection between symmetry, Lie groups, and the equations of mathematical physics. A symmetry operation of mathematical physics is expressed in terms of the action of a Lie group on an underlying manifold  $M$ , i.e., as a group of transformations of  $M$ . The Lie algebra of such a Lie group consists of the infinitesimal generators of the corresponding transformation. These generators can be expressed as first-order differential operators as in Equation (27.20). It is therefore natural to choose as the carrier space of a representation the Hilbert space  $\mathcal{L}^2(M)$  of the square-integrable functions on  $M$ , which, through the local identification of  $M$  with  $\mathbb{R}^m$  ( $m = \dim M$ ), can be identified with functions on  $\mathbb{R}^m$ . Then the infinitesimal generators act directly on the functions of  $\mathcal{L}^2(M)$  as first-order differential operators.

The Casimir operators  $\{\mathbf{C}_\alpha\}_{\alpha=1}^r$ , where  $r$  is the rank of the Lie algebra, are polynomials in the infinitesimal generators, i.e., differential operators of higher order. On the irreducible invariant subspaces of  $\mathcal{L}^2(M)$ , each  $\mathbf{C}_\alpha$  acts as a multiple of the identity, so if  $f(\mathbf{r})$  belongs to such an invariant subspace, we have

$$\mathbf{C}_\alpha f(\mathbf{r}) = \lambda(\alpha) f(\mathbf{r}), \quad \alpha = 1, 2, \dots, r. \tag{27.62}$$

This is a set of differential equations that are invariant under the symmetry of the physical system, i.e., its solutions transform among themselves under the action of the group of symmetries.

It is a stunning reality and a fact of profound significance that many of the differential equations of mathematical physics are, as in Equation (27.62), expressions of the invariance of the Casimir operators of some Lie algebra in a particular representation. Moreover, all the standard functions of mathematical physics, such as Bessel, hypergeometric, and confluent hypergeometric functions, are related to matrix elements in the representations of a few of the simplest Lie groups (see [Mill 68] for a thorough discussion of this topic).

---

**Claude Chevalley** (1909–1984) was the only son of Abel and Marguerite Chevalley who were the authors of the Oxford Concise French Dictionary. He studied under Emile Picard

at the Ecole Normale Supérieure in Paris, graduating in 1929 and becoming the youngest of the mathematicians of the Bourbaki school.

After graduation, Chevalley went to Germany to continue his studies under Artin at Hamburg during the session 1931–32. He then went to the University of Marburg to work with Hasse, who had been appointed to fill Hensel's chair there in 1930. He was awarded his doctorate in 1937. A year later Chevalley went to the Institute for Advanced Study at Princeton, where he also served on the faculty of Princeton University. From July 1949 until June 1957 he served as professor of mathematics at Columbia University, afterwards returning to the University of Paris.



Chevalley had a major influence on the development of several areas of mathematics. His papers of 1936 and 1941 led to major advances in class field theory and also in algebraic geometry. He did pioneering work in the theory of local rings in 1943, developing the ideas of Krull into a theorem bearing his name. Chevalley's theorem was important in applications made in 1954 to quasi-algebraically closed fields and the following year to algebraic groups. Chevalley groups play a central role in the classification of finite simple groups. His name is also attached to Chevalley decompositions and to a Chevalley type of semisimple algebraic group.

Many of his texts have become classics. He wrote *Theory of Lie Groups* in three volumes which appeared in 1946, 1951, and 1955. He also published *Theory of Distributions* (1951), *Introduction to the Theory of Algebraic Functions of one Variable* (1951), *The Algebraic Theory of Spinors* (1954), *Class Field Theory* (1954), *Fundamental Concepts of Algebra* (1956), and *Foundations of Algebraic Geometry* (1958).

Chevalley was awarded many honors for his work. Among these was the Cole Prize of the American Mathematical Society. He was elected a member of the London Mathematical Society in 1967.

### 27.5.3 Representation of $\mathfrak{so}(3)$ and $\mathfrak{so}(3, 1)$

Because of their importance in physical applications, we study the representations of  $\mathfrak{so}(3)$ , the rotation, and  $\mathfrak{so}(3, 1)$ , the Lorentz, algebras. For rotations, we define  $\mathbf{J}_1 \equiv -i\mathbf{M}_{23}$ ,  $\mathbf{J}_2 \equiv i\mathbf{M}_{13}$ , and  $\mathbf{J}_3 \equiv -i\mathbf{M}_{12}$ ,<sup>13</sup> and note that the  $\mathbf{J}_i$ 's satisfy exactly the same commutation relations as the angular momentum operators of Chapter 12. Therefore, the irreducible representations of  $\mathfrak{so}(3)$  are labeled by  $j$ , which can be an integer or a half-odd integer (see also Example 27.5.6). These representations are finite-dimensional because  $SO(3)$  is a compact group (Example 27.3.3 and Theorem 27.3.9). The dimension of the irreducible representation of  $\mathfrak{so}(3)$  labeled by  $j$  is  $2j + 1$ .

Because of local isomorphism of Lie groups and their Lie algebras, the same irreducible spaces found for Lie algebras can be used to represent the Lie groups.

<sup>13</sup>Sometimes we use  $\mathbf{J}_x$ ,  $\mathbf{J}_y$ , and  $\mathbf{J}_z$  instead of  $\mathbf{J}_1$ ,  $\mathbf{J}_2$ , and  $\mathbf{J}_3$ .

In particular, the states  $\{|jm\rangle\}_{m=-j}^j$ , where  $m$  is the eigenvalue of  $J_z$ , can also be used as a basis of the  $j$ -th irreducible representation.

The flow of each infinitesimal generator of  $\mathfrak{so}(3)$  is a one-parameter subgroup of  $SO(3)$ . For example,  $\exp(\mathbf{M}_{12}\varphi)$  is a rotation of angle  $\varphi$  about the  $z$ -axis. Using Euler angles, we can write a general rotation as

$$\mathbf{R}(\psi, \theta, \varphi) = \exp(\mathbf{M}_{12}\psi) \exp(\mathbf{M}_{31}\theta) \exp(\mathbf{M}_{12}\varphi).$$

The corresponding rotation operator acting on the vectors of the carrier space is

$$\mathbf{R}(\psi, \theta, \varphi) = \exp(\mathbf{M}_{12}\psi) \exp(\mathbf{M}_{31}\theta) \exp(\mathbf{M}_{12}\varphi) = e^{iJ_z\psi} e^{iJ_y\theta} e^{iJ_z\varphi}.$$

rotation matrix

The rotation matrix corresponding to the above operator is obtained by sandwiching  $\mathbf{R}(\psi, \theta, \varphi)$  between basis vectors of a given irreducible representation:

$$\begin{aligned} D_{m'm}^{(j)}(\psi, \theta, \varphi) &\equiv \langle jm' | \mathbf{R}(\psi, \theta, \varphi) | jm \rangle = \langle jm' | e^{iJ_z\psi} e^{iJ_y\theta} e^{iJ_z\varphi} | jm \rangle \\ &= e^{im'\psi} e^{im\varphi} \langle jm' | e^{iJ_y\theta} | jm \rangle = e^{i(m'\psi+m\varphi)} d_{m'm}^{(j)}(\theta). \end{aligned} \tag{27.63}$$

Wigner formula for rotation matrices

Thus, the calculation of rotation matrices is reduced to finding  $d_{m'm}^{(j)}(\theta)$ . These are given by the **Wigner formula** (see [Hame 89, pp. 348–357]):

$$d_{m'm}^{(j)}(\theta) = \sum_{\mu} \phi(j, m, m'; \mu) \left(\cos \frac{\theta}{2}\right)^{2(j-\mu)+m-m'} \left(\sin \frac{\theta}{2}\right)^{2\mu+m'-m} \tag{27.64}$$

where

$$\phi(j, m, m'; \mu) \equiv (-1)^{\mu} \frac{[(j+m)!(j-m)!(j+m')!(j-m')!]^{1/2}}{(j+m-\mu)!\mu!(j-m'-\mu)!(m'-m+\mu)!}$$

and the summation extends over all integral values of  $\mu$  for which the factorials have a meaning. The number of terms in the summation is equal to  $1 + \tau$ , where  $\tau$  is the smallest of the four integers  $j \pm m, j \pm m'$ .

From the rotation matrices, we can obtain the characters of the rotation group. However, an easier way is to use Euler's theorem (Theorem 4.7.7), Example 23.2.19, and Box 24.2.7 to conclude that the character of a rotation depends only on the angle of rotation, and not on the direction of the rotation axis. Choosing the  $z$ -axis as our only axis of rotation, we obtain

$$\begin{aligned} \chi^{(j)}(\varphi) &= \sum_{m=-j}^j \langle jm | e^{iJ_z\varphi} | jm \rangle = \sum_{m=-j}^j e^{im\varphi} = e^{-ij\varphi} \sum_{m=-j}^j e^{i(j+m)\varphi} \\ &= e^{-ij\varphi} \sum_{k=0}^{2j} e^{ik\varphi} = e^{-ij\varphi} \frac{e^{i(2j+1)\varphi} - 1}{e^{i\varphi} - 1} \\ &= \frac{e^{i(j+1)\varphi} - e^{-ij\varphi}}{e^{i\varphi} - 1} = \frac{\sin(j + \frac{1}{2})\varphi}{\sin(\varphi/2)}. \end{aligned} \tag{27.65}$$

Equation (27.65) can be used to obtain the celebrated **addition theorem** for angular momenta. Suppose that initially we have two physical systems corresponding to angular momenta  $j_1$  and  $j_2$ . When these systems are made to interact with one another, the total system will be described by the tensor product states. These states are vectors in the tensor product of the irreducible representations  $T^{(j_1)}$  and  $T^{(j_2)}$  of the rotation group, as discussed in Section 24.7. This product is reducible. To find the factors into which it reduces, we consider its character corresponding to angle  $\varphi$ . Using Equation (24.44), we have

$$\begin{aligned}\chi^{(j_1 \times j_2)}(\varphi) &= \chi^{(j_1)}(\varphi) \cdot \chi^{(j_2)}(\varphi) = \sum_{m_1=-j_1}^{j_1} e^{im_1\varphi} \sum_{m_2=-j_2}^{j_2} e^{im_2\varphi} \\ &= \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} e^{i(m_1+m_2)\varphi} \\ &= \sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^J e^{iM\varphi} = \sum_{J=|j_1-j_2|}^{j_1+j_2} \chi^{(J)}(\varphi),\end{aligned}$$

where the double sum on the third line is an equivalent way of writing the double summation of the second line, as the reader may verify. From this equation we read off the Clebsch–Gordan decomposition of the tensor product:

addition theorem for  
angular momenta

$$T^{(j_1)} \otimes T^{(j_2)} = \sum_{J=|j_1-j_2|}^{j_1+j_2} T^{(J)}, \quad (27.66)$$

which is also known as the addition theorem for angular momenta. Equation (27.66) shows that (see page 702)

**27.5.7. Box.** *The rotation group is simply reducible.*

The RHS of Equation (27.66) tells us which irreducible representations result from multiplying  $T^{(j_1)}$  and  $T^{(j_2)}$ . In particular, if  $j_1 = j_2 \equiv l$ , the RHS includes the  $J = 0$  representation, i.e., a scalar. In terms of the states, this says that we can combine two states with angular momentum  $l$  to obtain a scalar state. Let us find this combination. We use Equation (24.48) in the form

$$|JM\rangle = \sum_{m_1, m_2} C(j_1 j_2; J | m_1 m_2; M) |j_1, m_1; j_2, m_2\rangle, \quad m_1 + m_2 = M. \quad (27.67)$$

In the case under investigation,  $J = 0 = M$ , so (27.67) becomes

$$|00\rangle = \sum_{m=-l}^l C(l l; 0 | m, -m; 0) |lm; l, -m\rangle.$$

Problem 27.33 shows that  $C(l; 0|m, -m; 0) = (-1)^{l-m} / \sqrt{2l+1}$ , so that

$$|00\rangle = \sum_{m=-l}^l \frac{(-1)^{l-m}}{\sqrt{2l+1}} |lm; l, -m\rangle.$$

Take the “inner product” of this with  $\langle \theta, \varphi; \theta', \varphi' |$  to obtain

$$\begin{aligned} \langle \theta, \varphi; \theta', \varphi' | 00 \rangle &= \sum_{m=-l}^l \frac{(-1)^{l-m}}{\sqrt{2l+1}} \langle \theta, \varphi; \theta', \varphi' | lm; l, -m \rangle \\ &= \sum_{m=-l}^l \frac{(-1)^{l-m}}{\sqrt{2l+1}} \underbrace{\langle \theta, \varphi | lm \rangle}_{Y_{lm}(\theta, \varphi)} \underbrace{\langle \theta', \varphi' | l, -m \rangle}_{Y_{l,-m}(\theta', \varphi')}, \end{aligned} \quad (27.68)$$

where we have used  $\langle \theta, \varphi; \theta', \varphi' | = \langle \theta, \varphi | \langle \theta', \varphi' |$  and contracted each bra with a ket. We can evaluate the LHS of (27.68) by noting that since it is a scalar, the choice of orientation of coordinates is immaterial. So, let  $\theta = 0$  to get  $\theta' = \gamma$ , the angle between the two directions. Then using the facts

$$Y_{lm}(0, \varphi) = \delta_{m0} \sqrt{\frac{2l+1}{4\pi}} \quad \text{and} \quad Y_{l0}(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos \theta)$$

on the RHS of (27.68), we obtain

$$\langle \theta, \varphi; \theta', \varphi' | 00 \rangle = \frac{(-1)^l}{4\pi} \sqrt{2l+1} P_l(\cos \gamma).$$

Substituting this in the LHS of Equation (27.68), we get

$$P_l(\cos \gamma) = \frac{4\pi}{2l+1} \sum_{m=-l}^l (-1)^m Y_{lm}(\theta, \varphi) Y_{l,-m}(\theta', \varphi'),$$

which is the addition theorem for spherical harmonics discussed in Chapter 12. Let us now turn to  $\mathfrak{so}(3, 1)$ . We collect the generators in two categories

$$\mathbf{M} \equiv (M_1, M_2, M_3) \equiv (\mathbf{M}_{23}, \mathbf{M}_{31}, \mathbf{M}_{12}),$$

$$\mathbf{N} \equiv (N_1, N_2, N_3) \equiv (\mathbf{M}_{01}, \mathbf{M}_{02}, \mathbf{M}_{03}),$$

and verify that

$$[M_i, M_j] = -\epsilon_{ijk} M_k, \quad [N_i, N_j] = \epsilon_{ijk} M_k, \quad [M_i, N_j] = -\epsilon_{ijk} N_k,$$

and that there are two Casimir operators:  $\mathbf{M}^2 - \mathbf{N}^2$  and  $\mathbf{M} \cdot \mathbf{N}$ . It follows that the irreducible representations of  $\mathfrak{so}(3, 1)$  are labeled by two numbers. To find these numbers, define the generators

$$\mathbf{J} \equiv \frac{1}{2i} (\mathbf{M} + i\mathbf{N}), \quad \mathbf{K} \equiv \frac{1}{2i} (\mathbf{M} - i\mathbf{N}),$$

and show that

$$[J_i, J_m] = \epsilon_{imk} J_k, \quad [K_i, K_j] = \epsilon_{ijm} K_m, \quad [J_i, K_j] = 0.$$

It follows that the  $J$ 's and the  $K$ 's generate two completely independent Lie algebras isomorphic to the angular momentum algebras and that  $\mathfrak{so}(3, 1)$  is a direct sum of these algebras. Since each one requires a (half-odd) integer to designate its irreducible representations, we can choose these two numbers as the eigenvalues of the Casimir operators needed to label the irreducible representations of  $\mathfrak{so}(3, 1)$ . Thus, the irreducible representations of  $\mathfrak{so}(3, 1)$  are of the form  $T^{(jj')}$ , where  $j$  and  $j'$  can each be an integer or a half-odd integer.

### 27.5.4 Representation of the Poincaré Algebra

The Poincaré algebra  $\mathfrak{p}(p, n - p)$ , introduced in Section 27.2.1, is the generalization of the Lie algebra of the invariance group of the special theory of relativity. It contains the Lorentz, the rotation, and the translation groups as its proper subgroups. Its irreducible representations are of direct physical significance, and we shall study them here.

As the first step in the construction of representations of  $\mathfrak{p}(p, n - p)$ , we shall try to find its Casimir operators. Equation (27.61) suggests one, but it works only for semisimple Lie algebras, and the Poincaré algebra is not semisimple. Nevertheless, let us try to find an operator based on that construction. From the commutation relations for  $\mathfrak{p}(p, n - p)$ , as given in Equation (27.40), and the double-indexed structure constants defined by,<sup>14</sup>

$$[M_{ij}, M_{kl}] = c_{ij,kl}^{mn} M_{mn}, \quad [M_{ij}, P_k] = c_{ij,k}^m P_m,$$

we obtain

$$\begin{aligned} c_{ij,kl}^{mn} &= \delta_j^m \delta_i^n \eta_{ik} - \delta_j^m \delta_k^n \eta_{il} + \delta_i^m \delta_k^n \eta_{jl} - \delta_i^m \delta_j^n \eta_{jk}, \\ c_{ij,k}^m &= \delta_j^m \eta_{ik} - \delta_i^m \eta_{jk}. \end{aligned} \tag{27.69}$$

From these structure constants, we can construct a double indexed "metric"

$$g_{ij,kl} = c_{ij,mn}^{rs} c_{kl,rs}^{mn} + c_{ij,m}^r c_{kl,r}^m,$$

which the reader may verify to be equal to

$$g_{ij,kl} = 2(n - 1)(\eta_{jk}\eta_{il} - \eta_{ik}\eta_{jl}).$$

<sup>14</sup>Please make sure to differentiate between the pair  $(M_{ij}, P_k)$  (which acts on  $\mathfrak{p}$ ) and the pair  $(\mathbf{M}_{ij}, \mathbf{P}_k)$ , which acts on the state vectors in the Hilbert space of representation.

There is no natural way of constructing a single-indexed metric. Therefore, we can only contract the  $M$ 's. In doing so, it is understood that the indices are raised and lowered by  $\eta_{ij}$ . So, the first candidate for a Casimir operator is

$$M^2 \equiv g_{ij,kl} M^{ij} M^{kl} = 2(n-1)(\eta_{jk}\eta_{il} - \eta_{ik}\eta_{jl}) M^{ij} M^{kl} = -4(n-1) M^{ij} M_{ij}$$

The reader may verify that  $M^2$  commutes with all the  $M^{ij}$ 's but not with the  $P^i$ 's. This is to be expected because  $M^2$ , the total "angular momentum" operator<sup>15</sup> is a scalar and should commute with all its components. But commutation with the  $P^i$ 's is not guaranteed.

The construction above, although a failure, gives us a clue for a successful construction. We can make another scalar out of the  $P$ 's. The reader may check that  $P^2 \equiv \eta^{ij} P_i P_j$  indeed commutes with all elements of the Poincaré algebra. We have thus found one Casimir operator. Can we find more? We have exhausted the polynomials of degree two. The only third-degree polynomials that we can construct are  $M^{ij} P_i P_j$  and  $\eta_{il} M^{ij} M_{jk} M^{kl}$ . The first one is identically zero (why?), and the second one will not commute with the  $P$ 's.

To find higher-order polynomials in the infinitesimal generators, we build new tensors out of them and contract these tensors with one another. For example, consider the vector

$$C_i \equiv M_{ij} P^j = \eta^{kj} M_{ij} P_k. \tag{27.70}$$

Then  $C^i C_i$ , a fourth-degree polynomial in the generators, is a scalar, and therefore, it commutes with the  $M_{ij}$ 's, but unfortunately, not with  $P_i$ 's.

Another common way to construct tensors is to contract various numbers of the generators with the Levi-Civita tensor. For example,

$$W^{i_1 \dots i_{n-3}} \equiv \epsilon^{i_1 \dots i_{n-3} jkl} M_{jk} P_l \tag{27.71}$$

is a contravariant tensor of rank  $n - 3$ . Let us contract  $W$  with itself to find a scalar (which we expect to commute with all the  $M_{ij}$ 's):

$$\begin{aligned} W^2 &\equiv W^{i_1 \dots i_{n-3}} W_{i_1 \dots i_{n-3}} \\ &= \epsilon^{i_1 \dots i_{n-3} jkl} M_{jk} P_l \epsilon_{i_1 \dots i_{n-3} rst} M^{rs} P^t \\ &= (-1)^{n-} \sum_{\pi} \epsilon_{\pi} \delta_{\pi(i_1)}^{i_1} \delta_{\pi(i_2)}^{i_2} \dots \delta_{\pi(i_{n-3})}^{i_{n-3}} \delta_{\pi(r)}^j \delta_{\pi(s)}^k \delta_{\pi(t)}^l M_{jk} P_l M^{rs} P^t \\ &= (-1)^p (n-3)! \sum_{\pi} \epsilon_{\pi} \delta_{\pi(r)}^j \delta_{\pi(s)}^k \delta_{\pi(t)}^l M_{jk} P_l M^{rs} P^t, \end{aligned}$$

where we used Equation (25.21). The sum above can be carried out, with the final result

$$\begin{aligned} W^2 &= 2(-1)^p (n-3)! (M_{ij} M^{ij} P^2 - 2C_i C^i) \\ &= 2(-1)^p (n-3)! (M^2 P^2 - 2C^2), \end{aligned} \tag{27.72}$$

<sup>15</sup>This "angular momentum" includes ordinary rotations as well as the Lorentz boosts.

where  $C_i$  was defined in Equation (27.70). We have already seen that  $M^2$ ,  $P^2$ , and  $C^2$  all commute with the  $M_{jk}$ 's. The reader may check that  $W^2$  commutes with the  $P_j$ 's as well. In fact,  $W^{1 \dots n-3}$  itself commutes with all the  $P_j$ 's. Other tensors and Casimir operators can be constructed in a similar fashion.

We now want to construct the irreducible vector spaces that are labeled by the eigenvalues of the Casimir operators. We take advantage of the fact that the Poincaré algebra has a commutative subalgebra, the translation generators. Since the  $P_k$ 's commute among themselves and with  $P^2$  and  $W^2$ , we can choose simultaneous eigenvectors of  $\{P_k\}_{k=1}^n$ ,  $P^2$ , and  $W^2$ . In particular, we can label the vectors of an irreducible invariant subspace by the eigenvalues of these operators. The  $P^2$  and  $W^2$  labels will be the same for all vectors in each irreducible invariant subspace, while the  $P_k$ 's will label different vectors of the same invariant subspace.

Let us concentrate on the momentum labels and let  $|\psi_{\mathbf{p}}^\mu\rangle$  be a vector in an irreducible representation of  $\mathfrak{p}(p, n - p)$ , where  $\mathbf{p}$  labels momenta and  $\mu$  distinguishes among all different vectors that have the same momentum label. We thus have

$$P_k |\psi_{\mathbf{p}}^\mu\rangle = p_k |\psi_{\mathbf{p}}^\mu\rangle \quad \text{for } k = 1, 2, \dots, n, \tag{27.73}$$

where  $p_k$  is the eigenvalue of  $P_k$ . We also need to know how the "rotation" operators act on  $|\psi_{\mathbf{p}}^\mu\rangle$ . Instead of the full operator  $e^{M_{ij}\theta^{ij}}$ , we apply its small-angle approximation  $1 + M_{ij}\theta^{ij}$ . Since all states are labeled by momentum, we expect the rotated state to have a new momentum label, i.e., to be an eigenstate of  $P_k$ . We want to show that  $(1 + M_{ij}\theta^{ij})|\psi_{\mathbf{p}}^\mu\rangle$  is an eigenvector of  $P_k$ . Let the eigenvalue be  $\mathbf{p}'$ , which should be slightly different from  $\mathbf{p}$ . Then, the problem reduces to determining  $\delta\mathbf{p}' \equiv \mathbf{p}' - \mathbf{p}$ . Ignoring the index  $\mu$  for a moment, we have

$$P_k |\psi_{\mathbf{p}'}\rangle = p'_k |\psi_{\mathbf{p}'}\rangle = (p_k + \delta p_k) (1 + M_{ij}\theta^{ij}) |\psi_{\mathbf{p}}\rangle.$$

Using the commutation relations between  $P_k$  and  $M_{ij}$ , we can write the LHS as

$$LHS = P_k (1 + M_{ij}\theta^{ij}) |\psi_{\mathbf{p}}\rangle = [p_k + \theta^{ij} (M_{ij} p_k + \eta_{jk} p_i - \eta_{ik} p_j)] |\psi_{\mathbf{p}}\rangle.$$

The RHS, to first order in infinitesimal quantities, can be expressed as

$$RHS = (p_k + \delta p_k + p_k \theta^{ij} M_{ij}) |\psi_{\mathbf{p}}\rangle.$$

Comparison of the last two equations shows that

$$\delta p_k = \theta^{ij} (\eta_{jk} p_i - \eta_{ik} p_j) = \theta^{ij} (\eta_{jk} \eta_{il} - \eta_{ik} \eta_{jl}) p^l = \theta^{ij} (M_{ij})_{kl} p^l,$$

where we used Equation (27.37). It follows that

$$\mathbf{p}' = \mathbf{p} + \delta\mathbf{p} = (1 + \theta^{ij} M_{ij})\mathbf{p},$$

stating that the rotation operator of the carrier Hilbert space rotates the momentum label of the state. Note that since "rotations" do not change the length (induced by  $\eta$ ),  $\mathbf{p}'$  and  $\mathbf{p}$  have the same length.



construction and properties of the little group

To obtain all the vectors of an irreducible representation of  $\mathfrak{p}(p, n - p)$ , we must apply the rotation operators to vectors such as  $|\psi_{\mathfrak{p}}^{\mu}\rangle$ . But not all rotations will change the label  $\mathfrak{p}$ ; for example, in three dimensions, the vector  $\mathfrak{p}$  will not be affected<sup>16</sup> by a rotation about  $\mathfrak{p}$ . This motivates the following definition.

little group and little algebra

**27.5.8. Definition.** Let  $\mathfrak{p}_0$  be any given eigenvalue of the translation generators. The set  $\mathcal{R}_{\mathfrak{p}_0}$  of all rotations  $A^{\mathfrak{p}_0}$  that do not change  $\mathfrak{p}_0$ , is a subgroup of the rotation group  $O(p, n - p)$ , called the **little group** corresponding to  $\mathfrak{p}_0$ . The **little algebra** consists of the generators  $M_{ij}^{\mathfrak{p}_0}$  satisfying

$$M_{ij}^{\mathfrak{p}_0} \mathfrak{p}_0 = 0.$$

induced representations

The significance of the little group resides in the fact that a representation of  $\mathcal{R}_{\mathfrak{p}_0}$  induces a representation of the whole Poincaré group. We shall only sketch the proof in the following and refer the reader to Mackey [Mack 68] for a full and rigorous discussion of **induced representations**.

Suppose we have found an irreducible representation of  $\mathcal{R}_{\mathfrak{p}_0}$  with operators  $A^{\mathfrak{p}_0}$ . Let  $A^{\mathfrak{p}\mathfrak{p}_0}$  be the rotation that carries<sup>17</sup>  $\mathfrak{p}_0$  to  $\mathfrak{p}$ , i.e.,  $\mathfrak{p} = A^{\mathfrak{p}\mathfrak{p}_0} \mathfrak{p}_0$ . Consider any rotation  $A$  and let  $\mathfrak{p}'$  be the momentum obtained when  $A$  acts on  $\mathfrak{p}$ , i.e.,  $A\mathfrak{p} \equiv \mathfrak{p}'$ . Then

$$\underbrace{A A^{\mathfrak{p}\mathfrak{p}_0}}_{=\mathfrak{p}} \mathfrak{p}_0 = \underbrace{A^{\mathfrak{p}'\mathfrak{p}_0}}_{=\mathfrak{p}'} \mathfrak{p}_0 \Rightarrow (A^{\mathfrak{p}'\mathfrak{p}_0})^{-1} A A^{\mathfrak{p}\mathfrak{p}_0} \mathfrak{p}_0 = \mathfrak{p}_0.$$

This shows that  $(A^{\mathfrak{p}'\mathfrak{p}_0})^{-1} A A^{\mathfrak{p}\mathfrak{p}_0}$  belongs to the little group. So,

$$(A^{\mathfrak{p}'\mathfrak{p}_0})^{-1} A A^{\mathfrak{p}\mathfrak{p}_0} = A^{\mathfrak{p}_0}$$

for some  $A^{\mathfrak{p}_0} \in \mathcal{R}_{\mathfrak{p}_0}$ . Thus,  $A = A^{\mathfrak{p}'\mathfrak{p}_0} A^{\mathfrak{p}_0} (A^{\mathfrak{p}\mathfrak{p}_0})^{-1}$ , and

$$\begin{aligned} T(A) |\psi_{\mathfrak{p}}^{\mu}\rangle &\equiv A |\psi_{\mathfrak{p}}^{\mu}\rangle = A^{\mathfrak{p}'\mathfrak{p}_0} A^{\mathfrak{p}_0} (A^{\mathfrak{p}\mathfrak{p}_0})^{-1} |\psi_{\mathfrak{p}}^{\mu}\rangle = A^{\mathfrak{p}'\mathfrak{p}_0} A^{\mathfrak{p}_0} |\psi_{\mathfrak{p}_0}^{\mu}\rangle \\ &= A^{\mathfrak{p}'\mathfrak{p}_0} \sum_{\nu} T_{\nu\mu}(A^{\mathfrak{p}_0}) |\psi_{\mathfrak{p}_0}^{\nu}\rangle = \sum_{\nu} T_{\nu\mu}(A^{\mathfrak{p}_0}) A^{\mathfrak{p}'\mathfrak{p}_0} |\psi_{\mathfrak{p}_0}^{\nu}\rangle \\ &= \sum_{\nu} T_{\nu\mu}(A^{\mathfrak{p}_0}) |\psi_{\mathfrak{p}'}^{\nu}\rangle = \sum_{\nu} T_{\nu\mu}(A^{\mathfrak{p}_0}) |\psi_{A\mathfrak{p}}^{\nu}\rangle. \end{aligned}$$

Note how the matrix elements of the representation of the little group alone have entered in the last line. We therefore consider

$$T(A) |\psi_{\mathfrak{p}}^{\mu}\rangle \equiv \sum_{\nu} R_{\nu\mu}(A^{\mathfrak{p}_0}) |\psi_{A\mathfrak{p}}^{\nu}\rangle, \quad \text{where} \quad \begin{cases} A = A^{\mathfrak{p}'\mathfrak{p}_0} A^{\mathfrak{p}_0} (A^{\mathfrak{p}\mathfrak{p}_0})^{-1}, \\ \mathfrak{p}' \equiv A\mathfrak{p}. \end{cases} \quad (27.74)$$

<sup>16</sup>The reader should be warned that although such a rotation does not change  $\mathfrak{p}$ , the rotation operator may change the state  $|\psi_{\mathfrak{p}}^{\mu}\rangle$ . However, the resulting state will be an eigenstate of the  $\mathbf{P}_k$ 's with eigenvalue  $\mathfrak{p}$ .  
<sup>17</sup>We are using the fact that  $O(p, n - p)$  is transitive (see Problem 27.39).

To avoid confusion, we have used  $R$  for the representation of the little group. We claim that Equation (27.74) defines a (matrix) representation of the whole group. In fact,

$$\begin{aligned} T(A_1)T(A_2)|\psi_{\mathfrak{p}}^{\mu}\rangle &= T(A_1)\sum_{\nu}R_{\nu\mu}(A_2^{P_0})|\psi_{A_2\mathfrak{p}}^{\nu}\rangle \\ &= \sum_{\nu}R_{\nu\mu}(A_2^{P_0})\sum_{\rho}R_{\rho\nu}(A_1^{P_0})|\psi_{A_1A_2\mathfrak{p}}^{\rho}\rangle \\ &= \sum_{\rho}\left(\underbrace{\sum_{\nu}R_{\rho\nu}(A_1^{P_0})R_{\nu\mu}(A_2^{P_0})}_{=R_{\rho\mu}(A_1^{P_0}A_2^{P_0})\text{ since }R\text{ is a rep.}}\right)|\psi_{A_1A_2\mathfrak{p}}^{\rho}\rangle. \end{aligned}$$

The reader may check that  $A_1^{P_0}A_2^{P_0} \equiv (A_1A_2)^{P_0}$ . Therefore,

$$T(A_1)T(A_2)|\psi_{\mathfrak{p}}^{\mu}\rangle = \sum_{\rho}R_{\rho\mu}((A_1A_2)^{P_0})|\psi_{A_1A_2\mathfrak{p}}^{\rho}\rangle \equiv T(A_1A_2)|\psi_{\mathfrak{p}}^{\mu}\rangle,$$

and  $T$  is indeed a representation. It turns out that if  $R$  is irreducible, then so is  $T$ . The discussion above shows that the irreducible representations of the Poincaré group are entirely determined by those of the little group and Equation (27.73). The recipe for the construction of the irreducible representations of  $\mathfrak{p}(p, n - p)$  is now clear:

**27.5.9. Theorem.** *Choose any simultaneous eigenvector  $\mathfrak{p}_0$  of the  $\mathfrak{P}_k$ 's. Find the little algebra  $\mathfrak{A}_{\mathfrak{p}_0}$  at  $\mathfrak{p}_0$  by finding all  $M_{ij}$ 's satisfying  $M_{ij}\mathfrak{p}_0 = 0$ . Find all irreducible representations of  $\mathfrak{A}_{\mathfrak{p}_0}$ . The same eigenvalues that label the irreducible representations of  $\mathfrak{A}_{\mathfrak{p}_0}$  can be used, in addition to those of  $\mathfrak{P}^2$  and  $\mathfrak{W}^2$ , to label the irreducible representations of  $\mathfrak{p}(p, n - p)$ .*

We are particularly interested in  $\mathfrak{p}(3, 1)$ , the symmetry group of the special theory of relativity. In applying the formalism developed above, we need to make contact with the physical world. This always involves interpretations. Borrowing from the angular momentum theory, in which a physical system was given the attribute of angular momentum, the label of the irreducible representation of the rotation group, we attribute the labels of an irreducible representation of the Poincaré group, i.e., the eigenvalues of the four translation generators and the two Casimir operators, to a physical system. Since the four translation generators are identified as the three components of momentum and energy, and their specification implies their constancy over time, we have to come to the conclusion that

**27.5.10. Box.** *An irreducible representation of the Poincaré group specifies a free relativistic particle.*

There may be some internal interactions between constituents of a (composite) particle, e.g. between quarks inside a proton, but as a whole, the composite will be interpreted as a single particle. To construct the little group, we have to specify a 4-momentum  $\mathbf{p}_0$ . We shall consider two cases: In the first case,  $\mathbf{p}_0 \cdot \mathbf{p}_0 \neq 0$ , whereby the particle is deduced to be massive and we can choose<sup>18</sup>  $\mathbf{p}_0 = (0, 0, 0, m)$ . In the second case,  $\mathbf{p}_0 \cdot \mathbf{p}_0 = 0$ , in which case the particle is massless, and we can choose  $\mathbf{p}_0 = (p, 0, 0, p)$ . We consider these two cases separately.

The little group (really, the little Lie algebra) for  $\mathbf{p}_0 = (0, 0, 0, m)$  is obtained by searching for those rotations that leave  $\mathbf{p}_0$  fixed. This is equivalent to searching for  $M_{ij}$ 's that annihilate  $(0, 0, 0, m)$ , namely, the solutions to

$$(M_{ij}\mathbf{p}_0)_l = (M_{ij})_{lr}(\mathbf{p}_0)^r = (M_{ij})_{l0}m = 0 \Rightarrow (M_{ij})_{l0} = 0.$$

Since  $(M_{ij})_{l0} = \eta_{i0}\eta_{jl} - \eta_{j0}\eta_{il}$ , we conclude that  $(M_{ij})_{l0} = 0$  if and only if  $i \neq 0$  and  $j \neq 0$ . Thus the little group is generated by  $(M_{23}, M_{31}, M_{12})$  which are the components of angular momentum. The reader may also verify directly that when the 4-momentum has only a time component, the Casimir operator  $W^2$  reduces essentially to the total angular momentum operator. Since we are dealing with a single particle, the total angular momentum can only be spin. Therefore, we have the following theorem.

**27.5.11. Theorem.** *In the absence of any interactions, a massive relativistic particle is specified by its mass  $m$  and its spin  $s$ , the former being any positive number, the latter taking on integer or half-odd-integer values.*

The case of the massless particle can be handled in the same way. We seek those  $M_{ij}$ 's that annihilate  $(p, 0, 0, p)$ , namely, the solutions to

$$(M_{ij}\mathbf{p}_0)_k = (M_{ij})_{kr}(\mathbf{p}_0)^r = (M_{ij})_{k0}p + (M_{ij})_{k1}p = 0.$$

The reader may check that

$$\begin{aligned} (M_{01}\mathbf{p}_0)_k &= \eta_{1k}p - \eta_{0k}p, & (M_{02}\mathbf{p}_0)_k &= \eta_{2k}p, & (M_{03}\mathbf{p}_0)_k &= \eta_{3k}p, \\ (M_{23}\mathbf{p}_0)_k &= 0, & (M_{12}\mathbf{p}_0)_k &= \eta_{2k}p, & (M_{13}\mathbf{p}_0)_k &= \eta_{3k}p. \end{aligned}$$

Clearly,  $M_{23}$  is one of the generators of the little group. Subtracting the middle terms and the last terms of each line, we see that  $M_{02} - M_{12}$  and  $M_{03} - M_{13}$  are the other two generators. These happen to be the components of  $W$ . In fact, it is easily verified that

$$W^0 = W^1 = M_{23}p, \quad W^2 = 2p(M_{13} - M_{03}), \quad W^3 = 2p(M_{02} - M_{12}). \quad (27.75)$$

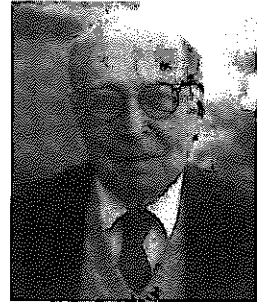
Therefore, the little group is generated by all the components of  $W$ . Furthermore,  $W^2$  has zero eigenvalue for  $|\psi_{\mathbf{p}_0}\rangle$  when  $\mathbf{p}_0 = (p, 0, 0, p)$ . Since both Casimir

<sup>18</sup>We use units in which  $c = 1$ .

operators annihilate the state  $|\psi_{p_0}\rangle$ , we need to come up with another way of labeling the states.

---

**Eugene Paul Wigner** (1902–1995) was the second of three children born to Hungarian Jewish parents in Budapest. His father operated a large leather tannery and hoped that his son would follow him in that vocation, but the younger Wigner soon discovered both a taste and an aptitude for mathematics and physics. Although Wigner tried hard to accommodate his father's wishes, he clearly heard his calling, and the world of physics is fortunate that he did.



Wigner began his education in what he said “may have been the finest high school in the world.” He later studied chemical engineering and returned to Budapest to apply that training in his father's tannery. He kept track of the seminal papers during the early years of quantum theory and, when the lure of physics became too strong, returned to Berlin to work in a crystallography lab. He lectured briefly at the University of Göttingen before moving to America to escape the Nazis.

Wigner accepted a visiting professorship to Princeton in 1930. When the appointment was not made permanent, the disappointed young professor moved to the University of Wisconsin, where he served happily until his new wife died suddenly of cancer only a few months after their marriage. As Wigner prepared, quite understandably, to leave Wisconsin, Princeton corrected its earlier mistake and offered him a permanent position. Except for occasional visiting appointments in America and abroad, he remained at Princeton until his death.

Wigner's contributions to mathematical physics began during his studies in Berlin, where his supervisor suggested a problem dealing with the symmetry of atoms in a crystal. John von Neumann, a fellow Hungarian physicist, pointed out the relevance of papers by Frobenius and Schur on representation theory. Wigner soon became enamored with the group theory inherent in the problem and began to apply that approach to quantum mechanical problems. Largely at the urging of Leo Szilard (another Hungarian physicist and Wigner's best friend), Wigner collected many of his results into the classic textbook *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*.

The decades that followed were filled with important contributions to mathematical physics, with applications of group theory comprising a large share: angular momentum; nuclear physics and  $SU(4)$  or “supermultiplet” theory; parity; and studies of the Lorentz group and Wigner's classic definition of an elementary particle. Other work included early efforts in many-body theory and a paper on level spacings derived from the properties of Hermitian matrices that later proved useful to workers in quantum chaos.

As with most famous figures, Wigner's personality became as well known as his professional accomplishments. His insistence on “reasonable” behavior, for instance, made him refuse to pay a relative's hospital bill until after the patient was released—it was obviously unreasonable to hold a sick person hostage. His gentleness is exemplified in an anecdote in which on getting into an argument about a tip with a New York City cab driver, Wigner loses his patience, stamps his foot, and says, “Oh, go to hell, . . . please!”

He held others' feelings in such high regard that it was said to be impossible to follow Wigner through a door. He was light-hearted and fun-loving, but also devoted to his family and concerned about the future of the planet. This combination of exceptional skill and laudable humanity ensures Wigner's place among the most highly regarded of his field. (Taken from E. Vogt, *Phys. Today* 48 (12) (1995) 40–44.)

Define the new quantities

$$H_{\pm} \equiv \frac{1}{2}(W_1 \pm iW_2), \quad H_0 \equiv \frac{1}{2p}W_0$$

and the corresponding operators acting on the carrier space. From Equation (27.75), it follows that  $[W_1, W_2] = 0$ ,  $W^2 = -4H_+H_-$ , and that

$$[H_+, H_0] = -H_+, \quad [H_+, H_0] = H_-, \quad [H_+, H_-] = 0.$$

Denote the eigenstates of  $W^2$  and  $H_0$  by  $|\alpha, \beta\rangle$ :

$$W^2 |\alpha, \beta\rangle = \alpha |\alpha, \beta\rangle, \quad H_0 |\alpha, \beta\rangle = \beta |\alpha, \beta\rangle.$$

Then the reader may check that  $H_{\pm} |\alpha, \beta\rangle$  has eigenvalues  $\alpha$  and  $\beta \pm 1$ . By applying  $H_{\pm}$  repeatedly, we can generate all eigenvalues of  $H_0$  and note that they are of the form

$$\beta = r + n, \quad \text{where } n = 0, \pm 1, \pm 2, \dots \quad \text{and} \quad 1 > r \geq 0.$$

Since  $H_0 = M_{23}$ ,  $H_0$  is recognized as an angular momentum operator whose eigenvalues are integer (for bosons) and half-odd integer (for fermions). Therefore,  $r = 0$  for bosons and  $r = \frac{1}{2}$  for fermions.

Now, within an irreducible representation, only those  $|\alpha, \beta\rangle$ 's can occur that have the same  $\alpha$ . Therefore, if we relabel the  $\beta$  values by integers, then

$$\langle \alpha, n | H_0 | \alpha, m \rangle = (r + n)\delta_{nm}.$$

Similarly,

$$\begin{aligned} \langle \alpha, n | H_+ | \alpha, m \rangle &= a_n \delta_{n, m+1}, \\ \langle \alpha, n | H_- | \alpha, m \rangle &= b_n \delta_{n, m-1}, \end{aligned}$$

where  $a_n$  and  $b_n$  are some constants. It follows that

$$\begin{aligned} \alpha &= \langle \alpha, n | W^2 | \alpha, n \rangle = \langle \alpha, n | H_+ H_- | \alpha, n \rangle \\ &= \langle \alpha, n | H_+ | \alpha, n-1 \rangle \langle \alpha, n-1 | H_- | \alpha, n \rangle \\ &= a_n b_n. \end{aligned}$$

If we assume that the representation is unitary, then all  $W_j$ 's will be hermitian,  $(H_+)^{\dagger} = H_-$ , so  $a_n = b_n^*$  and  $\alpha = |a_n|^2 \geq 0$ .

helicity of massless particles

If  $\alpha = 0$ , then  $a_n = 0$  and  $b_n = 0$  for all  $n$ . Consequently,  $H_+ = 0 = H_-$ , i.e., there are no raising or lowering operators. It follows that there are only two spin states, corresponding to the maximum and the minimum eigenvalues of  $H_0$ . A natural axis for the projection of spin is the direction of motion of the particle. Then the projection of spin is called **helicity**. We summarize our discussion in the following theorem.

**27.5.12. Theorem.** *In the absence of any interactions, a massless relativistic particle is specified by its spin and its helicity. The former taking on integer or half-odd-integer values  $s$ , the latter having values  $+s$  and  $-s$ .*

Theorems 27.5.11 and 27.5.12 are beautiful examples of the fruitfulness of the interplay between mathematics and physics. Physics has provided mathematics with a group, the Poincaré group, and mathematics, through its theory of group representation, has provided physics with the deep result that all particles must have a spin that takes on a specific value, and none other; that massive particles are allowed to have  $2s + 1$  different values for the projection of their spin; and that massless particles are allowed to have only two values for their spin projection. Such far-reaching results that are both universal and specific makes physics unique among all other sciences. It also provides impetus for the development of mathematics as the only dialect through which nature seems to communicate to us her deepest secrets.

If  $\alpha > 0$ , then the resulting representations will have continuous spin variables. Such representations do not correspond to particles found in nature; therefore, we shall not pursue them any further.

## 27.6 Problems

**27.1.** Show that the set  $G = GL(n, \mathbb{R}) \times \mathbb{R}^n$  equipped with the “product”

$$(A, \mathbf{u})(B, \mathbf{v}) \equiv (AB, A\mathbf{v} + \mathbf{u})$$

forms a group. This is called the **affine group**.

**27.2.** Show that  $m : U \times U \rightarrow \mathbb{R}$  defined in Example 27.1.5 is a local Lie group.

**27.3.** Find the multiplication law for the groups in (b) and (c) of Example 27.1.8.

**27.4.** Show that the one-dimensional projective group of Example 27.1.8 satisfies all the group properties. In particular, find the identity and the inverse of an element in the group.

**27.5.** Let  $G$  be a Lie group. Let  $S$  be a subgroup of  $G$  that is also a submanifold of  $G$ . Show that  $S$  is a Lie group.

**27.6.** Show that the differential map of  $\psi : GL(\mathcal{V}) \rightarrow \mathbb{H}$ , defined by  $\psi(\mathbf{A}) = \mathbf{A}\mathbf{A}^\dagger$ , where  $\mathbb{H}$  is the set of hermitian operators on  $\mathcal{V}$ , is surjective. Derive Equation (27.11).

27.7. Verify that  $I_g \equiv R_g^{-1} \circ L_g$  is an isomorphism.

27.8. Prove Proposition 27.1.21.

27.9. Start with Equation (27.19) and use the fact that second derivative is independent of the order of differentiation to obtain

$$u_{ik} \left[ \frac{\partial \theta_{\kappa\mu}^{-1}}{\partial a_\lambda} - \frac{\partial \theta_{\kappa\lambda}^{-1}}{\partial a_\mu} \right] + \theta_{\kappa\mu}^{-1} \frac{\partial u_{ik}}{\partial a_\lambda} - \theta_{\kappa\lambda}^{-1} \frac{\partial u_{ik}}{\partial a_\mu} = 0.$$

Now use the chain rule  $\partial u_{ik}/\partial a_\lambda = (\partial u_{ik}/\partial x_j)(\partial x_j/\partial a_\lambda)$  and Equation (27.19) to get

$$u_{ik} \left[ \frac{\partial \theta_{\kappa\mu}^{-1}}{\partial a_\lambda} - \frac{\partial \theta_{\kappa\lambda}^{-1}}{\partial a_\mu} \right] + \left[ u_{jv} \frac{\partial u_{ik}}{\partial x_j} - u_{jk} \frac{\partial u_{iv}}{\partial x_j} \right] \theta_{\kappa\mu}^{-1} \theta_{\nu\lambda}^{-1} = 0,$$

or

$$u_{j\sigma} \frac{\partial u_{i\tau}}{\partial x_j} - u_{j\tau} \frac{\partial u_{i\sigma}}{\partial x_j} = c_{\tau\sigma}^\kappa(\mathbf{a}) u_{ik}(\mathbf{x}), \quad (27.76)$$

where

$$c_{\tau\sigma}^\kappa(\mathbf{a}) = \left[ \frac{\partial \theta_{\kappa\mu}^{-1}}{\partial a_\lambda} - \frac{\partial \theta_{\kappa\lambda}^{-1}}{\partial a_\mu} \right] \theta_{\mu\tau} \theta_{\lambda\sigma}.$$

Substituting Equation (27.76) in Equation (27.21) leads to (27.22). Now differentiate both sides of Equation (27.76) with respect to  $a_\rho$  to get

$$\frac{\partial c_{\tau\sigma}^\kappa}{\partial a_\rho} u_{ik} = 0.$$

With the assumption that the  $u_{ik}$  are linearly independent, conclude that the structure “constants” are indeed constants.

27.10. Find the invariant Haar measure of the general linear group in two dimensions.

27.11. Show that the invariant Haar measure for a compact group satisfies  $d\mu_g = d\mu_{g^{-1}}$ . Hint: Define a measure  $\nu$  by  $d\nu_g \equiv d\mu_{g^{-1}}$  and show that  $\nu$  is left-invariant. Now use the uniqueness of the left-invariant Haar measure for compact groups.

27.12. Show that  $O(p, n-p)$  is a group. Use this and the fact that  $\eta^{-1} = \eta$  to show that  $A\eta A^t = \eta$ .

27.13. Show that the orthogonal group  $O(p, n-p)$  has dimension  $n(n-1)/2$ . Hint: Look at its algebra  $\mathfrak{o}(p, n-p)$ .

**27.14.** Let  $x = (x_1, x_2, x_3, x_0)$  be a timelike or (null) 4-vector with  $x_0 > 0$ . Let  $A$  be a proper orthochronous transformation. Show that  $x' = Ax$  is also timelike (null). Hint: Consider the zeroth component of  $x'$  as an inner product of  $(x_1, x_2, x_3, x_0)$  and another vector and use Schwarz inequality.

**27.15.** Starting with the definition of each matrix, derive Equation (27.40).

**27.16.** Let  $D_1$  and  $D_2$  be derivations of a Lie algebra  $\mathfrak{v}$ . Show that  $D_1 D_2 \equiv D_1 \circ D_2$  is not a derivation, but  $[D_1, D_2]$  is.

**27.17.** Let  $\mathfrak{v}$  be a Lie algebra. Verify that  $\text{ad}_X$  is a derivation of  $\mathfrak{v}$  for any  $X \in \mathfrak{v}$ , and that  $\text{ad}_{[X, Y]} = [\text{ad}_X, \text{ad}_Y]$

**27.18.** Show that  $\psi : \mathfrak{v} \rightarrow \mathfrak{ad}_{\mathfrak{v}}$  given by  $\psi(X) = \text{ad}_X$  is (a) a homomorphism, (b)  $\ker \psi$  is the center of  $\mathfrak{v}$ , and (c)  $\mathfrak{ad}_{\mathfrak{v}}$  is an ideal of  $\mathfrak{D}_{\mathfrak{v}}$ .

**27.19.** Suppose that a Lie algebra  $\mathfrak{v}$  can be decomposed into a direct sum of Lie subalgebras. Show that each subalgebra is necessarily an ideal of  $\mathfrak{v}$ .

**27.20.** Show that if  $\psi$  is an automorphism of  $\mathfrak{v}$ , then

$$\text{ad}_{\psi(X)} = \psi \circ \text{ad}_X \circ \psi^{-1} \quad \forall X \in \mathfrak{v}.$$

Hint: Apply both sides to an arbitrary element of  $\mathfrak{v}$ .

**27.21.** Show that for any Lie algebra,

$$c_{ijk} = c_{js}^l c_{il}^r c_{kr}^s + c_{si}^l c_{lj}^r c_{rk}^s$$

is completely antisymmetric in all its indices.

**27.22.** Show that the Killing form of  $\mathfrak{v}$  is invariant under all automorphisms of  $\mathfrak{v}$ .

**27.23.** Show that the translation generators  $\mathbf{P}_j$  of the Poincaré algebra  $\mathfrak{p}(p, n-p)$  form a commutative ideal.

**27.24.** Find the Cartan metrics for  $\mathfrak{o}(3, 1)$  and  $\mathfrak{p}(3, 1)$ , and show directly that the first is semisimple but the second is not.

**27.25.** Show that the operation on a compact group defined by

$$(u|v) \equiv \int_G \langle \mathbf{T}_g u | \mathbf{T}_g v \rangle d\mu_g$$

is an inner product.

**27.26.** Show that the Weyl operator  $\mathbf{K}_u$  is hermitian.

**27.27.** Derive Equations (27.53) and (27.54). Hint: Follow the finite-group analogy.



**27.28.** Suppose that a Lie group  $G$  acts on a Euclidean space  $\mathbb{R}^n$  as well as on the space of (square-integrable) functions  $\mathcal{L}(\mathbb{R}^n)$ . Let  $\phi_i^{(\alpha)}$  transform as the  $i$ th row of the  $\alpha$ th irreducible representation. Verify that the relation

$$\mathbf{T}_g \phi_i^{(\alpha)}(\mathbf{x}) = \sum_{j=1}^{n_\alpha} T_{ji}^{(\alpha)}(g) \phi_j^{(\alpha)}(\mathbf{x} \cdot g^{-1})$$

defines a representation of  $G$ .

**27.29.** Show that  $GL(\mathcal{V})$  is not a compact group. Hint: Find a continuous function  $GL(\mathcal{V}) \rightarrow \mathbb{C}$  whose image is not compact.

**27.30.** Suppose that  $T : G \rightarrow GL(\mathcal{V})$  is a representation, and let

$$\mathcal{V}^{\otimes r} \equiv \underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{r \text{ times}}$$

be the  $r$ -fold tensor product of  $\mathcal{V}$ . Show that  $T^{\otimes r} : G \rightarrow GL(\mathcal{V}^{\otimes r})$ , given by

$$\mathbf{T}_g^{\otimes r}(\mathbf{v}_1, \dots, \mathbf{v}_r) = \mathbf{T}_g(\mathbf{v}_1) \otimes \cdots \otimes \mathbf{T}_g(\mathbf{v}_r),$$

is also a representation.

**27.31.** Suppose that in Example 27.4.2, we set  $k_1 = 2$  for our treatment of  $n = 2$ ,  $r = 3$ . Show that  $\mathbf{Y}_2(\mathbf{e}_{k_1} \otimes \mathbf{e}_{k_2} \otimes \mathbf{e}_{k_3})$  does not produce any new vector beyond what we obtained for  $k_1 = 1$ .

**27.32.** Show that  $g^{ij} g^{sr} c_{iks}$  is antisymmetric in  $j$  and  $r$ .

**27.33.** Operate  $L_+$  on  $|00\rangle = \sum_{m=-l}^l C(l; 0|m, -m; 0) |lm; l, -m\rangle$  and use  $L_+ |00\rangle = 0$  to find a recursive relation among  $C(l; 0|m, -m; 0)$ . Use normalization and the convention that  $C(l; 0|m, -m; 0) > 0$  to show that  $C(l; 0|m, -m; 0) = (-1)^{l-m} / \sqrt{2l+1}$  (see Section 12.3).

**27.34.** Show that the generators of  $\mathfrak{so}(3, 1)$ ,

$$\mathbf{M} \equiv (M_1, M_2, M_3) \equiv (M_{23}, M_{31}, M_{12}),$$

$$\mathbf{N} \equiv (N_1, N_2, N_3) \equiv (M_{01}, M_{02}, M_{03}),$$

satisfy the commutation relations

$$[M_i, M_j] = -\epsilon_{ijk} M_k, \quad [N_i, N_j] = \epsilon_{ijk} M_k, \quad [M_i, N_j] = -\epsilon_{ijk} N_k,$$

and that  $\mathbf{M}^2 - \mathbf{N}^2$  and  $\mathbf{M} \cdot \mathbf{N}$  commute with all the  $M$ 's and the  $N$ 's.

**27.35.** Let the double-indexed "metric" of the Poincaré algebra be defined as

$$g_{ij,kl} = c_{ij,mn}^{rs} c_{kl,rs}^{mn} + c_{ij,m}^r c_{kl,r}^m,$$

where the structure constants are given in Equation (27.69). Show that

$$g_{ij,kl} = 2(n-1)(\eta_{jk}\eta_{il} - \eta_{ik}\eta_{jl}).$$

27.36. Show that  $[\mathbf{M}^2, \mathbf{M}^{ij}] = 0$ , and

$$[\mathbf{M}^2, \mathbf{P}_k] = 4M_{kj}\mathbf{P}^j + 2(n-1)\mathbf{P}_k.$$

27.37. Show that the vector operator

$$\mathbf{C}_i \equiv M_{ij}\mathbf{P}^j = \eta^{kj}M_{ij}\mathbf{P}_k$$

satisfies the following commutation relations:

$$[\mathbf{C}_i, \mathbf{P}_j] = \eta_{ij}\mathbf{P}^2 - \mathbf{P}_i\mathbf{P}_j, \quad [\mathbf{C}_i, M_{jk}] = \eta_{ik}\mathbf{C}_j - \eta_{ij}\mathbf{C}_k, \quad [\mathbf{C}_i, \mathbf{C}_j] = M_{ij}\mathbf{P}^2.$$

Show also that  $[\mathbf{C}^2, M_{jk}] = 0$ ,  $\mathbf{C}^i\mathbf{P}_i = 0$ , and

$$\mathbf{P}^i\mathbf{C}_i = -(n-1)\mathbf{P}^2, \quad [\mathbf{C}^2, \mathbf{P}_i] = \{2\mathbf{C}_i + (n-1)\mathbf{P}_i\}\mathbf{P}^2.$$

27.38. Derive Equation (27.72) and show that  $\mathbf{W}^{i_1 \dots i_{n-3}}$  commutes with all the  $\mathbf{P}_j$ 's.

27.39. Let  $\hat{\mathbf{e}}_x \equiv (x_1, \dots, x_n)$  be any unit vector in  $\mathbb{R}^n$ .

(a) Show that a matrix is  $\eta$ -orthogonal, i.e., it satisfies Equation (27.33), if and only if its columns are  $\eta$ -orthogonal.

(b) Show that there exists an  $A \in O(p, n-p)$  such that  $\hat{\mathbf{e}}_x = A\hat{\mathbf{e}}_1$  where  $\hat{\mathbf{e}}_1 = (1, 0, \dots, 0)$ . Hint: Find the first column of  $A$  and use (a).

(c) Conclude that  $O(p, n-p)$  is transitive in its action on the collection of all vectors of the same length.

27.40. Verify directly that when the 4-momentum has only a time component, the Casimir operator  $\mathbf{W}^2 = \mathbf{W} \cdot \mathbf{W}$  reduces essentially to the total angular momentum operator.

27.41. Verify that for the case of a massless particle, when  $\mathbf{p}_0 = (p, 0, 0, p)$ ,

$$W_0 = W_1 = M_{23}, \quad W_2 = 2p(M_{13} - M_{03}), \quad W_3 = 2p(M_{02} - M_{12}),$$

and that  $\mathbf{W}^2 = \mathbf{W} \cdot \mathbf{W}$  annihilates  $|\psi_{\mathbf{p}_0}\rangle$ .

## Additional Reading

1. Barut, A. and Raczka, R *Theory of Group Representations and Applications*, World Scientific, 1986. A comprehensive introduction to Lie groups and Lie algebras using the modern language of manifolds. Intended for physicists and mathematicians alike, the first chapter is a long introduction to Lie algebras.
2. Chevalley, C. *Theory of Lie Groups*, Princeton University Press, 1946. Still a relevant classic for Lie groups and manifold theory written by one of the major contributors to the subject.

3. Gilmore, R. *Lie Groups, Lie Algebras, and Some of Their Applications*, Wiley, 1974. Although the concept of manifold is introduced, no heavy use of it is made in the treatment of Lie groups and Lie algebras, and the “parametric” method of Sophus Lie is employed throughout the book. Nevertheless, the book does a good job of classifying the Lie algebras for outsiders.
4. Hamermesh, M. *Group Theory and Its Application to Physical Problems*, Dover, 1989. Does not use the modern language of manifolds, but does a good job of introducing Lie groups and Lie algebras via “parameters.”

# Differential Geometry

The elegance of the geometrical expression of physical ideas has attracted much attention ever since Einstein proposed his geometrical theory of gravity in 1916. Such an interpretation was, however, confined to the general theory of relativity until the 1970s when the language of geometry was found to be most suitable, not only for gravity, but also for the other three fundamental forces of nature. Geometry, in the form of gauge field theories of electroweak and strong interactions, has been successful not only in creating a model—the so-called **standard model**—that explains all experimental results to remarkable accuracy, but also in providing a common language for describing all fundamental forces of nature, and with that a hope for unifying these forces into a single all-embracing force. This hope is encouraged by the successful unification of electromagnetism with the weak nuclear force through the medium of geometry and gauge field theory.

The word “geometry” is normally used in the mathematics literature for a manifold on which a “machine” is defined with the property that it gives a number when two vectors are fed into it. Symplectic geometry’s machine was a nondegenerate 2-form. Riemannian (or pseudo-Riemannian) geometry has a symmetric bilinear form (metric, inner product). Both of these geometries are important: Symplectic geometry is the natural setting for Hamiltonian dynamics, and Riemannian geometry is the basis of the general theory of relativity. In this chapter, we shall study the latter.<sup>1</sup> However, before introducing the metric, let us investigate some related structures that are independent of a metric.

---

<sup>1</sup>This chapter really belongs to the previous part of the book; however, because of our use of certain Lie-group theoretic ideas in Sections 28.4 and 28.6, we have included it here.

## 28.1 Vector Fields and Curvature

A manifold is, in general, not flat (flatness will be defined later). One way to “feel” the curvature of a space *intrinsically* is to translate a vector parallel to itself along different paths and compare the final vectors. In a flat space the two vectors at the end will be the same, but not in a general space. An illustration is provided by the surface of a sphere. Assume that we have a vector perpendicular to the equator. To exaggerate the effect of curvature, we move the vector parallel to itself on the equator a quarter of the way around the sphere, then all the way to the north pole. Alternatively, we start with the vector again perpendicular to the equator, but this time we move it parallel to itself directly to the north pole. Clearly, the two final vectors will not be the same; in fact, they will be perpendicular to one another.

The above intuitive discussion should help to make it clear that to find the curvature of space, we look at how vectors change. In analogy with the exterior derivative and forms, we want to introduce a derivative that operates on vectors. In fact, since we already have the exterior derivative available, let us see if we can extend it to vectors.

Consider an arbitrary 1-form  $\omega$  and an arbitrary vector field  $\mathbf{v}$ . The pairing  $\langle \omega, \mathbf{v} \rangle$  is a real-valued function  $f$  on which we know how the exterior derivative acts. A natural extension of  $d$  (which we denote by the same symbol) is given as follows:

$$df = d \langle \omega, \mathbf{v} \rangle \equiv \langle d\omega, \mathbf{v} \rangle' + \langle \omega, d\mathbf{v} \rangle', \quad (28.1)$$

where we have used a prime to designate the new pairing. In general, this new pairing may be different from the ordinary pairing, because for the participants in the latter, no exterior differentiation is defined. As we shall see below, we indeed have to change the old pairing slightly for the new pairing to make sense. The LHS of Equation (28.1) is a 1-form. The first term on the RHS is a 2-form contracted with a vector, i.e., a 1-form. For the second term to be a 1-form,  $d\mathbf{v}$  must be a tensor that contracts with a 1-form to give a 1-form. We say that  $d\mathbf{v}$  is a **vector-valued 1-form**.

vector-valued  
1-forms

Let us take a basis  $\{\mathbf{e}_i\}$  and its dual  $\{\epsilon^j\}$  and express Equation (28.1) in terms of components in those bases. Then, on the one hand, with  $\omega = \omega_i \epsilon^i$  and  $\mathbf{v} = v^j \mathbf{e}_j$ , we have  $\langle \omega, \mathbf{v} \rangle = \omega_i v^i$  and

$$df = d \langle \omega, \mathbf{v} \rangle = (d\omega_i) v^i + \omega_i dv^i. \quad (28.2)$$

On the other hand,

$$\begin{aligned} df &= \langle d\omega_i \wedge \epsilon^i + \omega_i d\epsilon^i, v^j \mathbf{e}_j \rangle' + \langle \omega_i \epsilon^i, dv^j \mathbf{e}_j + v^j d\mathbf{e}_j \rangle' \\ &= v^j \langle d\omega_i \wedge \epsilon^i, \mathbf{e}_j \rangle' + v^j \omega_i \langle d\epsilon^i, \mathbf{e}_j \rangle' + \omega_i \langle \epsilon^i, dv^j \mathbf{e}_j \rangle' + \omega_i v^j \langle \epsilon^i, d\mathbf{e}_j \rangle' \\ &= v^j \langle d\omega_i \wedge \epsilon^i, \mathbf{e}_j \rangle' + \omega_i dv^j \underbrace{\langle \epsilon^i, \mathbf{e}_j \rangle'}_{\delta_j^i} + v^j \omega_i \{ \langle d\epsilon^i, \mathbf{e}_j \rangle' + \langle \epsilon^i, d\mathbf{e}_j \rangle' \}, \end{aligned} \quad (28.3)$$

where we have assumed that the primed pairing is equivalent to the old pairing when no derivatives are inside. If Equations (28.2) and (28.3) are to hold for arbitrary  $\omega$  and  $\mathbf{v}$ , we must have<sup>2</sup>

$$\begin{aligned} 0 &= \langle d\epsilon^i, \mathbf{e}_j \rangle' + \langle \epsilon^i, d\mathbf{e}_j \rangle', \\ (d\omega_i)v^i &= v^j \langle d\omega_i \wedge \epsilon^i, \mathbf{e}_j \rangle'. \end{aligned} \quad (28.4)$$

The first relation is simply the fact that the exterior derivative of  $\delta_j^i = \langle \epsilon^i, \mathbf{e}_j \rangle$  is zero. The second relation defines the order of contraction of vectors and higher-order forms in the new pairing. Since  $d\omega_i$  is a 1-form, we can write it as  $d\omega_i = \alpha_{ik}\epsilon^k$ , so that

$$d\omega_i \wedge \epsilon^i = \alpha_{ik}\epsilon^k \wedge \epsilon^i, \quad \text{where} \quad \alpha_{ik} = -\alpha_{ki}.$$

Then, the second equation of (28.4) gives

$$\alpha_{ik} \langle \epsilon^k \wedge \epsilon^i, \mathbf{e}_j \rangle' = d\omega_j = \alpha_{jk}\epsilon^k = \frac{1}{2}\alpha_{ik}(\delta_{ij}\epsilon^k - \delta_{kj}\epsilon^k).$$

This equation demands that the new pairing of a vector with a wedge product of 1-forms be defined as

$$\langle \epsilon^k \wedge \epsilon^i, \mathbf{e}_j \rangle' \equiv -i_{\mathbf{e}_j}(\epsilon^k \wedge \epsilon^i) = -\frac{1}{2}(\delta_{kj}\epsilon^i - \delta_{ij}\epsilon^k),$$

where  $i_{\mathbf{e}_j}$  is the interior product of Definition 26.5.8. We summarize the properties of the new pairing in the following equation:

$$\begin{aligned} \langle d\epsilon^i, \mathbf{e}_j \rangle' + \langle \epsilon^i, d\mathbf{e}_j \rangle' &= 0, \\ \langle \epsilon^k \wedge \epsilon^i, \mathbf{e}_j \rangle' &\equiv -i_{\mathbf{e}_j}(\epsilon^k \wedge \epsilon^i) = -\frac{1}{2}(\delta_{kj}\epsilon^i - \delta_{ij}\epsilon^k). \end{aligned} \quad (28.5)$$

In differential structures everything takes place locally, and translations and movements are all infinitesimal. Let us look at the exterior derivative from this point of view. For a real-valued function on  $M$ , we have

$$d^2 f = d\left(\frac{\partial f}{\partial x^i} dx^i\right) = \sum_{j < i} \left( \frac{\partial^2 f}{\partial x^j \partial x^i} - \frac{\partial^2 f}{\partial x^i \partial x^j} \right) dx^i \wedge dx^j.$$

Thus,  $d^2 f = 0$  means that the mixed partial derivatives are independent of the order of differentiation—a familiar result. Geometrically, this means that for small displacements,  $dx^i$  and  $dx^j$ , the value of a function is the same if one moves in two (perpendicular) directions, once in a given order and then in reverse order. This is

<sup>2</sup>These consistency relations apply to the first application of exterior derivative to the pairing. Higher-order applications, or, equivalently, application of  $d$  to a pairing involving wedge products of forms, may require new consistency relations. Fortunately, for our purposes, the first-order consistency relations will be sufficient.

true even if the space is curved. For flat spaces, we know that the same conclusion holds for displacement (parallel to themselves) of *vectors*. When we interpret  $d^2\mathbf{v}$  as the change in a vector as it is displaced in two different directions, then the example of the sphere above suggests that  $d^2\mathbf{v}$  *must be related to curvature*. Let us find this relation.

Starting with a basis  $\{\mathbf{e}_i\}$  and an arbitrary vector  $\mathbf{v} = v^i\mathbf{e}_i$ , operate on it with  $d$  twice, keeping in mind that its action on functions and differential forms is exactly the same as the exterior derivative defined before:  $d\mathbf{v} = dv^i\mathbf{e}_i + v^i d\mathbf{e}_i$  and

$$d^2\mathbf{v} = \underbrace{d^2v^i}_{=0}\mathbf{e}_i + \underbrace{(-1)^1 dv^i \wedge d\mathbf{e}_i + dv^i \wedge d\mathbf{e}_i}_{=0} + v^i d^2\mathbf{e}_i = v^i d^2\mathbf{e}_i. \quad (28.6)$$

This equation has a remarkable property: It leaves the components of  $\mathbf{v}$  undifferentiated! In other words, regardless of how any given two vectors  $\mathbf{v}$  and  $\mathbf{w}$  vary away from the point  $P$  of the manifold,  $d^2\mathbf{v} = d^2\mathbf{w}$  as long as the two vectors are equal at  $P$ . More importantly, if we take the linearity of  $d$  (and therefore,  $d^2$ ) into account, we obtain

$$\boxed{d^2(f\mathbf{v} + g\mathbf{u}) = f d^2\mathbf{v} + g d^2\mathbf{u}.} \quad (28.7)$$

It appears that  $d^2\mathbf{v}$  depends not on external objects (vectors), but on the intrinsic property of the manifold, i.e., how it “curves” away from  $P$ . To find this curvature, expand the vector-valued 1-form  $d\mathbf{e}_i$  as

$$d\mathbf{e}_i = \mathbf{e}_j \otimes \omega^j_i, \quad \text{where} \quad \omega^j_i \equiv \Gamma^j_{ik} \epsilon^k. \quad (28.8)$$

As we shall see shortly, one has to be cautious to know which index is raised or lowered. In the formulas above, this caution has been observed by leaving blank the original position of the raised (or lowered) index. Differentiating Equation (28.8) once more, we obtain the vector-valued 2-form

extracting the curvature 2-form from second exterior derivative of basis vectors

$$\begin{aligned} d^2\mathbf{e}_i &= d\mathbf{e}_j \wedge \omega^j_i + \mathbf{e}_j \otimes d\omega^j_i = (\mathbf{e}_k \otimes \omega^k_j) \wedge \omega^j_i + \mathbf{e}_j \otimes d\omega^j_i \\ &\equiv \mathbf{e}_k \otimes (\omega^k_j \wedge \omega^j_i) + \mathbf{e}_j \otimes d\omega^j_i = \mathbf{e}_j \otimes (d\omega^j_i + \omega^j_k \wedge \omega^k_i). \end{aligned} \quad (28.9)$$

The expression in parentheses is a 2-form, called the **curvature two-form**:

$$\theta^j_i \equiv d\omega^j_i + \omega^j_k \wedge \omega^k_i, \quad \text{or} \quad \theta_{ij} = d\omega_{ij} + \omega_{ik} \wedge \omega^k_j. \quad (28.10)$$

With this notation, Equation (28.6) becomes

$$d^2\mathbf{v} = v^i d^2\mathbf{e}_i = v^i \mathbf{e}_j \otimes \theta^j_i = \mathbf{e}_j \otimes \theta^j_i v^i \equiv \Theta \otimes \mathbf{v}, \quad (28.11)$$

where the last expression is simply an abbreviation.

So far, we have been dealing with the exterior derivatives of the basis vectors  $\{e_i\}$ . What about the exterior derivatives of the dual basis vectors  $\{\epsilon^j\}$ ? They are closely related to  $\{de_i\}$ , as the following argument shows. The first relation of (28.5), as well as Equation (28.8) and the fact that  $\epsilon^j$  pairs up with  $e_k$  yield

$$\langle e_i, d\epsilon^j \rangle' = -\langle de_i, \epsilon^j \rangle' = -\langle e_k \otimes \omega^k_i, \epsilon^j \rangle' = -\omega^k_i \langle e_k, \epsilon^j \rangle = -\omega^j_i \tag{28.12}$$

On the other hand, since  $d\epsilon^j$  is a 2-form, it can be written as

$$d\epsilon^j = \sigma^j_k \wedge \epsilon^k \equiv \gamma^j_{km} \epsilon^m \wedge \epsilon^k, \quad \text{where} \quad \gamma^j_{km} = -\gamma^j_{mk}$$

and

$$\begin{aligned} \langle e_i, d\epsilon^j \rangle' &= \langle e_i, \gamma^j_{km} \epsilon^m \wedge \epsilon^k \rangle' = \gamma^j_{km} \langle e_i, \epsilon^m \wedge \epsilon^k \rangle' = -\gamma^j_{km} i_{e_i}(\epsilon^m \wedge \epsilon^k) \\ &= \frac{1}{2} \gamma^j_{km} (\delta^k_i \epsilon^m - \delta^m_i \epsilon^k) = \gamma^j_{im} \epsilon^m = \sigma^j_i, \end{aligned} \tag{28.13}$$

where we used the second equation in (28.5). Comparing (28.13) with Equation (28.12) yields  $\sigma^j_i = -\omega^j_i$ . We therefore have

$$d\epsilon^j = -\omega^j_k \wedge \epsilon^k. \tag{28.14}$$

Equations (28.8) and (28.14) show that the  $\{\omega_{ij}\}$  give all the information about how the bases  $\{e_i\}$  and  $\{\epsilon^j\}$  change with infinitesimal movement away from a point  $P$ .

**28.1.1. Box.** *If we can find the  $\{\omega_{ij}\}$ , we will know the (local) geometry of the manifold.*

connection coefficients

From the definition of  $\Gamma^i_{jk}$  in Equation (28.8), we have  $\omega_{ij} = \Gamma_{ijk} \epsilon^k$ . The functions  $\Gamma_{ijk}$  are called the **connection coefficients**. Because of (28.8), these coefficients are antisymmetric in their first two indices. On the other hand, (28.14) gives

$$\begin{aligned} d\epsilon^k &= -\omega^k_j \wedge \epsilon^j = -g^{ik} \omega_{ij} \wedge \epsilon^j = -g^{ik} \Gamma_{ijm} \epsilon^m \wedge \epsilon^j \\ &= \frac{1}{2} g^{ik} (\Gamma_{ijm} - \Gamma_{imj}) \epsilon^j \wedge \epsilon^m. \end{aligned} \tag{28.15}$$

If coordinate frames are used for basis vectors, so that  $\epsilon^k = dx^k$ , the LHS of (28.15) will be zero and the coefficients on the RHS of Equation (28.15) must vanish, i.e., the connection coefficients are symmetric in their last two indices:

$$\Gamma_{ijm} = \Gamma_{imj} \quad \text{in coordinate frames.} \tag{28.16}$$



Equation (28.15) shows further that by calculating  $d\epsilon^k$  we merely determine the antisymmetric combination  $\Gamma_{ijm} - \Gamma_{imj}$ . However, the antisymmetry of  $\Gamma_{ijm}$  in its first two indices can be used to determine it completely. Let  $C_{ijm} \equiv \Gamma_{ijm} - \Gamma_{imj}$  be the coefficients that can be read off from (28.15). Then one can show that

$$\Gamma_{ijm} = \frac{1}{2}(C_{ijm} + C_{jmi} - C_{mij}). \tag{28.17}$$

Once the  $\Gamma$ 's are determined by this relation, they can be used to write  $\omega_{ji}$ 's as a linear combination of the dual basis vectors.

## 28.2 Riemannian Manifolds

As mentioned before, manifolds that possess a metric are important. The general theory of relativity, for example, is entirely based on the existence of a metric. In fact, it is the job of that theory to *determine* the metric of 4-dimensional space-time from a knowledge of the distribution of matter.

Riemannian and pseudo-Riemannian manifolds defined

**28.2.1. Definition.** A *Riemannian manifold* is a differentiable manifold  $M$  with a symmetric tensor field  $\mathbf{g} \in T_2^0(M)$ , called the *metric*, such that at each point  $P \in M$ ,  $\mathbf{g}|_P$  is a positive definite inner product. A manifold with an indefinite inner product at each point is called a *pseudo-Riemannian manifold*.

orthonormal frames

With  $\mathbf{g}$  defined on  $M$ , we can obtain orthonormal vectors at each point of  $M$ . That is, we can construct **orthonormal frames**  $\{\mathbf{e}_i\}$  such that

$$\mathbf{e}_i \cdot \mathbf{e}_j \equiv \mathbf{g}(\mathbf{e}_i, \mathbf{e}_j) = \eta_{ij} \equiv \pm\delta_{ij}$$

at each point  $P \in M$ .

### 28.2.1 Curvature via Connection

We can find a relation between the metric tensor and  $\omega_{ij}$  by taking the exterior derivative of both sides of  $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$  and using Equation (28.8):

$$\begin{aligned} dg_{ij} &= (d\mathbf{e}_i) \cdot \mathbf{e}_j + \mathbf{e}_i \cdot (d\mathbf{e}_j) = (\mathbf{e}_k \otimes \omega^k_i) \cdot \mathbf{e}_j + \mathbf{e}_i \cdot (\mathbf{e}_k \otimes \omega^k_j) \\ &\equiv \omega^k_i \otimes (\mathbf{e}_k \cdot \mathbf{e}_j) + \omega^k_j \otimes (\mathbf{e}_i \cdot \mathbf{e}_k) = \omega^k_i g_{kj} + \omega^k_j g_{ik} \equiv \omega_{ji} + \omega_{ij}. \end{aligned}$$

structure equations, integrability condition, and curvature matrix

In particular, if we work in an orthonormal basis, then  $g_{ij} = \pm\delta_{ij}$ , and the LHS will be zero.<sup>3</sup> In such a case, we obtain the following antisymmetry condition for the 1-forms  $\omega_{ij}$ :

$$\omega_{ij} + \omega_{ji} = 0. \tag{28.18}$$

<sup>3</sup>Orthonormality is not really required. All that is necessary is for  $g_{ij}$  to be constant.

We now develop an algorithm to determine the local curvature of the manifold. Choose an *orthonormal* basis and its dual and introduce the matrices

$$\mathbf{e} = (\mathbf{e}_1 \dots \mathbf{e}_m), \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^m \end{pmatrix}, \quad \boldsymbol{\Omega} = \begin{pmatrix} 0 & \omega_{12} & \omega_{13} & \dots & \omega_{1m} \\ -\omega_{12} & 0 & \omega_{23} & \dots & \omega_{2m} \\ -\omega_{13} & -\omega_{23} & 0 & \dots & \omega_{3m} \\ \vdots & \vdots & & \ddots & \vdots \\ -\omega_{1m} & -\omega_{2m} & -\omega_{3m} & \dots & 0 \end{pmatrix},$$

whose elements are one-forms, or vectors. Write Equations (28.8), (28.18), (28.14), and (28.9) as

$$\begin{aligned} d\mathbf{e} &= \mathbf{e}\mathbf{G}\boldsymbol{\Omega}, & \boldsymbol{\Omega} + \boldsymbol{\Omega}^t &= 0, & d\boldsymbol{\varepsilon} &= -\mathbf{G}\boldsymbol{\Omega} \wedge \boldsymbol{\varepsilon}; \\ d^2\mathbf{e} &= \mathbf{e}\boldsymbol{\Theta}, & \text{where} & & \boldsymbol{\Theta} &= d\boldsymbol{\Omega} + \boldsymbol{\Omega} \wedge (\mathbf{G}\boldsymbol{\Omega}). \end{aligned} \quad (28.19)$$

The matrix  $\mathbf{G}$  is the matrix of the metric with components  $g_{ij} = \pm\delta_{ij}$ . It is introduced to raise the indices when the equations are written in component form.<sup>4</sup> The first two equations in (28.19) are called the **structure equations**, the third is called the **integrability condition**, and  $\boldsymbol{\Theta}$  is called the **curvature matrix**.

No great mind of the past has exerted a deeper influence on the mathematics of the twentieth century than **Georg Friedrich Bernhard Riemann** (1826–1866), the son of a poor country minister in northern Germany. He studied the works of Euler and Legendre while he was still in secondary school, and it is said that he mastered Legendre's treatise on the theory of numbers in less than a week. But he was shy and modest, with little awareness of his own extraordinary abilities, so at the age of 19 he went to the University of Göttingen with the aim of pleasing his father by studying theology and becoming a minister himself. Fortunately, this worthy purpose soon stuck in his throat, and with his father's willing permission he switched to mathematics.



The presence of the legendary Gauss automatically made Göttingen the center of the mathematical world. But Gauss was remote and unapproachable—particularly to beginning students—and after only a year Riemann left this unsatisfying environment and went to the University of Berlin. There he attracted the friendly interest of Dirichlet and Jacobi, and learned a great deal from both men. Two years later he returned to Göttingen, where he obtained his doctor's degree in 1851. During the next 8 years, despite debilitating poverty, he created his greatest works. In 1854 he was appointed Privatdozent (unpaid lecturer), which at that time was the necessary first step on the academic ladder. Gauss died in 1855, and Dirichlet was called to Göttingen as his successor. Dirichlet helped Riemann in every way he could, first with a small salary (about one-tenth of that paid to a full professor) and then with a promotion to an assistant professorship. In 1859 he also died, and Riemann was

<sup>4</sup>Strictly speaking, we should use  $\mathbf{G}^{-1}$  instead of  $\mathbf{G}$ . But since  $g_{ij} = \pm\delta_{ij}$ , the two are identical.

appointed as a full professor to replace him. Riemann's years of poverty were over, but his health was broken. At the age of 39 he died of tuberculosis in Italy, on the last of several trips he undertook in order to escape the cold, wet climate of northern Germany. Riemann had a short life and published comparatively little, but his works permanently altered the course of mathematics in analysis, geometry, and number theory.

It is said that the three greatest mathematicians of modern times are Euler, Gauss, and Riemann. It is a curiosity of nature that these three names are among the most frequently mentioned names in the physics literature as well. Aside from the indirect use of his name in the application of complex analysis in physics, **Riemannian geometry** has become the most essential building block of all theories of fundamental interactions, starting with gravity, which Einstein formulated in this language in 1916. As part of the requirement to become a Privatdozent, Riemann had to write a probationary essay and to present a trial lecture to the faculty. It was the custom for the candidate to offer three titles, and the head of his department usually accepted the first. However, Riemann rashly listed as his third topic the foundations of geometry. Gauss, who had been turning this subject over in his mind for 60 years, was naturally curious to see how this particular candidate's "gloriously fertile originality" would cope with such a challenge, and to Riemann's dismay he designated this as the subject of the lecture. Riemann quickly tore himself away from his other interests at the time—"my investigations of the connection between electricity, magnetism, light, and gravitation"—and wrote his lecture in the next two months. The result was one of the great classical masterpieces of mathematics, and probably the most important scientific lecture ever given. It is recorded that even Gauss was surprised and enthusiastic.

**Bianchi identity** We can derive further integrability conditions. For instance, applying  $d$  to the third equation of (28.19) gives

$$0 = d^2\varepsilon = -d(G\Omega) \wedge \varepsilon + G\Omega \wedge d\varepsilon = -Gd\Omega \wedge \varepsilon - (G\Omega) \wedge (G\Omega) \wedge \varepsilon.$$

Multiplying both sides on the left by  $G = G^{-1}$ , we obtain

$$0 = -[d\Omega + \Omega \wedge (G\Omega)] \wedge \varepsilon = -\Theta \wedge \varepsilon. \tag{28.20}$$

Similarly, the reader may show that

$$d\Theta = \Omega \wedge (G\Theta) - \Theta \wedge (G\Omega). \tag{28.21}$$

This is called the **Bianchi identity**.

**Riemann curvature tensor**

The  $ij$ th element of the matrix of  $\Theta$  can be written as

$$\theta_{ij} = \frac{1}{2}R_{ijkl}\epsilon^k \wedge \epsilon^l, \tag{28.22}$$

which defines the components  $R_{ijkl}$  of the **Riemann curvature tensor**. The antisymmetry of the matrix  $\Theta$  (showing this is left as a problem for the reader) and Equation (28.22) give

$$R_{ijkl} + R_{jikl} = 0, \quad R_{ijkl} + R_{ijlk} = 0. \tag{28.23}$$

Square brackets mean antisymmetrization.

Similarly, the relation  $\Theta \wedge \varepsilon = 0$  of Equation (28.20) can be shown to be equivalent to

$$R_{ijkl} + R_{iklj} + R_{iljk} = 0 \quad \text{and} \quad R_{i[jkl]} = 0, \tag{28.24}$$

where

**28.2.2. Box.** *The enclosure of indices in square brackets means complete antisymmetrization of those indices.*

When coordinate bases are used,  $\omega_{ij}$  is no longer antisymmetric, but we still have

$$dg_{ij} = \omega_{ij} + \omega_{ji} \equiv \Gamma_{ijk}dx^k + \Gamma_{jik}dx^k.$$

Since  $dg_{ij} = (\partial g_{ij}/\partial x^k)dx^k$ , we get

$$g_{ij,k} \equiv \frac{\partial g_{ij}}{\partial x^k} = \Gamma_{ijk} + \Gamma_{jik}. \tag{28.25}$$

Christoffel symbol

Using Equations (28.16) and (28.25), we can readily show that

$$\Gamma_{ijk} = \frac{1}{2}(g_{ij,k} + g_{ik,j} - g_{kj,i}) \equiv \frac{1}{2} \left( \frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{kj}}{\partial x^i} \right). \tag{28.26}$$

connection between infinitesimal displacement, arc length, and metric tensor

This is the **Christoffel symbol** used in classical tensor analysis. Now we consider the connection between an infinitesimal displacement and a metric. Let  $P$  be a point of  $M$ . Let  $\gamma$  be a curve through  $P$  such that  $\gamma(c) = P$ . For an infinitesimal number  $\delta u$ , let  $P' = \gamma(c + \delta u)$  be a point on  $\gamma$  close to  $P$ . Since the  $x^i$  are well-behaved functions,  $x^i(P') - x^i(P)$  are infinitesimal real numbers. Let  $\xi^i = x^i(P') - x^i(P)$ , and construct the vector  $\mathbf{v} = \xi^i \partial_i$ , where  $\{\partial_i\}$  consists of tangent vectors at  $P$ . We call  $\mathbf{v}$  the **infinitesimal displacement** at  $P$ . The length of this vector,  $\mathbf{g}(\mathbf{v}, \mathbf{v})$ , is shown to be  $g_{ij}\xi^i\xi^j$ . This is called the **arc length** from  $P$  to  $P'$ , and is naturally written as  $ds^2 = g_{ij}\xi^i\xi^j$ . It is customary to write  $dx^i$  (not a 1-form!) in place of  $\xi^i$ :

$$ds^2 = g_{ij}dx^i dx^j, \tag{28.27}$$

where the  $dx^i$  are infinitesimal real numbers.

**Elwin Bruno Christoffel** (1829–1900) came from a family in the cloth trade. He attended an elementary school in Montjoie (which was renamed Monschau in 1918) but then spent a number of years being tutored at home in languages, mathematics, and classics. He attended secondary schools from 1844 until 1849. At first he studied at the Jesuit gymnasium in

Cologne but moved to the Friedrich-Wilhelms Gymnasium in the same town for at least the three final years of his school education. He was awarded the final school certificate with a distinction in 1849. The next year he went to the University of Berlin and studied under a number of distinguished mathematicians, including Dirichlet.

After one year of military service in the Guards Artillery Brigade, he returned to Berlin to study for his doctorate, which was awarded in 1856 with a dissertation on the motion of electricity in homogeneous bodies. His examiners included mathematicians and physicists, Kummer being one of the mathematics examiners.

At this point Christoffel spent three years outside the academic world. He returned to Montjoie, where his mother was in poor health, but read widely from the works of Dirichlet, Riemann, and Cauchy. It has been suggested that this period of academic isolation had a major effect on his personality and on his independent approach towards mathematics. It was during this time that he published his first two papers on numerical integration, in 1858, in which he generalized Gauss's method of quadrature and expressed the polynomials that are involved as a determinant. This is now called Christoffel's theorem.



In 1859 Christoffel took the qualifying examination to become a university teacher and was appointed a lecturer at the University of Berlin. Four years later, he was appointed to a chair at the Polytechnicum in Zurich, filling the post left vacant when Dedekind went to Brunswick. Christoffel was to have a huge influence on mathematics at the Polytechnicum, setting up an institute for mathematics and the natural sciences there.

In 1868 Christoffel was offered the chair of mathematics at the Gewerbsakademie in Berlin, which is now the University of Technology of Berlin. However, after three years at the Gewerbsakademie in Berlin, Christoffel moved to the University of Strasbourg as the chair of mathematics, a post he held until he was forced to retire due to ill health in 1892.

Some of Christoffel's early work was on conformal mappings of a simply connected region bounded by polygons onto a circle. He also wrote important papers that contributed to the development of the tensor calculus of Gregorio Ricci-Curbastro and Tullio Levi-Civita. The Christoffel symbols that he introduced are fundamental in the study of tensor analysis. The Christoffel reduction theorem, so named by Klein, solves the local equivalence problem for two quadratic differential forms. The procedure Christoffel employed in his solution of the equivalence problem is what Ricci later called *covariant differentiation*; Christoffel also used the latter concept to define the basic Riemann-Christoffel *curvature tensor*. His approach allowed Ricci and Levi-Civita to develop a coordinate-free differential calculus which Einstein, with the help of Grossmann, turned into the tensor analysis, the mathematical foundation of general relativity.

In applications, it is common to start with the metric tensor  $\mathbf{g}$  given in terms of coordinate differential forms:

$$\mathbf{g} = g_{ij} dx^i \otimes dx^j, \quad \text{where } g_{ij} = g_{ji} = \mathbf{g}(\partial_i, \partial_j). \quad (28.28)$$

Then the orthonormal bases  $\{\mathbf{e}_i\}$  and  $\{\mathbf{e}^i\}$  are constructed in terms of  $\{\partial_i\}$  and  $\{dx^i\}$ , respectively, and are utilized as illustrated in the following examples. The

equivalence of the arc length [Equation (28.27)] and the metric [Equation (28.28)] is the reason why it is the arc length that is given in most practical problems. Once the arc length is known, the metric  $g_{ij}$  can be read off, and all the relevant geometric quantities can be calculated from it.

**28.2.3. Example.** Let us look at a few examples of arc lengths and the corresponding metrics.

- (a) For  $ds^2 = dx^2 + dy^2 + dz^2$ ,  $\mathbf{g}$  is the Euclidean metric of  $\mathbb{R}^3$ , with  $g_{ij} = \delta_{ij}$ .
- (b) For  $ds^2 = -dx^2 - dy^2 - dz^2 + dt^2$ ,  $\mathbf{g}$  is the Minkowski (or Lorentz) metric of  $\mathbb{R}^4$ , with  $g_{ij} = \eta_{ij}$ , where  $\eta_{xx} = \eta_{yy} = \eta_{zz} = -\eta_{tt} = -1$  and  $\eta_{ij} = 0$  for  $i \neq j$ .
- (c) For  $ds^2 = dr^2 + r^2(d\theta^2 + \sin^2 \theta d\varphi^2)$ , the metric is the Euclidean metric given in spherical coordinates in  $\mathbb{R}^3$  with  $g_{rr} = 1$ ,  $g_{\theta\theta} = r^2$ ,  $g_{\varphi\varphi} = r^2 \sin^2 \theta$ , and all other components zero.
- (d) For  $ds^2 = a^2 d\theta^2 + a^2 \sin^2 \theta d\varphi^2$ , the metric is that of a two-dimensional spherical surface, with  $g_{\theta\theta} = a^2$ ,  $g_{\varphi\varphi} = a^2 \sin^2 \theta$ , and all other components zero.

Friedmann metric

(e) For

$$ds^2 = dt^2 - a^2(t)[d\chi^2 + \sin^2 \chi(d\theta^2 + \sin^2 \theta d\varphi^2)],$$

the metric is the **Friedmann metric** used in cosmology. Here  $g_{tt} = 1$ ,  $g_{\chi\chi} = -[a(t)]^2$ ,  $g_{\theta\theta} = -[a(t)]^2 \sin^2 \chi$ ,  $g_{\varphi\varphi} = -[a(t)]^2 \sin^2 \chi \sin^2 \theta$ , and all other components are zero.

Schwarzschild metric

(f) For

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2 \theta d\varphi^2),$$

the metric is the **Schwarzschild metric** with  $g_{tt} = 1 - 2M/r$ ,  $g_{rr} = -(1 - 2M/r)^{-1}$ ,  $g_{\theta\theta} = -r^2$ ,  $g_{\varphi\varphi} = -r^2 \sin^2 \theta$ , and all other components zero.

For each of the arc lengths above, we have an orthonormal basis of one-forms:

- (a)  $\mathbf{g} = \epsilon^1 \otimes \epsilon^1 + \epsilon^2 \otimes \epsilon^2 + \epsilon^3 \otimes \epsilon^3$  with  $\epsilon^1 = dx$ ,  $\epsilon^2 = dy$ ,  $\epsilon^3 = dz$ ;
- (b)  $\mathbf{g} = -\epsilon^1 \otimes \epsilon^1 - \epsilon^2 \otimes \epsilon^2 - \epsilon^3 \otimes \epsilon^3 + \epsilon^0 \otimes \epsilon^0$  with  $\epsilon^1 = dx$ ,  $\epsilon^2 = dy$ ,  $\epsilon^3 = dz$ ,  $\epsilon^0 = dt$ ;
- (c)  $\mathbf{g} = \epsilon^r \otimes \epsilon^r + \epsilon^\theta \otimes \epsilon^\theta + \epsilon^\varphi \otimes \epsilon^\varphi$  with  $\epsilon^r = dr$ ,  $\epsilon^\theta = r d\theta$ ,  $\epsilon^\varphi = r \sin \theta d\varphi$ ;
- (d)  $\mathbf{g} = \epsilon^\theta \otimes \epsilon^\theta + \epsilon^\varphi \otimes \epsilon^\varphi$  with  $\epsilon^\theta = a d\theta$ ,  $\epsilon^\varphi = a \sin \theta d\varphi$ ;
- (e)  $\mathbf{g} = \epsilon^t \otimes \epsilon^t - \epsilon^\chi \otimes \epsilon^\chi - \epsilon^\theta \otimes \epsilon^\theta - \epsilon^\varphi \otimes \epsilon^\varphi$  with  $\epsilon^t = dt$ ,  $\epsilon^\chi = a(t) d\chi$ ,  $\epsilon^\theta = a(t) \sin \chi d\theta$ ,  $\epsilon^\varphi = a(t) \sin \chi \sin \theta d\varphi$ ;
- (f)  $\mathbf{g} = \epsilon^t \otimes \epsilon^t - \epsilon^r \otimes \epsilon^r - \epsilon^\theta \otimes \epsilon^\theta - \epsilon^\varphi \otimes \epsilon^\varphi$  with  $\epsilon^t = (1 - 2M/r)^{1/2} dt$ ,  $\epsilon^r = (1 - 2M/r)^{-1/2} dr$ ,  $\epsilon^\theta = r d\theta$ ,  $\epsilon^\varphi = r \sin \theta d\varphi$ . ■

**28.2.4. Example.** In this example, we examine the curvilinear coordinates used in vector analysis. Recall that in terms of these coordinates the displacement is given by  $ds^2 = h_1^2(dq_1)^2 + h_2^2(dq_2)^2 + h_3^2(dq_3)^2$ . Therefore, the orthonormal one-forms are  $\epsilon^1 = h_1 dq_1$ ,  $\epsilon^2 = h_2 dq_2$ ,  $\epsilon^3 = h_3 dq_3$ . We also note (see Problem 26.24) that

$$d * df = \left( \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right) dx \wedge dy \wedge dz = \nabla^2 f dx \wedge dy \wedge dz. \tag{28.29}$$

We use this equation to find the Laplacian in terms of  $q_1$ ,  $q_2$ , and  $q_3$ :

$$\begin{aligned} df &= \frac{\partial f}{\partial q_1} dq_1 + \frac{\partial f}{\partial q_2} dq_2 + \frac{\partial f}{\partial q_3} dq_3 \\ &= \left( \frac{1}{h_1} \frac{\partial f}{\partial q_1} \right) \epsilon^1 + \left( \frac{1}{h_2} \frac{\partial f}{\partial q_2} \right) \epsilon^2 + \left( \frac{1}{h_3} \frac{\partial f}{\partial q_3} \right) \epsilon^3, \end{aligned}$$

where we substituted orthonormal 1-forms so we could apply the Hodge star operator. It follows now that

$$\begin{aligned} *df &= \left( \frac{1}{h_1} \frac{\partial f}{\partial q_1} \right) * \epsilon^1 + \left( \frac{1}{h_2} \frac{\partial f}{\partial q_2} \right) * \epsilon^2 + \left( \frac{1}{h_3} \frac{\partial f}{\partial q_3} \right) * \epsilon^3 \\ &= \left( \frac{1}{h_1} \frac{\partial f}{\partial q_1} \right) \epsilon^2 \wedge \epsilon^3 + \left( \frac{1}{h_2} \frac{\partial f}{\partial q_2} \right) \epsilon^3 \wedge \epsilon^1 + \left( \frac{1}{h_3} \frac{\partial f}{\partial q_3} \right) \epsilon^1 \wedge \epsilon^2 \\ &= \left( \frac{h_2 h_3}{h_1} \frac{\partial f}{\partial q_1} \right) dq_2 \wedge dq_3 + \left( \frac{h_1 h_3}{h_2} \frac{\partial f}{\partial q_2} \right) dq_3 \wedge dq_1 + \left( \frac{h_1 h_2}{h_3} \frac{\partial f}{\partial q_3} \right) dq_1 \wedge dq_2. \end{aligned}$$

Differentiating once more, we get

$$\begin{aligned} d * df &= \frac{\partial}{\partial q_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial f}{\partial q_1} \right) dq_1 \wedge dq_2 \wedge dq_3 + \frac{\partial}{\partial q_2} \left( \frac{h_1 h_3}{h_2} \frac{\partial f}{\partial q_2} \right) dq_2 \wedge dq_3 \wedge dq_1 \\ &\quad + \frac{\partial}{\partial q_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial f}{\partial q_3} \right) dq_3 \wedge dq_1 \wedge dq_2 \\ &= \left\{ \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial q_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial f}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left( \frac{h_1 h_3}{h_2} \frac{\partial f}{\partial q_2} \right) \right. \right. \\ &\quad \left. \left. + \frac{\partial}{\partial q_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial f}{\partial q_3} \right) \right] \right\} \epsilon^1 \wedge \epsilon^2 \wedge \epsilon^3. \end{aligned} \tag{28.30}$$

Since  $\{\epsilon^1, \epsilon^2, \epsilon^3\}$  are orthonormal one-forms (as are  $\{dx, dy, dz\}$ ), the volume elements  $\epsilon^1 \wedge \epsilon^2 \wedge \epsilon^3$  and  $dx \wedge dy \wedge dz$  are equal. Thus, we substitute the latter for the former in (28.30), compare with (28.29), and conclude that

$$\nabla^2 f = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial q_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial f}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left( \frac{h_1 h_3}{h_2} \frac{\partial f}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial f}{\partial q_3} \right) \right],$$

which is the result obtained in curvilinear vector analysis. ■

**28.2.5. Example.** Let  $M = \mathbb{R}^2$ , and suppose that the arc length is given by  $ds^2 = (dx^2 + dy^2)/y^2$ . We can write the metric as  $\mathbf{g} = \epsilon^1 \otimes \epsilon^1 + \epsilon^2 \otimes \epsilon^2$  if we define

$$\epsilon^1 = \frac{dx}{y} \quad \text{and} \quad \epsilon^2 = \frac{dy}{y}.$$

The dual vectors  $\{\epsilon^1, \epsilon^2\}$  are orthonormal one-forms and  $G = 1$ , so we need not worry about raising and lowering indices. Inspection of the definition of  $\epsilon^1$  and  $\epsilon^2$ , along with the fact that  $dx(\partial_x) = dy(\partial_y) = 1$  and  $dx(\partial_y) = dy(\partial_x) = 0$ , immediately gives  $\mathbf{e}_1 = y\partial_x$  and  $\mathbf{e}_2 = y\partial_y$ .

To find the curvature tensor, we take the exterior derivative of the  $\epsilon^i$ 's:

$$d\epsilon^1 = d\left(\frac{1}{y}dx\right) = -\frac{1}{y^2}dy \wedge dx = \epsilon^1 \wedge \epsilon^2, \quad d\epsilon^2 = 0. \tag{28.31}$$

From these equations, the antisymmetry of the  $\omega$ 's, and Equation (28.14), we can read off  $\omega_{ij}$ . They are  $\omega_{11} = \omega_{22} = 0$  and  $\omega_{12} = -\omega_{21} = -\epsilon^1$ . Thus, the matrix  $\Omega$  is

$$\Omega = \begin{pmatrix} 0 & \omega_{12} \\ -\omega_{12} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\epsilon^1 \\ \epsilon^1 & 0 \end{pmatrix},$$

which gives

$$d\Omega = \begin{pmatrix} 0 & -d\epsilon^1 \\ d\epsilon^1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\epsilon^1 \wedge \epsilon^2 \\ \epsilon^1 \wedge \epsilon^2 & 0 \end{pmatrix},$$

and

$$\Omega \wedge \Omega = \begin{pmatrix} 0 & -\epsilon^1 \\ \epsilon^1 & 0 \end{pmatrix} \wedge \begin{pmatrix} 0 & -\epsilon^1 \\ \epsilon^1 & 0 \end{pmatrix} = 0.$$

Therefore, the curvature matrix is

$$\Theta = d\Omega = \begin{pmatrix} 0 & -\epsilon^1 \wedge \epsilon^2 \\ \epsilon^1 \wedge \epsilon^2 & 0 \end{pmatrix} \equiv \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix}.$$

This shows that the only nonzero independent component of the Riemann curvature tensor is  $R_{1212} = -1$ .  $\blacksquare$

**28.2.6. Example.** For a spherical surface of radius  $a$ , the element of length is

$$ds^2 = a^2 d\theta^2 + a^2 \sin^2 \theta d\varphi^2.$$

The orthonormal forms are  $\epsilon^\theta = a d\theta$  and  $\epsilon^\varphi = a \sin \theta d\varphi$ , and we have

$$G = 1, \quad d\epsilon^\theta = 0, \quad d\epsilon^\varphi = a \cos \theta d\theta \wedge d\varphi = \frac{1}{a} \cot \theta \epsilon^\theta \wedge \epsilon^\varphi.$$

The matrix  $\Omega$  can now be read off:<sup>5</sup>

$$\Omega = \begin{pmatrix} 0 & -\frac{\cot \theta}{a} \epsilon^\varphi \\ \frac{\cot \theta}{a} \epsilon^\varphi & 0 \end{pmatrix}.$$

A straightforward exterior differentiation yields

$$d\Omega = \begin{pmatrix} 0 & \frac{1}{a^2} \epsilon^\theta \wedge \epsilon^\varphi \\ -\frac{1}{a^2} \epsilon^\theta \wedge \epsilon^\varphi & 0 \end{pmatrix} = \frac{1}{a^2} \begin{pmatrix} 0 & \epsilon^\theta \wedge \epsilon^\varphi \\ -\epsilon^\theta \wedge \epsilon^\varphi & 0 \end{pmatrix}.$$

Similarly,  $\Omega \wedge \Omega = 0$ . Therefore, the curvature matrix is

$$\Theta = d\Omega = \frac{1}{a^2} \begin{pmatrix} 0 & \epsilon^\theta \wedge \epsilon^\varphi \\ -\epsilon^\theta \wedge \epsilon^\varphi & 0 \end{pmatrix}$$

The only independent component of the Riemann curvature tensor is  $R_{\theta\varphi\theta\varphi} = 1/a^2$ , which is constant, as expected for a spherical surface.  $\blacksquare$



What is a flat manifold?

It is clear that when the  $g_{ij}$  in the expression for the line element are all constants for all points in the manifold, then  $\epsilon^i$  will be proportional to  $dx^i$  and  $d\epsilon^i = 0$ , for all  $i$ . This immediately tells us that  $\Omega = 0$ , and therefore  $\Theta = 0$ ; that is, the manifold has no curvature. We call such a manifold **flat**. Thus, for  $ds^2 = dx^2 + dy^2 + dz^2$ , the space is flat. However, arc lengths of a flat space come in various guises with nontrivial coefficients. Does the curvature matrix  $\Theta$  recognize the flat arc length, or is it possible to fool it into believing that it is privileged with a curvature when in reality the curvature is still zero? The following example shows that the curvature matrix can detect flatness no matter how disguised the line element is!

**28.2.7. Example.** In spherical coordinates, the line element (arc length) of the flat Euclidean space  $\mathbb{R}^3$  is  $ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2$ . To calculate the curvature matrix, we first need an orthonormal set of one-forms. These are immediately obtained from the expression above:

$$\epsilon^r = dr, \quad \epsilon^\theta = r d\theta, \quad \epsilon^\varphi = r \sin \theta d\varphi, \quad G = 1.$$

Taking the exterior derivatives of these one-forms, we obtain

$$\begin{aligned} d\epsilon^r &= d^2 r = 0, \\ d\epsilon^\theta &= d(r d\theta) = dr \wedge d\theta + r \underbrace{d^2 \theta}_{=0} = \epsilon^r \wedge \left(\frac{\epsilon^\theta}{r}\right) = \frac{1}{r} \epsilon^r \wedge \epsilon^\theta, \\ d\epsilon^\varphi &= d(r \sin \theta) \wedge d\varphi = \sin \theta dr \wedge d\varphi + r \cos \theta d\theta \wedge d\varphi \\ &= \sin \theta \epsilon^r \wedge \left(\frac{\epsilon^\varphi}{r \sin \theta}\right) + r \cos \theta \left(\frac{\epsilon^\theta}{r}\right) \wedge \left(\frac{\epsilon^\varphi}{r \sin \theta}\right) \\ &= \frac{1}{r} \epsilon^r \wedge \epsilon^\varphi + \frac{\cot \theta}{r} \epsilon^\theta \wedge \epsilon^\varphi. \end{aligned}$$

We can now use Equation (28.14) to find the matrix of one-forms  $\Omega$ . In calculating the elements of  $\Omega$ , we remember that it is a skew-symmetric matrix, so all diagonal elements are zero. We also note that  $d\epsilon^k = 0$  does not imply that  $\omega^k_j = 0$ . Keeping these facts in mind, we can easily obtain  $\Omega$  (the calculation is left as a problem for the reader):

$$\Omega = \begin{pmatrix} 0 & -\frac{1}{r} \epsilon^\theta & -\frac{1}{r \sin \theta} \epsilon^\varphi \\ \frac{1}{r} \epsilon^\theta & 0 & -\frac{\cot \theta}{r} \epsilon^\varphi \\ \frac{1}{r \sin \theta} \epsilon^\varphi & \frac{\cot \theta}{r} \epsilon^\varphi & 0 \end{pmatrix}.$$

The exterior derivative of this matrix is found to be

$$d\Omega = \begin{pmatrix} 0 & 0 & -\frac{\cot \theta}{r^2} \epsilon^\theta \wedge \epsilon^\varphi \\ 0 & 0 & \frac{1}{r^2} \epsilon^\theta \wedge \epsilon^\varphi \\ \frac{\cot \theta}{r^2} \epsilon^\theta \wedge \epsilon^\varphi & -\frac{1}{r^2} \epsilon^\theta \wedge \epsilon^\varphi & 0 \end{pmatrix}.$$

<sup>5</sup>Note that  $d\epsilon^\theta = 0$  does not imply  $\omega_{12} = 0$ .

which is precisely (the negative of) the exterior product  $\Omega \wedge \Omega$ , as the reader may wish to verify. Thus,  $\Theta = d\Omega + \Omega \wedge \Omega = 0$ , and the space is indeed flat!  $\blacksquare$

In all the foregoing examples, the curvature was calculated intrinsically. We never had to leave the space and go to a higher dimension to “see” the curvature. For example, in the case of the sphere, the only information we had was the line element in terms of the coordinates on the sphere. We never had to resort to any three-dimensional analysis to discover a globe embedded in the Euclidean  $\mathbb{R}^3$ . As mentioned earlier, if a space has line elements with constant  $g_{ij}$ , then the Riemann curvature vanishes trivially. We have also seen examples in which the components of a metric tensor were by no means trivial, but  $\Theta$  was smart enough to detect the flatness in disguise. Under what conditions can we choose coordinate systems in terms of which the line elements have  $g_{ij} = \pm\delta_{ij}$ ? To answer this question we need the following lemma (proved in [Flan 89, pp. 135–136]):

**28.2.8. Lemma.** *If  $\Omega$  is a matrix of 1-forms such that  $d\Omega + \Omega \wedge (\mathcal{G}\Omega) = 0$ , then there exists an orthogonal matrix  $A$  such that  $dA = A\mathcal{G}\Omega$ .*

The question raised above is intimately related to the connection between coordinate and orthonormal frames. We have seen the usefulness of both. Coordinate frames, due to the existence of the related coordinate *functions*, are useful for many analytical calculations, for example in Hamiltonian dynamics. Orthonormal frames are useful because of the simplicity of expressions inherent in all orthonormal vectors. Furthermore, we saw how curvature was easily calculated once we constructed orthonormal dual frames. Naturally, we would like to have both. Is it possible to construct frames that are both coordinate and orthonormal? The following theorem answers this question:

**28.2.9. Theorem.** *Let  $M$  be a Riemannian manifold. Then  $M$  is flat, i.e.,  $\Theta = 0$  if and only if there exists a local coordinate system  $\{x^i\}$  for which  $\{\partial_i\}$  is an orthonormal basis.*

*Proof.* The existence of orthonormal coordinate frames implies that  $\{dx^i\}$  are orthonormal. Thus, we can use them to find the curvature. But since  $d(dx^i) = 0$  for all  $i$ , it follows from Equation (28.12) that  $\omega^i_j = 0$  and  $\Omega = 0$ . So the curvature must vanish. Conversely, suppose that  $\Theta = 0$ . Then by Lemma 28.2.8, there exists an orthogonal matrix  $A$  such that  $dA = A\mathcal{G}\Omega$ . Now we define the one-form column matrix  $\tau$  by  $\tau = A\varepsilon$ , where  $\varepsilon$  is the one-form column matrix of Equation (28.14). Then, using (28.19), we have

$$d\tau = d(A\varepsilon) = dA \wedge \varepsilon + Ad\varepsilon = (A\mathcal{G}\Omega) \wedge \varepsilon - A(\mathcal{G}\Omega \wedge \varepsilon) = 0.$$

Thus,  $d\tau^i = 0$  for all  $i$ . By Theorem 26.5.14 there must exist zero-forms (functions)  $x^i$  such that  $\tau^i = dx^i$ . These  $x^i$  are the coordinates we are after. The basis  $\{\partial_i\}$  is obtained using the inverse of  $A$  (see the discussion following Proposition 25.1.1). Since  $A$  is orthogonal, both  $\{dx^i\}$  and  $\{\partial_i\}$  are orthonormal bases.  $\square$

## 28.3 Covariant Derivative and Geodesics

The essence of all geometries are straight lines. The familiar Euclidean geometry is developed entirely based on a number of postulates concerning certain attributes and properties for straight lines. From the physical standpoint, straight lines are those trajectories on which “free” particles—including light—travel. If geometry is the basis of physical theories (general theory of relativity, and, to a lesser degree, electromagnetism and the nuclear interactions), knowledge of straight lines will be crucial.

### 28.3.1 Covariant Derivative

Straight lines are characterized by the “least amount of bending.” In flat space, this involves zero bending; but if space is curved, the bending cannot be eliminated. The bending of space is gauged by a test vector as it moves (infinitesimally) along some trajectory. The infinitesimal character of any trajectory is encapsulated in the vector tangent to it at the point of interest. Thus the concept of straight line is tied to the way one vector (the test vector) changes along a second vector (the tangent vector).

covariant derivative  
of a vector

**28.3.1. Definition.** Let  $\mathbf{u}, \mathbf{v} \in \mathcal{X}(M)$  be vector fields on  $M$ . The *covariant derivative* of  $\mathbf{v}$  with respect to (or along)  $\mathbf{u}$ , denoted by  $\nabla_{\mathbf{u}}\mathbf{v}$ , is defined as

$$\nabla_{\mathbf{u}}\mathbf{v} = \langle d\mathbf{v}, \mathbf{u} \rangle. \tag{28.32}$$

So,  $\nabla_{\mathbf{u}} : \mathcal{X}(M) \rightarrow \mathcal{X}(M)$ , i.e.,  $\nabla_{\mathbf{u}}$  maps vector fields to vector fields.

Certain properties of the covariant derivative follow from this definition. In most books on differential geometry and relativity, these properties are used to *define* the covariant derivative. We collect these properties in the following theorem.

Properties of  
covariant derivative

**28.3.2. Theorem.** Let  $\mathbf{u}, \mathbf{v}$ , and  $\mathbf{w}$  be vector fields, and  $f$  and  $h$  functions on a manifold  $M$ . Then the covariant derivative has the following properties:

1.  $\nabla_{\mathbf{u}}\mathbf{v} - \nabla_{\mathbf{v}}\mathbf{u} = [\mathbf{u}, \mathbf{v}]$ .
2.  $\nabla_{\mathbf{u}}(f\mathbf{v}) = \mathbf{u}(f)\mathbf{v} + f\nabla_{\mathbf{u}}\mathbf{v}$ .
3.  $\nabla_{\mathbf{u}}(\mathbf{v}+\mathbf{w}) = \nabla_{\mathbf{u}}\mathbf{v} + \nabla_{\mathbf{u}}\mathbf{w}$ .
4.  $\nabla_{f\mathbf{u}+h\mathbf{w}}(\mathbf{v}) = f\nabla_{\mathbf{u}}\mathbf{v} + h\nabla_{\mathbf{w}}\mathbf{v}$ .

*Proof.* We shall prove the first property, which happens to be the hardest. The rest are easy consequences of the definition and the linearity of pairing. To prove the first property, choose a basis<sup>6</sup>  $\{\mathbf{e}_i\}$  and its dual  $\{\epsilon^j\}$  and write  $\mathbf{v} = v^i\mathbf{e}_i$ ,  $\mathbf{u} = u^i\mathbf{e}_i$ ,

<sup>6</sup>Eventually, we shall take the basis to be a coordinate frame. But for notational convenience, we first work in a general basis.

and

$$dv = dv^i e_i + v^i de_i, \quad du = du^i e_i + u^i de_i.$$

Then we have

$$\begin{aligned} \langle dv, u \rangle &= \langle dv^i, u \rangle e_i + v^i \langle de_i, u \rangle = [u(v^i)]e_i + v^i u^j \langle de_i, e_j \rangle, \\ \langle du, v \rangle &= \langle du^i, v \rangle e_i + u^i \langle de_i, v \rangle = [v(u^i)]e_i + u^i v^j \langle de_i, e_j \rangle, \end{aligned}$$

and

$$\langle dv, u \rangle - \langle du, v \rangle = [u(v^i) - v(u^i)]e_i + v^i u^j (\langle de_i, e_j \rangle - \langle de_j, e_i \rangle). \quad (28.33)$$

Let us evaluate the term in parentheses. Using Equation (28.8), we have

$$\begin{aligned} \langle de_i, e_j \rangle - \langle de_j, e_i \rangle &= \langle e_k \otimes \omega^k_i, e_j \rangle - \langle e_k \otimes \omega^k_j, e_i \rangle \\ &= e_k \langle \omega^k_i, e_j \rangle - e_k \langle \omega^k_j, e_i \rangle = e_k (\Gamma^k_{ij} - \Gamma^k_{ji}). \end{aligned}$$

The basis chosen above was general. However, we are free to choose any convenient basis to prove a vector or tensor identity. If we choose the basis to be a coordinate frame, then by Equation (28.16),  $\Gamma^k_{ij} - \Gamma^k_{ji} = 0$ . Furthermore,

$$[u(v^i) - v(u^i)]e_i = [u, v] \quad \text{in a coordinate frame.}$$

Substituting these two relations in Equation (28.33), we obtain the first part of the theorem.  $\square$

**Tullio Levi-Civita** (1873–1941), the son of Giacomo Levi-Civita, a lawyer who from 1908 was a senator, was an outstanding student at the liceo in Padua. In 1890 he enrolled in the Faculty of Mathematics of the University of Padua. Giuseppe Veronese and Gregorio Ricci-Curbastro were among his teachers. He received his diploma in 1894 and in 1895 became resident professor at the teachers' college annexed to the Faculty of Science at Pavia. From 1897 to 1918 Levi-Civita taught rational mechanics at the University of Padua. His years in Padua (where in 1914 he married a pupil, Libera Trevisani) were scientifically the most fruitful of his career. In 1918 he became professor of higher analysis at Rome and, in 1920, of rational mechanics. In 1938, struck by the fascist racial laws against Jews, he was forced to give up teaching.



The breadth of his scientific interests, his scruples regarding the fulfillment of his academic responsibilities, and his affection for young people made Levi-Civita the leader of a flourishing school of mathematicians.

Levi-Civita's approximately 200 memoirs in pure and applied mathematics deal with analytical mechanics, celestial mechanics, hydrodynamics, elasticity, electromagnetism, and atomic physics. His most important contribution to science was rooted in the memoir "Sulle trasformazioni delle equazioni dinamiche" (1896), which was characterized by the use of the methods of absolute differential calculus that Ricci had applied only to differential geometry. In the "Méthodes de calcul différentiel absolu et leurs applications," written with Ricci and published in 1900 in *Mathematische Annalen*, there is a complete exposition of the new calculus, which consists of a particular algorithm designed to express geometric and physical laws in Euclidean and non-Euclidean spaces, particularly in Riemannian curved spaces. The memoir concerns a very general but laborious type of calculus that made it possible to deal with many difficult problems, including, according to Einstein, the formulation of the general theory of relativity.

Although Levi-Civita had expressed certain reservations concerning relativity in the first years after its formulation (1905), he gradually came to accept the new views. His own original research culminated in 1917 in the introduction of the concept of *parallel transport* in curved spaces. With this new concept, absolute differential calculus, having absorbed other techniques, became tensor calculus, now the essential instrument of the unitary relativistic theories of gravitation and electromagnetism.

In his memoirs of 1903–1916 Levi-Civita contributed to celestial mechanics in the study of the three-body problem: the determination of the motion of three bodies, considered as reduced to their centers of mass and subject to mutual Newtonian attraction. In 1914–1916 he succeeded in eliminating the singularities present at the points of possible collisions, past or future. His research in relativity led Levi-Civita to mathematical problems suggested by atomic physics, which in the 1920s was developing outside the traditional framework: the general theory of adiabatic invariants, the motion of a body of variable mass, the extension of the Maxwellian distribution to a system of corpuscles, and the Schrödinger equation.

The covariant differentiation of vectors can be extended to arbitrary tensors once it is defined for 1-forms. For this, we make the extra assumption that  $\nabla_{\mathbf{u}}$  can be "pushed" inside a pairing, and that when acting on a tensor product, it obeys the product rule of differentiation, i.e.,  $\nabla_{\mathbf{u}}$  acts as a derivation on the algebra of tensors. Let  $\omega$  be a 1-form and  $\mathbf{v}$  a vector. Then

$$\nabla_{\mathbf{u}} \langle \omega, \mathbf{v} \rangle = \langle \nabla_{\mathbf{u}} \omega, \mathbf{v} \rangle + \langle \omega, \nabla_{\mathbf{u}} \mathbf{v} \rangle \quad (28.34)$$

defines the covariant derivative of  $\omega$ . Using Equation (28.34) the reader may readily show the primary covariant derivative relation:

$$\nabla_{\mathbf{e}_i} \epsilon^j = -\Gamma_{ki}^j \epsilon^k. \quad (28.35)$$

Since an arbitrary tensor of a given kind can be expressed as a linear combination of tensor product of vectors and 1-forms, our knowledge of the action of  $\nabla_{\mathbf{u}}$  on functions (coefficients of expansion), vectors, and 1-forms, plus the assumed derivation property of  $\nabla_{\mathbf{u}}$ , is enough to uniquely define the action of  $\nabla_{\mathbf{u}}$  on any tensor. Equation (28.34) shows that the covariant derivative of a 1-form with respect

to a vector is another 1-form, because when it pairs up with a vector it gives a number. We have already pointed out that the covariant derivative of a vector with respect to another vector is a third vector. We therefore conclude that

**28.3.3. Box.** *The covariant derivative of a tensor of a given kind is another tensor of the same kind.*

**28.3.4. Example.** Using the definition of the covariant derivative, the reader may check that

$$\nabla_{\mathbf{e}_i} \mathbf{e}_j = \mathbf{e}_k \Gamma^k_{ji}. \quad \text{Note the change in order of indices!} \quad (28.36)$$

Now consider two bases  $\{\mathbf{e}_i\}$  and  $\{\mathbf{e}_{i'}\}$ . Write the primed basis in terms of the other:  $\mathbf{e}_{i'} = R^j_{i'} \mathbf{e}_j$ . Then

$$\begin{aligned} \mathbf{e}_{k'} \Gamma^{k'}_{i'j'} &\equiv \nabla_{\mathbf{e}_{j'}} \mathbf{e}_{i'} = \nabla_{R^l_{j'} \mathbf{e}_l} (R^j_{i'} \mathbf{e}_j) = R^l_{j'} \nabla_{\mathbf{e}_l} (R^j_{i'} \mathbf{e}_j) \\ &= R^l_{j'} \{ R^j_{i'} \nabla_{\mathbf{e}_l} (\mathbf{e}_j) + \nabla_{\mathbf{e}_l} (R^j_{i'}) \mathbf{e}_j \} \\ &= R^l_{j'} R^j_{i'} \mathbf{e}_m \Gamma^m_{jl} + R^l_{j'} \underbrace{\mathbf{e}_l (R^m_{i'})}_{\equiv R^m_{i',l}} \mathbf{e}_m. \end{aligned}$$

Connection coefficients are not tensors!

Writing  $\mathbf{e}_{k'} = R^m_{k'} \mathbf{e}_m$  on the LHS, equating the components on both sides, and multiplying both sides by the inverse of the transformation matrix  $R$ , we obtain

$$\Gamma^{k'}_{i'j'} = \underbrace{R^{k'}_m R^l_{j'} R^j_{i'} \Gamma^m_{jl}}_{\text{how a (1, 2)-tensor transforms}} + \underbrace{R^{k'}_m R^l_{j'} R^m_{i',l}}_{\text{nontensorial term}}, \quad (28.37)$$

where  $R^{k'}_m \equiv (R^{-1})^{k'}_m$ . Equation (28.37) shows that the connection coefficients *are not tensors*. ■

**28.3.5. Example.** Equation (28.36) connects  $\Gamma^i_{jk}$  with the structure constants of the Lie algebra of vector fields on a manifold. To see this connection, use the first property of the covariant derivative in Theorem 28.3.2 and Equation (28.36) to obtain

$$[\mathbf{e}_i, \mathbf{e}_j] = \nabla_{\mathbf{e}_i} \mathbf{e}_j - \nabla_{\mathbf{e}_j} \mathbf{e}_i = \mathbf{e}_k (\Gamma^k_{ji} - \Gamma^k_{ij}).$$

It follows from this equation that

$$c^k_{ij} = \Gamma^k_{ji} - \Gamma^k_{ij}. \quad (28.38)$$

In particular, in a coordinate frame,  $c^k_{ij} = 0$  and the connection coefficients are symmetric in their lower indices, a result we obtained earlier. ■

gradient operator for tensors

Let  $\mathbf{T}$  be a tensor of type  $(r, s)$ . It is convenient to think of  $\nabla_{\mathbf{u}}\mathbf{T}$  as the contraction of another tensor  $\mathbf{S} \equiv \nabla\mathbf{T}$  with  $\mathbf{u}$ . Then  $\nabla$  may naturally be treated as a linear operator that maps  $T_s^r(M)$ , the bundle of tensor fields of type  $(r, s)$ , to  $T_{s+1}^r(M)$ , the bundle of tensor fields of type  $(r, s + 1)$ . One then writes  $\nabla : T_s^r(M) \rightarrow T_{s+1}^r(M)$  and calls  $\nabla$  the generalized gradient operator. If

$$\mathbf{T} = T_{j_1 \dots j_s}^{i_1 \dots i_r} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s},$$

then

$$\nabla\mathbf{T} \equiv T_{j_1 \dots j_s; k}^{i_1 \dots i_r} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s} \otimes \epsilon^k, \tag{28.39}$$

and, with  $\mathbf{u} = u^k \mathbf{e}_k$ ,

$$\nabla_{\mathbf{u}}\mathbf{T} = T_{j_1 \dots j_s; k}^{i_1 \dots i_r} u^k \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s}. \tag{28.40}$$

Using these relations, we can calculate the components of the covariant derivative of a general tensor. It is clear that if we use  $\mathbf{e}_k$  instead of  $\mathbf{u}$ , we obtain the  $k$ th component of the covariant derivative. So, on the one hand, we have

$$\nabla_{\mathbf{e}_k}\mathbf{T} = T_{j_1 \dots j_s; k}^{i_1 \dots i_r} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s}, \tag{28.41}$$

and on the other hand,

$$\begin{aligned} \nabla_{\mathbf{e}_k}\mathbf{T} &= \nabla_{\mathbf{e}_k} (T_{j_1 \dots j_s}^{i_1 \dots i_r} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s}) \\ &= T_{j_1 \dots j_s; k}^{i_1 \dots i_r} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s} \\ &\quad + T_{j_1 \dots j_s}^{i_1 \dots i_r} \sum_{m=1}^r \mathbf{e}_{i_1} \otimes \dots \otimes \underbrace{\nabla_{\mathbf{e}_k} \mathbf{e}_{i_m}}_{\mathbf{e}_n \Gamma_{imk}^n} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \epsilon^{j_s} \\ &\quad + T_{j_1 \dots j_s}^{i_1 \dots i_r} \sum_{m=1}^s \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \epsilon^{j_1} \otimes \dots \otimes \underbrace{\nabla_{\mathbf{e}_k} \epsilon^{j_m}}_{-\Gamma_{nk}^{j_m} \epsilon^n} \otimes \dots \otimes \epsilon^{j_s}, \end{aligned} \tag{28.42}$$

where  $T_{j_1 \dots j_s; k}^{i_1 \dots i_r} \equiv \mathbf{e}_k (T_{j_1 \dots j_s}^{i_1 \dots i_r})$ . Equating the components of Equations (28.41) and (28.42) yields

$$\begin{aligned} T_{j_1 \dots j_s; k}^{i_1 \dots i_r} &= T_{j_1 \dots j_s; k}^{i_1 \dots i_r} + \sum_{m=1}^r T_{j_1 \dots j_{m-1} j_{m+1} \dots j_s}^{i_1 \dots i_{m-1} i_{m+1} \dots i_r} \Gamma_{nk}^{i_m} \\ &\quad - \sum_{m=1}^s T_{j_1 \dots j_{m-1} j_{m+1} \dots j_s}^{i_1 \dots i_{m-1} i_{m+1} \dots i_r} \Gamma_{jmk}^n, \end{aligned} \tag{28.43}$$

where only the sum over the subindex  $m$  has been explicitly displayed; the (hidden) sum over repeated indices is, as always, understood.

There is a useful relation between the covariant and the Lie derivative that we derive now. First, let  $\mathbf{T}$  be of type  $(2, 0)$  and write it in some frame as  $\mathbf{T} = T^{ij} \mathbf{e}_i \otimes \mathbf{e}_j$ . Apply the covariant derivative with respect to  $\mathbf{u}$  to both sides to obtain

$$\nabla_{\mathbf{u}} \mathbf{T} = \mathbf{u}(T^{ij}) \mathbf{e}_i \otimes \mathbf{e}_j + T^{ij} (\nabla_{\mathbf{u}} \mathbf{e}_i) \otimes \mathbf{e}_j + T^{ij} \mathbf{e}_i \otimes (\nabla_{\mathbf{u}} \mathbf{e}_j).$$

derivation of the relation between the Lie and the covariant derivatives

Similarly,

$$L_{\mathbf{u}} \mathbf{T} = \mathbf{u}(T^{ij}) \mathbf{e}_i \otimes \mathbf{e}_j + T^{ij} (L_{\mathbf{u}} \mathbf{e}_i) \otimes \mathbf{e}_j + T^{ij} \mathbf{e}_i \otimes (L_{\mathbf{u}} \mathbf{e}_j).$$

Now use  $L_{\mathbf{u}} \mathbf{e}_j = [\mathbf{u}, \mathbf{e}_j] = \nabla_{\mathbf{u}} \mathbf{e}_j - \nabla_{\mathbf{e}_j} \mathbf{u}$  to get

$$L_{\mathbf{u}} \mathbf{T} = \nabla_{\mathbf{u}} \mathbf{T} - T^{ij} [(\nabla_{\mathbf{e}_i} \mathbf{u}) \otimes \mathbf{e}_j + \mathbf{e}_i \otimes (\nabla_{\mathbf{e}_j} \mathbf{u})]. \tag{28.44}$$

On the other hand, if we apply  $\nabla_{\mathbf{u}}$  and  $L_{\mathbf{u}}$  to both sides of  $\delta_j^i = \langle \mathbf{e}^i, \mathbf{e}_j \rangle$  and use  $[\mathbf{u}, \mathbf{e}_i] = \nabla_{\mathbf{u}} \mathbf{e}_i - \nabla_{\mathbf{e}_i} \mathbf{u}$ , we obtain

$$\nabla_{\mathbf{u}} \mathbf{e}^i = L_{\mathbf{u}} \mathbf{e}^i - (\nabla_{\mathbf{e}_j} \mathbf{u})^i \mathbf{e}^j.$$

It follows that for  $\mathbf{T} = T_{ij} \mathbf{e}^i \otimes \mathbf{e}^j$ , we have

$$L_{\mathbf{u}} \mathbf{T} = \nabla_{\mathbf{u}} \mathbf{T} + T_{ij} [(\nabla_{\mathbf{e}_k} \mathbf{u})^i \mathbf{e}^k \otimes \mathbf{e}^j + (\nabla_{\mathbf{e}_k} \mathbf{u})^j \mathbf{e}^i \otimes \mathbf{e}^k]. \tag{28.45}$$

One can use Equations (28.44) and (28.45) to generalize to a tensor of type  $(r, s)$ . If  $\mathbf{u}$  happens to be tangent to a curve  $t \mapsto \gamma(t)$ , Equation (28.40) is written as

$$\nabla_{\mathbf{u}} \mathbf{T} = \frac{DT_{j_1 \dots j_s}^{i_1 \dots i_r}}{dt} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_r} \otimes \mathbf{e}^{j_1} \otimes \dots \otimes \mathbf{e}^{j_s}, \tag{28.46}$$

where  $DT_{j_1 \dots j_s}^{i_1 \dots i_r} / dt \equiv T_{j_1 \dots j_s; k}^{i_1 \dots i_r} u^k$ . In a coordinate frame, with  $u^i = \dot{x}^i = dx^i / dt$ , Equations (28.40) and (28.43) give

$$\begin{aligned} \frac{DT_{j_1 \dots j_s}^{i_1 \dots i_r}}{dt} &= T_{j_1 \dots j_s; k}^{i_1 \dots i_r} \frac{dx^k}{dt} = \frac{dT_{j_1 \dots j_s}^{i_1 \dots i_r}}{dt} + \sum_{m=1}^r T_{j_1 \dots j_{m-1} j_{m+1} \dots j_s}^{i_1 \dots i_{m-1} i_{m+1} \dots i_r} \Gamma_{nk}^{i_m} \frac{dx^k}{dt} \\ &\quad - \sum_{m=1}^s T_{j_1 \dots j_{m-1} j_{m+1} \dots j_s}^{i_1 \dots i_{m-1} i_m i_{m+1} \dots i_r} \Gamma_{j_m k}^n \frac{dx^k}{dt} \end{aligned} \tag{28.47}$$

For the case of a vector (28.47) becomes

$$\boxed{\frac{Dv^k}{dt} = \frac{dv^k}{dt} + v^j \Gamma_{ji}^k \frac{dx^i}{dt}}. \tag{28.48}$$

This is an important equation, to which we shall return shortly.

With the generalized gradient operator defined, we can construct the divergence of a tensor just as in vector analysis. Given a vector, the divergence operator  $\nabla \cdot$  acts on it and gives a scalar, or, in the language of tensor analysis, it lowers the upper indices by 1. This takes place by differentiating components and contracting the upper index with the newly introduced index of differentiation. The divergence of an arbitrary tensor is defined in precisely the same way:



divergence of a  
tensor field

**28.3.6. Definition.** Given a tensor field  $\mathbf{T}$ , define its *divergence*  $\nabla \cdot \mathbf{T}$  to be the tensor obtained from  $\nabla \mathbf{T}$  by contracting the last upper index with the covariant derivative index. In components,

$$(\nabla \cdot \mathbf{T})_{j_1 \dots j_s}^{i_1 \dots i_{r-1}} = T_{j_1 \dots j_s; k}^{i_1 \dots i_{r-1} k}.$$

The covariant derivative replaces the ordinary derivative when functions are generalized to tensor fields, and in many respects it is very similar to ordinary derivatives. The aspect of the covariant derivative that is in contrast to the ordinary derivative, namely, its lack of commutativity, is related to an important geometrical object that we encountered before: curvature. Let us find this relation. Start with an ordinary 1-form  $\omega = \omega_i \epsilon^i$ , and note that

$$\begin{aligned} \langle d\omega, \mathbf{u} \wedge \mathbf{v} \rangle &= \langle d\omega_i \wedge \epsilon^i + \omega_i d\epsilon^i, \mathbf{u} \wedge \mathbf{v} \rangle \\ &= \langle d\omega_i \wedge \epsilon^i, \mathbf{u} \wedge \mathbf{v} \rangle + \langle \omega_i d\epsilon^i, \mathbf{u} \wedge \mathbf{v} \rangle, \end{aligned} \quad (28.49)$$

where  $\mathbf{u}$  and  $\mathbf{v}$  are arbitrary vectors. For the first term of Equation (28.49), we obtain

$$\langle d\omega_i \wedge \epsilon^i, \mathbf{u} \wedge \mathbf{v} \rangle = \langle d\omega_i, \mathbf{u} \rangle v^i - \langle d\omega_i, \mathbf{v} \rangle u^i = v^i \mathbf{u}(\omega_i) - u^i \mathbf{v}(\omega_i).$$

Now we use

$$\begin{aligned} \mathbf{u}(\langle \mathbf{v}, \omega \rangle) &= \mathbf{u}(v^i \omega_i) = \omega_i \mathbf{u}(v^i) + v^i \mathbf{u}(\omega_i), \\ \mathbf{v}(\langle \mathbf{u}, \omega \rangle) &= \mathbf{v}(u^i \omega_i) = \omega_i \mathbf{v}(u^i) + u^i \mathbf{v}(\omega_i), \end{aligned}$$

to get

$$\langle d\omega_i \wedge \epsilon^i, \mathbf{u} \wedge \mathbf{v} \rangle = \mathbf{u}(\langle \mathbf{v}, \omega \rangle) - \mathbf{v}(\langle \mathbf{u}, \omega \rangle) + \omega_i [\mathbf{v}(u^i) - \mathbf{u}(v^i)]. \quad (28.50)$$

With  $d\epsilon^i = -\omega^j \wedge \epsilon^j = -\Gamma^i_{jk} \epsilon^k \wedge \epsilon^j$ , the second term of (28.49) becomes

$$\begin{aligned} \langle \omega_i d\epsilon^i, \mathbf{u} \wedge \mathbf{v} \rangle &= -\omega_i \Gamma^i_{jk} \langle \epsilon^k \wedge \epsilon^j, \mathbf{u} \wedge \mathbf{v} \rangle = -\omega_i \Gamma^i_{jk} (u^k v^j - u^j v^k) \\ &= -\omega_i u^k v^j \underbrace{(\Gamma^i_{jk} - \Gamma^i_{kj})}_{\langle \epsilon^i, [\mathbf{e}_k, \mathbf{e}_j] \rangle} = -u^k v^j \langle \omega, [\mathbf{e}_k, \mathbf{e}_j] \rangle. \end{aligned}$$

We also note that

$$\begin{aligned} [\mathbf{u}, \mathbf{v}] &= [u^k \mathbf{e}_k, v^j \mathbf{e}_j] = u^k \mathbf{e}_k (v^j \mathbf{e}_j) - v^j \mathbf{e}_j (u^k \mathbf{e}_k) \\ &= [\mathbf{u}(v^j)] \mathbf{e}_j + u^k v^j \mathbf{e}_k \mathbf{e}_j - [\mathbf{v}(u^k)] \mathbf{e}_k - u^k v^j \mathbf{e}_j \mathbf{e}_k \\ &= [\mathbf{u}(v^j) - \mathbf{v}(u^j)] \mathbf{e}_j + u^k v^j [\mathbf{e}_k, \mathbf{e}_j]. \end{aligned}$$

It follows that

$$\langle \omega, [\mathbf{u}, \mathbf{v}] \rangle = \omega_j [\mathbf{u}(v^j) - \mathbf{v}(u^j)] + u^k v^j \langle \omega, [\mathbf{e}_k, \mathbf{e}_j] \rangle, \tag{28.51}$$

and the second term of (28.49) becomes

$$\langle \omega_i d\epsilon^i, \mathbf{u} \wedge \mathbf{v} \rangle = - \langle \omega, [\mathbf{u}, \mathbf{v}] \rangle + \omega_j [\mathbf{u}(v^j) - \mathbf{v}(u^j)]. \tag{28.52}$$

Combining Equations (28.49), (28.50), and (28.52), we obtain

$$\langle d\omega, \mathbf{u} \wedge \mathbf{v} \rangle = \mathbf{u}(\langle \mathbf{v}, \omega \rangle) - \mathbf{v}(\langle \mathbf{u}, \omega \rangle) - \langle \omega, [\mathbf{u}, \mathbf{v}] \rangle. \tag{28.53}$$

Equation (28.53) is the basis of our generalization involving the covariant derivative: We replace  $\omega$  with a vector-valued 1-form  $\mathbf{T}$ , and the derivatives  $\mathbf{u}(\dots)$  and  $\mathbf{v}(\dots)$  with  $\nabla_{\mathbf{u}}(\dots)$  and  $\nabla_{\mathbf{v}}(\dots)$ . The result will be

$$\langle d\mathbf{T}, \mathbf{u} \wedge \mathbf{v} \rangle = \nabla_{\mathbf{u}}(\langle \mathbf{v}, \mathbf{T} \rangle) - \nabla_{\mathbf{v}}(\langle \mathbf{u}, \mathbf{T} \rangle) - \langle \mathbf{T}, [\mathbf{u}, \mathbf{v}] \rangle. \tag{28.54}$$

Choose a vector  $\mathbf{w}$  and replace  $\mathbf{T}$  with  $d\mathbf{w}$  to obtain

$$\begin{aligned} \langle d^2\mathbf{w}, \mathbf{u} \wedge \mathbf{v} \rangle &= \nabla_{\mathbf{u}}(\langle \mathbf{v}, d\mathbf{w} \rangle) - \nabla_{\mathbf{v}}(\langle \mathbf{u}, d\mathbf{w} \rangle) - \langle d\mathbf{w}, [\mathbf{u}, \mathbf{v}] \rangle \\ &\quad \underbrace{\equiv \nabla_{\mathbf{v}}\mathbf{w}} \qquad \underbrace{\equiv \nabla_{\mathbf{u}}\mathbf{w}} \qquad \underbrace{\equiv \nabla_{[\mathbf{u}, \mathbf{v}]} \mathbf{w}} \\ &= \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\mathbf{w} - \nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\mathbf{w} - \nabla_{[\mathbf{u}, \mathbf{v}]} \mathbf{w} \equiv \mathbf{R}(\mathbf{u}, \mathbf{v})\mathbf{w}, \end{aligned} \tag{28.55}$$

where we have introduced the linear operator  $\mathbf{R}(\mathbf{u}, \mathbf{v}) : \mathcal{X}(M) \rightarrow \mathcal{X}(M)$  in the last line. We rewrite the definition of this operator as

$$\mathbf{R}(\mathbf{u}, \mathbf{v}) = \nabla_{\mathbf{u}}\nabla_{\mathbf{v}} - \nabla_{\mathbf{v}}\nabla_{\mathbf{u}} - \nabla_{[\mathbf{u}, \mathbf{v}]}. \tag{28.56}$$

Equations (28.55) and (28.56) can be used to derive the following identity (see Problem 28.11):

$$\mathbf{R}(\mathbf{u}, \mathbf{v})\mathbf{w} + \mathbf{R}(\mathbf{w}, \mathbf{u})\mathbf{v} + \mathbf{R}(\mathbf{v}, \mathbf{w})\mathbf{u} = 0. \tag{28.57}$$

Define the map  $\mathbf{R} : \mathcal{X}^*(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \rightarrow \mathbb{R}$  by

$$\mathbf{R}(\omega, \mathbf{w}, \mathbf{u}, \mathbf{v}) \equiv \langle \omega, \mathbf{R}(\mathbf{u}, \mathbf{v})\mathbf{w} \rangle. \tag{28.58}$$

The covariant derivatives in the definition of  $\mathbf{R}(\mathbf{u}, \mathbf{v})$  may give the impression that  $\mathbf{R}$  will differentiate whatever appears to its right. It is a remarkable property of  $\mathbf{R}$  that this will not happen: That  $\mathbf{R}$  is a *tensor* of type (1, 3) [see Box 26.4.11], the **Riemann curvature tensor**, follows from Equations (28.7) and (28.55). The reader may verify that the components of this tensor are precisely those introduced in Equation (28.22). In fact, it is now more appropriate to change  $\theta_{ij}$  to  $\mathbf{R}_{ij}$  and write

$$\mathbf{R}_{ij} \equiv \frac{1}{2} R_{ijkl} \epsilon^k \wedge \epsilon^l. \tag{28.59}$$

Riemann curvature "tensor" is indeed a tensor in the sense of Box 26.4.11!

From Equation (28.57) follows the cyclic property of the lower indices of the components of  $\mathbf{R}$  as given in Equation (28.24).

The components of  $\mathbf{R}$  can be conveniently evaluated in terms of connection coefficients:

$$\begin{aligned} R^i{}_{jkl} &\equiv \mathbf{R}(\epsilon^i, \mathbf{e}_j, \mathbf{e}_k, \mathbf{e}_l) = \langle \epsilon^i, \mathbf{R}(\mathbf{e}_k, \mathbf{e}_l)\mathbf{e}_j \rangle \\ &= \langle \epsilon^i, \nabla_{\mathbf{e}_k} \nabla_{\mathbf{e}_l} \mathbf{e}_j - \nabla_{\mathbf{e}_l} \nabla_{\mathbf{e}_k} \mathbf{e}_j - \nabla_{[\mathbf{e}_k, \mathbf{e}_l]} \mathbf{e}_j \rangle \\ &= \langle \epsilon^i, \nabla_{\mathbf{e}_k} \nabla_{\mathbf{e}_l} \mathbf{e}_j \rangle - \langle \epsilon^i, \nabla_{\mathbf{e}_l} \nabla_{\mathbf{e}_k} \mathbf{e}_j \rangle - \langle \epsilon^i, \nabla_{[\mathbf{e}_k, \mathbf{e}_l]} \mathbf{e}_j \rangle. \end{aligned} \quad (28.60)$$

We calculate each pairing separately. The vector in the first pairing is

$$\begin{aligned} \nabla_{\mathbf{e}_k} \nabla_{\mathbf{e}_l} \mathbf{e}_j &= \nabla_{\mathbf{e}_k} (\Gamma^m{}_{jl} \mathbf{e}_m) = [\mathbf{e}_k (\Gamma^m{}_{jl})] \mathbf{e}_m + \Gamma^m{}_{jl} \nabla_{\mathbf{e}_k} \mathbf{e}_m \\ &= \Gamma^m{}_{jl,k} \mathbf{e}_m + \Gamma^m{}_{jl} \Gamma^m{}_{mk} \mathbf{e}_n. \end{aligned}$$

Similarly, the second vector can be expressed as

$$\nabla_{\mathbf{e}_l} \nabla_{\mathbf{e}_k} \mathbf{e}_j = \Gamma^m{}_{jk,l} \mathbf{e}_m + \Gamma^m{}_{jk} \Gamma^m{}_{ml} \mathbf{e}_n.$$

For the third vector, we use the definition of the structure constants and write

$$\nabla_{[\mathbf{e}_k, \mathbf{e}_l]} \mathbf{e}_j = \nabla_{c^m{}_{kl} \mathbf{e}_m} \mathbf{e}_j = c^m{}_{kl} \nabla_{\mathbf{e}_m} \mathbf{e}_j = c^m{}_{kl} \Gamma^m{}_{jm} \mathbf{e}_n.$$

Substituting these in Equation (28.60), we obtain

$$R^i{}_{jkl} = \Gamma^i{}_{jl,k} - \Gamma^i{}_{jk,l} + \Gamma^m{}_{jl} \Gamma^i{}_{mk} - \Gamma^m{}_{jk} \Gamma^i{}_{ml} - c^m{}_{kl} \Gamma^i{}_{jm}. \quad (28.61)$$

If we use coordinate frames, the structure constants are zero, and we get

$$R^i{}_{jkl} = \frac{\partial \Gamma^i{}_{jl}}{\partial x^k} - \frac{\partial \Gamma^i{}_{jk}}{\partial x^l} + \Gamma^i{}_{mk} \Gamma^m{}_{jl} - \Gamma^i{}_{ml} \Gamma^m{}_{jk}. \quad (28.62)$$

If the manifold has a metric (Riemannian or pseudo-Riemannian manifold), then a combination of (28.26) and (28.62) gives the curvature in terms of the metric tensor.

### 28.3.2 Geodesics

The reader may recall from elementary physics courses that in a flat space, one is allowed<sup>7</sup> to move a vector about as long as it is kept parallel to itself. In any kind of parallel displacement of a vector, one moves the (tail of the) vector along some curve. This curve is in our subconscious in flat space, because it plays no role in such a displacement—all curves give the same end result. However, in curved spaces, different curves “parallel transport” a vector differently, so that the

<sup>7</sup>Except in situations where the position of the vector is important, as in torques and angular momenta.

end result *will* be different. Each curve honestly “thinks” that it is indeed parallel transporting the vector, i.e., that it is not changing the direction of the vector as the latter moves along the former. No change means zero derivative; and since the concept of derivative is local, let us replace the curve with its tangent  $\mathbf{u}$ :

parallel transport of vectors and tensors along a vector

**28.3.7. Definition.** A vector  $\mathbf{v}$  is said to be *parallel transported along*  $\mathbf{u}$  if  $\nabla_{\mathbf{u}}\mathbf{v} = 0$ . Similarly, a tensor  $\mathbf{T}$  is said to be *parallel transported along*  $\mathbf{u}$  if  $\nabla_{\mathbf{u}}\mathbf{T} = 0$ .

If the manifold  $M$  has a metric  $\mathbf{g}$ , then it is desirable for the parallel transportation not to affect the metric, i.e., that the angle between any two vectors remain the same after transportation. Therefore, we demand that  $\mathbf{g}$  be constant for parallel transportation along any vector, i.e.,

$$\nabla_{\mathbf{u}}\mathbf{g} = 0 \quad \forall \mathbf{u} \Rightarrow \nabla\mathbf{g} = 0, \quad \text{or} \quad g_{ij;k} = 0 \quad \forall i, j, k. \quad (28.63)$$

A statement equivalent to this equation is that

**28.3.8. Box.** *The operation of raising and lowering of indices commutes with the operation of covariant differentiation.*

Consider two vectors  $\mathbf{v}$  and  $\mathbf{w}$ . If  $\mathbf{g}$  satisfies Equation (28.63), then

$$\begin{aligned} \nabla_{\mathbf{u}}(\mathbf{v} \cdot \mathbf{w}) &= \nabla_{\mathbf{u}}[\mathbf{g}(\mathbf{v}, \mathbf{w})] \equiv \nabla_{\mathbf{u}}(\mathbf{g}, \mathbf{v} \otimes \mathbf{w}) \\ &= \langle \nabla_{\mathbf{u}}\mathbf{g}, \mathbf{v} \otimes \mathbf{w} \rangle + \langle \mathbf{g}, (\nabla_{\mathbf{u}}\mathbf{v}) \otimes \mathbf{w} \rangle + \langle \mathbf{g}, \mathbf{v} \otimes (\nabla_{\mathbf{u}}\mathbf{w}) \rangle \\ &= \mathbf{g}(\nabla_{\mathbf{u}}\mathbf{v}, \mathbf{w}) + \mathbf{g}(\mathbf{v}, \nabla_{\mathbf{u}}\mathbf{w}), \end{aligned}$$

or

$$\nabla_{\mathbf{u}}(\mathbf{v} \cdot \mathbf{w}) = (\nabla_{\mathbf{u}}\mathbf{v}) \cdot \mathbf{w} + \mathbf{v} \cdot (\nabla_{\mathbf{u}}\mathbf{w}). \quad (28.64)$$

In particular, if  $\mathbf{v}$  and  $\mathbf{w}$  are parallel transported along  $\mathbf{u}$ , their dot product will not change. In the flat space of a large sheet of paper, construction of a straight line in a given direction starting at a given point  $P_0$  is done by laying down the end of a vector (a straight edge) at  $P_0$  pointing in the given direction, connecting  $P_0$  to a neighboring point  $P_1$  along the vector, moving the vector *parallel to itself* to  $P_1$ , connecting  $P_1$  to a neighboring point  $P_2$ , and continuing the process. In the language of the machinery of the covariant derivative, we might say that a straight line is constructed by transporting the tangent vector parallel to itself.

geodesics defined

**28.3.9. Definition.** Let  $M$  be a manifold and  $\gamma : [a, b] \rightarrow M$  a curve. Then  $\gamma$  is called a *geodesic* of  $M$  if the tangent vector  $\mathbf{u}$  at every point of  $\gamma$  is parallel transported along itself:  $\nabla_{\mathbf{u}}\mathbf{u} = 0$ .

geodesic equation

It follows from Equation (28.48)—with  $v^k = u^k = dx^k/dt$ —that a geodesic

curve satisfies the following DE:

$$\frac{d^2 x^k}{dt^2} + \Gamma^k_{ji} \frac{dx^i}{dt} \frac{dx^j}{dt} = 0. \tag{28.65}$$

This second-order DE, called the **geodesic equation**, will have a unique solution if  $x^i(a)$  and  $\dot{x}^i(a)$ , i.e., the initial point and the initial direction, are given. Thus,

**28.3.10. Box.** *Through a given point and in a given direction passes only one geodesic curve.*

**28.3.11. Example.** Let us determine the geodesics of the space whose arc length is given by  $ds^2 = (dx^2 + dy^2)/y^2$  (see Example 28.2.5). With  $x = x^1$  and  $y = x^2$ , we recognize the metric tensor as

$$g_{11} = g_{22} = \frac{1}{y^2}, \quad g_{12} = g_{21} = 0, \\ g^{11} = g^{22} = y^2 \quad g^{12} = g^{21} = 0.$$

Using Equation (28.26), we can readily calculate the nonzero connection coefficients:

$$\Gamma_{112} = \Gamma_{121} = -\Gamma_{211} = \Gamma_{222} = -\frac{1}{y^3}.$$

The geodesic equation for the first coordinate is

$$\frac{d^2 x}{dt^2} + \Gamma^1_{ji} \frac{dx^i}{dt} \frac{dx^j}{dt} = 0,$$

or

$$\frac{d^2 x}{dt^2} + \Gamma^1_{11} \left(\frac{dx}{dt}\right)^2 + 2\Gamma^1_{12} \frac{dx}{dt} \frac{dy}{dt} + \Gamma^1_{22} \left(\frac{dy}{dt}\right)^2 = 0.$$

To find the connection coefficients with raised indices, we multiply those with all indices lowered by the inverse of the metric tensor. For instance,

$$\Gamma^1_{12} = g^{1i} \Gamma_{i12} = g^{11} \Gamma_{112} + \underbrace{g^{12}}_{=0} \Gamma_{212} = y^2 \left(-\frac{1}{y^3}\right) = -\frac{1}{y}.$$

Similarly,  $\Gamma^1_{11} = 0 = \Gamma^1_{22}$ , and the geodesic equation for the first coordinate becomes

$$\frac{d^2 x}{dt^2} - 2\frac{1}{y} \frac{dx}{dt} \frac{dy}{dt} = 0. \tag{28.66}$$

For the second coordinate, we need  $\Gamma^2_{11}, \Gamma^2_{12}$ , and  $\Gamma^2_{22}$ . These can be readily evaluated as above, with the result

$$\Gamma^2_{11} = -\Gamma^2_{22} = \frac{1}{y}, \quad \Gamma^2_{12} = 0,$$

yielding the geodesic equation for the second coordinate

$$\frac{d^2y}{dt^2} + \frac{1}{y} \left(\frac{dx}{dt}\right)^2 - \frac{1}{y} \left(\frac{dy}{dt}\right)^2 = 0. \tag{28.67}$$

With  $\dot{x} \equiv dx/dt$ , Equation (28.66) can be written as

$$\frac{d\dot{x}}{dt} - 2\dot{x} \frac{dy/dt}{y} = 0 \Rightarrow \frac{d\dot{x}/dt}{\dot{x}} = 2 \frac{dy/dt}{y} \Rightarrow \dot{x} = Cy^2.$$

Using the chain rule and the notation  $y' \equiv dy/dx$ , we obtain

$$\begin{aligned} \frac{dy}{dt} &= y' \dot{x} = Cy^2 y', \\ \frac{d^2y}{dt^2} &= C(2y \frac{dy}{dt} y' + y^2 \frac{dy'}{dx} \dot{x}) = C^2(2y^3 y'^2 + y^4 y''). \end{aligned}$$

Substituting in Equation (28.67) yields

$$y^3 y'^2 + y^4 y'' + y^3 = 0 \Rightarrow (y')^2 + yy'' + 1 = 0 \Rightarrow \frac{d}{dx}(yy') + 1 = 0.$$

It follows that  $yy' = -x + A$  and  $x^2 + y^2 = 2Ax + B$ . Thus, the geodesics are circles with arbitrary radii whose centers lie on the  $x$ -axis. ■

## 28.4 Isometries and Killing Vector Fields

Whenever a structure is defined on a given family of sets, mappings between such sets that preserve that structure become the “natural” mappings. Thus, the natural maps among vector spaces are *linear* maps, among groups are *homomorphisms*, among algebras are *algebra homomorphisms*, and among manifolds are *smooth* maps. The introduction of a metric on a manifold makes certain maps more privileged than others.

**28.4.1. Definition.** Let  $M$  and  $N$  be Riemannian manifolds with metrics  $\mathfrak{g}_M$  and  $\mathfrak{g}_N$ , respectively. The smooth map  $\psi : M \rightarrow N$  is called **isometric at**  $P \in M$  if  $\mathfrak{g}_M(\mathbf{X}, \mathbf{Y}) = \mathfrak{g}_N(\psi_*\mathbf{X}, \psi_*\mathbf{Y})$  for all  $\mathbf{X}, \mathbf{Y} \in \mathcal{T}_P(M)$ . An **isometry** of  $M$  to  $N$  is a bijective smooth map that is isometric at every point of  $M$ , in which case we have  $\psi_*\mathfrak{g}_M = \mathfrak{g}_N$ .

Isometries of a manifold form a subgroup of  $\text{Diff}(M)$ .

Of special interest are isometries  $\psi : M \rightarrow M$  of a single manifold. These happen to be a subgroup of  $\text{Diff}(M)$ , the group of diffeomorphisms of  $M$ . In the subgroup of isometries, we concentrate on the one-parameter groups of transformations. These define (and are defined by) certain vector fields:

**28.4.2. Definition.** Let  $\mathbf{X} \in \mathcal{X}(M)$  be a vector field with integral curve  $F_t$ . Then  $\mathbf{X}$  is called a **Killing vector field** if  $F_t$  is an isometry of  $M$ .

The following proposition follows immediately from the definition of the Lie derivative [Equation (26.29)].

**28.4.3. Proposition.** *A vector field  $X \in \mathcal{X}(M)$  is a Killing vector field if and only if  $L_X \mathbf{g} = 0$ .*

Choosing a coordinate system  $\{x^i\}$ , we write  $\mathbf{g} = g_{ij} dx^i \otimes dx^j$  and conclude that  $X^j \partial_j$  is a Killing vector field if and only if

$$0 = L_X(g_{ij} dx^i \otimes dx^j) = X(g_{ij}) dx^i \otimes dx^j + g_{ij} (L_X dx^i) \otimes dx^j + g_{ij} dx^i \otimes (L_X dx^j).$$

Killing equation Using Equation (26.33) for the 1-form  $dx^k$ , we obtain the **Killing equation**

$$X^k \partial_k g_{ij} + \partial_i X^k g_{kj} + \partial_j X^k g_{ki} = 0. \tag{28.68}$$

If in Equation (28.45) we replace  $T$  with  $\mathbf{g}$  and  $\mathbf{u}$  with  $X$ , where  $X$  is a Killing vector field, we obtain

$$0 = 0 + g_{ij} [(\nabla_{e_k} X)^i \epsilon^k \otimes \epsilon^j + (\nabla_{e_k} X)^j \epsilon^i \otimes \epsilon^k], \tag{28.69}$$

where we have assumed that the covariant derivative is compatible with the metric tensor. The reader may check that Equation (28.69) leads to

Killing equation in terms of covariant derivative

$$X_{j;k} + X_{k;j} = 0. \tag{28.70}$$

This is another form of Killing equation.

**28.4.4. Proposition.** *If  $X$  is a Killing vector field, then its inner product with the tangent to any geodesic is constant along that geodesic, i.e., if  $\mathbf{u}$  is such a tangent, then  $\nabla_{\mathbf{u}}[\mathbf{g}(\mathbf{u}, X)] = 0$ .*

*Proof.* We write the desired covariant derivative in component form,

$$\begin{aligned} u^k (g_{ij} u^i X^j)_{;k} &= u^k \underbrace{g_{ij;k}}_{=0} u^i X^j + u^k g_{ij} \underbrace{u^i_{;k}}_{=0} X^j + u^k g_{ij} u^i X^j_{;k} \\ &= \frac{1}{2} \underbrace{(X_{i;k} + X_{k;i})}_{=0 \text{ by (28.70)}} u^i u^k, \end{aligned}$$

where the first term vanishes by assumption of the compatibility of the metric and the covariant derivative, and the second term by the geodesic equation.  $\square$

**28.4.5. Example.** In a flat  $m$ -dimensional manifold we can choose an orthonormal coordinate frame (Theorem 28.2.9), so that the Killing equation becomes

$$\partial_i X_j + \partial_j X_i = 0. \tag{28.71}$$

Setting  $i = j$ , we see that  $\partial_j X_j = 0$  (no sum). Differentiating Equation (28.71) with respect to  $x^i$ , we obtain

$$\partial_i^2 X_j + \underbrace{\partial_j \partial_i X_i}_{=0} = 0 \Rightarrow \partial_i^2 X_j = 0 \quad \forall i, j.$$

Therefore,  $X_j$  is linear in  $x^i$ , i.e.,  $X_j = a_{jk}x^k + b_j$  with  $a_{jk}$  and  $b_j$  arbitrary. Inserting this in (28.71), we get  $a_{ij} + a_{ji} = 0$ . The Killing vector is then

$$\mathbf{X} = (a^i_j x^j + b^i) \partial_i = \frac{1}{2} a^i_j (x^j \partial_i - x^i \partial_j) + b^i \partial_i.$$

maximally symmetric spaces

The first term is clearly the generator of a rotation and the second term that of translation. Altogether, there are  $m(m - 1)/2$  rotations and  $m$  translations. So the total number of independent Killing vectors is  $m(m + 1)/2$ . Manifolds that have this many Killing vectors are called **maximally symmetric spaces**. ■

Using Equations (26.34) and (26.35) one can show that the set of Killing vector fields on a manifold form a vector subspace of  $\mathcal{X}(M)$ , and that if  $\mathbf{X}$  and  $\mathbf{Y}$  are Killing vector fields, then so is  $[\mathbf{X}, \mathbf{Y}]$ . Thus,

**28.4.6. Box.** *The set of Killing vector fields forms a Lie algebra.*

**28.4.7. Example.** From  $\mathbf{g} = d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi$ , the metric of the unit sphere  $S^2$ , one writes the Killing equations

$$\begin{aligned} \partial_\theta X_\theta + \partial_\theta X_\theta &= 0, \\ \partial_\varphi X_\varphi + \partial_\varphi X_\varphi + 2 \sin \theta \cos \theta X_\theta &= 0, \\ \partial_\theta X_\varphi + \partial_\varphi X_\theta - 2 \cot \theta X_\varphi &= 0. \end{aligned} \tag{28.72}$$

The first equation implies that  $X_\theta = f(\varphi)$ , a function of  $\varphi$  only. Substitution in the second equation yields

$$X_\varphi = -\frac{1}{2} F(\varphi) \sin 2\theta + g(\theta), \quad \text{where} \quad f(\varphi) = \frac{dF}{d\varphi}.$$

Inserting this in the third equation of (28.72), we obtain

$$-F(\varphi) \cos 2\theta + \frac{dg}{d\theta} + \frac{df}{d\varphi} + 2 \cot \theta [\frac{1}{2} F(\varphi) \sin 2\theta - g(\theta)] = 0,$$

or

$$\left[ \frac{dg}{d\theta} - 2 \cot \theta g(\theta) \right] + \left[ \frac{df}{d\varphi} + F(\varphi) \right] = 0.$$

For this equation to hold for all  $\theta$  and  $\varphi$ , we must have

$$\frac{dg}{d\theta} - 2 \cot \theta g(\theta) = C = -\frac{df}{d\varphi} - F(\varphi),$$



where  $C$  is a constant. This gives

$$g(\theta) = (C_1 - C \cot \theta) \sin^2 \theta \quad \text{and} \quad f(\varphi) = X_\theta = A \sin \varphi + B \cos \varphi$$

with

$$X_\varphi = (A \cos \varphi - B \sin \varphi) \sin \theta \cos \theta + C_1 \sin^2 \theta.$$

A general Killing vector field is thus given by

$$\mathbf{X} = X^\theta \partial_\theta + X^\varphi \partial_\varphi = AL_y - BL_x + C_1 L_z,$$

where

$$L_x = -\cos \varphi \partial_\theta + \cot \theta \sin \varphi \partial_\varphi,$$

$$L_y = \sin \varphi \partial_\theta + \cot \theta \cos \varphi \partial_\varphi,$$

$$L_z = \partial_\varphi$$

are the generators of  $SO(3)$ . ■

discussion of conformal transformations and conformal Killing vector fields

Sometimes it is useful to relax the complete invariance of the metric tensor under the diffeomorphism of a manifold induced by a vector field and allow a change of scale in the metric. More precisely, we consider vector fields  $\mathbf{X}$  whose flow  $F_t$  changes the metric of  $M$ . So  $F_{t*} \mathbf{g} = e^{\phi(t)} \mathbf{g}$ , where  $\phi$  is a real-valued function on  $M$  that is also dependent on the parameter  $t$ . Such a transformation keeps angles unchanged but rescales all lengths. In analogy with those of the complex plane with the same property, we call such transformations **conformal transformations**. A vector field that generates a conformal transformation will satisfy

$$X^k \partial_k g_{ij} + \partial_i X^k g_{kj} + \partial_j X^k g_{ki} = -\psi g_{ij}, \quad \psi = \left. \frac{\partial \phi}{\partial t} \right|_{t=0}, \quad (28.73)$$

and is called a **conformal Killing vector field**.

We now specialize to a flat  $m$ -dimensional manifold and choose an orthonormal coordinate frame (Theorem 28.2.9). Then Equation (28.73) becomes

$$\partial_i X_j + \partial_j X_i = -\psi g_{ij}. \quad (28.74)$$

Remember Einstein's summation convention!

Multiply both sides by  $g^{ij}$  and sum over  $i$  to obtain

$$2\partial^i X_i = -m\psi \Rightarrow \partial^i X_i = -\frac{m}{2}\psi, \quad m = \dim M.$$

Apply  $\partial^i$  to both sides of Equation (28.74) and sum over  $i$ . This yields

$$\partial^i \partial_i X_j + \partial_j \underbrace{\partial^i X_i}_{-\frac{m}{2}\psi} = -\partial_j \psi \Rightarrow \partial^i \partial_i X_j = \frac{1}{2}(m-2)\partial_j \psi. \quad (28.75)$$

Differentiate both sides of the second equation in (28.75) with respect to  $x^k$  and symmetrize the result in  $j$  and  $k$  to obtain

$$(m - 2)\partial_j \partial_k \psi = \partial^i \partial_i (\partial_k X_j + \partial_j X_k) = -g_{jk} \partial^i \partial_i \psi.$$

Raising the index  $j$  and contracting it with  $k$  gives  $\partial^i \partial_i \psi = 0$  if  $m \neq 1$ . It follows that

$$(m - 2)\partial_j \partial_k \psi = 0. \tag{28.76}$$

Equations (28.74), (28.75), and (28.76) determine both  $\psi$  and  $X_j$  if  $m \neq 2$ . It follows from Equation (28.76) that  $\psi$  is linear in  $x$ , and, consequently [from (28.74)], that  $X_i$  is at most quadratic in  $x$ . The most general solution to Equation (28.74) satisfying (28.75) and (28.76) is

$$X^j(x) = b^j + \varepsilon x^j + a^j_k x^k + c^j x^k x_k - 2c_k x^k x^j, \tag{28.77}$$

where  $a_{ij} = -a_{ji}$  and indices of the constants are raised and lowered as usual.

conformal group

Equation (28.77) gives the generators of an  $[(m+1)(m+2)/2]$ -parameter group of transformations on  $\mathbb{R}^m$ ,  $m \neq 2$  called the **conformal group**. The reader should note that translation (represented by the parameters  $b^j$ ) and rotations (represented by the parameters  $a_{ij}$ ) are included in this group. The other finite (as opposed to infinitesimal) transformations of coordinates can be obtained by using Equation (27.25). For example, the finite transformation generated by the parameter  $\varepsilon$  is given by the solution to the DE  $dx'^j/dt = x'^j$ , which is  $x'^j = e^t x^j$ , or  $x'^j = e^\varepsilon x^j$ , and is called a **dilatation** of coordinates. Similarly, the finite transformation generated by the parameter  $c_i$  is given by the solution to the DE  $dx'^j/dc_i = \delta^{ij} x'^k x'_k - 2x'^i x'^j$ , or

dilatation

inversion, or special conformal transformation

$$x'^i = \frac{x^i - c^i x^2}{1 - 2c \cdot x + c^2 x^2}, \quad \text{where } c \cdot x \equiv c^k x_k, \quad c^2 \equiv c \cdot c, \quad x^2 \equiv x \cdot x,$$

which is called **inversion**, or the **special conformal transformation**. Equations (28.75), and (28.76) place no restriction on  $\psi$  (and therefore on  $X_i$ ) when  $m = 2$ . This means that

**28.4.8. Box.** *The conformal group is infinite-dimensional for  $\mathbb{R}^2$ .*

In fact, we encountered the conformal transformations of  $\mathbb{R}^2$  in the context of complex analysis, where we showed that *any* (therefore, infinitely many) analytic function is a conformal transformation of  $\mathbb{C} \cong \mathbb{R}^2$ . The conformal group of  $\mathbb{R}^2$  has important applications in string theory and statistical mechanics, but we shall not pursue them here.

## 28.5 Geodesic Deviation and Curvature

Geodesics are the straight lines of general manifolds on which, for example, free particles move. If  $\mathbf{u}$  represents the tangent to a given geodesic, one can say that  $\nabla_{\mathbf{u}}\mathbf{u} = 0$  is the equation of motion of a free particle. In flat spaces, the relative velocity of any pair of free particles will not change, so that their relative acceleration is always zero. In general, however, due to the effects of curvature, we expect a nonzero acceleration. Let us elaborate on this.

Consider some region of the manifold through whose points geodesics can be drawn in various directions. Concentrate on one geodesic and its neighboring geodesics. Let  $t$  designate the parameter that locates points of the geodesic. Let  $s$  be a continuous parameter that labels different geodesics (see Figure 28.1). One can connect the points on some neighboring geodesics corresponding to the same value of  $t$  and obtain a curve parametrized by  $s$ . The collection of all geodesics that pass through all points of this curve form a two-dimensional submanifold with coordinates  $t$  and  $s$ . Each such geodesic is thus described by the geodesic equation  $\nabla_{\mathbf{u}}\mathbf{u} = 0$  with  $\mathbf{u} = \partial/\partial t$  (because  $t$  is a coordinate). Furthermore, as we hop from one geodesic to its neighbor, the geodesic equation does not change; i.e., the geodesic equation is independent of  $s$ . Translated into the language of calculus, this means that differentiation of the geodesic equation with respect to  $s$  will give zero. Translated into the higher-level language of tensor analysis, it means that covariant differentiation of the geodesic equation will yield zero. We write this final translation as  $\nabla_{\mathbf{n}}(\nabla_{\mathbf{u}}\mathbf{u}) = 0$  where  $\mathbf{n} \equiv \partial/\partial s$ . This can also be written as

$$0 = \nabla_{\mathbf{n}}\nabla_{\mathbf{u}}\mathbf{u} = \nabla_{\mathbf{u}}\nabla_{\mathbf{n}}\mathbf{u} + [\nabla_{\mathbf{n}}, \nabla_{\mathbf{u}}]\mathbf{u} = \nabla_{\mathbf{u}}(\nabla_{\mathbf{n}}\mathbf{u} + [\mathbf{n}, \mathbf{u}]) + [\nabla_{\mathbf{n}}, \nabla_{\mathbf{u}}]\mathbf{u},$$

where we have used the first property of Theorem 28.3.2. Using the fact that  $\mathbf{n}$  and  $\mathbf{u}$  are coordinate frames, we conclude that  $[\mathbf{n}, \mathbf{u}] = 0$  which in conjunction with Equation (28.56) yields

$$\nabla_{\mathbf{u}}\nabla_{\mathbf{n}}\mathbf{u} + \mathbf{R}(\mathbf{n}, \mathbf{u})\mathbf{u} = 0. \tag{28.78}$$

relative acceleration

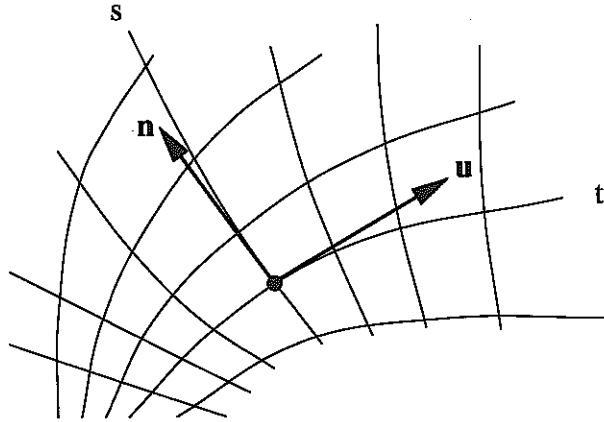
The first term can be interpreted as the **relative acceleration** of two geodesic curves (or free particles), because  $\nabla_{\mathbf{u}}$  is the generalization of the derivative with respect to  $t$ , and  $\nabla_{\mathbf{n}}\mathbf{u}$  is interpreted as relative velocity. In a flat manifold, the relative acceleration for any pair of free particles is zero. When curvature is present, it produces a nonzero relative acceleration.

equation of geodesic deviation

By writing  $\mathbf{u} = u^i\partial_i$  and  $\mathbf{n} = n^k\partial_k$  and substituting in Equation (28.78), we arrive at the **equation of geodesic deviation** in coordinate form:

$$u^i u^j n^k R^m_{ijk} = \frac{d^2 n^m}{dt^2} + \frac{d}{dt} (u^j n^k \Gamma^m_{kj}) + u^j \frac{dn^k}{dt} \Gamma^m_{kj} + u^i u^j n^k \Gamma^n_{kj} \Gamma^m_{ni}, \tag{28.79}$$

where we have used the fact that  $u^i\partial_i f = df/dt$  for any function defined on the manifold.



**Figure 28.1** A region of the manifold and the two-dimensional surface defined by  $s$  and  $t$ .

The chain that connects relative acceleration to curvature has another link that connects the latter two to gravity. From a Newtonian standpoint, gravity is the only force that accelerates all objects at the same rate (equivalence principle). From a geometric standpoint, this property allows one to include gravity in the structure of space-time: An object in free fall is considered “free,” and its path, a geodesic. Locally, this is in fact a better picture of reality, because inside a laboratory in free fall (such as a space shuttle in orbit around earth) one actually verifies the first law of motion on all objects floating in midair. One need not include an external phenomenon called gravity. Gravity becomes part of the fabric of space-time. But how *does* gravity manifest itself? Is there any observable effect that can indicate the presence of gravity, or by a mere transfer to a freely falling frame have we been able to completely eliminate gravity? The second alternative would be strange indeed, because the source of gravity is matter, and if we eliminate gravity completely, we have to eliminate matter as well! If the gravitational field were homogeneous, one *could* eliminate it—and the matter that produces it as well, but no such gravitational field exists. The inhomogeneity of gravitational fields has indeed an observable effect. Consider two test particles that are falling freely toward the source of gravity on two different radii. As they get closer and closer to the center, their relative distance—in fact, their relative velocity—changes: They experience a relative acceleration. Since as we saw in Equation (28.78), relative acceleration is related to curvature, we conclude that

Einstein's  
interpretation of  
gravity

**28.5.1. Box.** *Gravity manifests itself by giving space-time a curvature.*

This is Einstein’s interpretation of gravity. From a Newtonian standpoint, the relative acceleration is caused by the inhomogeneity of the gravitational field. Such inhomogeneity (in the field of the Moon and the Sun) is responsible for tidal waves. That is why the curvature term in Equation (28.78) is also called the **tide-producing** gravitational force.

### 28.5.1 Riemann Normal Coordinates

Starting with a point  $P$  of an  $n$ -dimensional manifold  $M$  on which a covariant derivative is defined, we can construct a unique geodesic in every direction, i.e., for every vector in  $\mathcal{T}_P(M)$ . By parallel transportation of the tangent vectors at  $P$ , we can construct a vector field in a neighborhood of  $P$ : The value of the vector field at  $Q$ —assumed to be close enough to  $P$ —is the tangent at  $Q$  on the geodesic starting at  $P$  and passing through  $Q$ .<sup>8</sup> The vector field so obtained makes it possible to define an exponential map from the tangent space to the manifold. In fact, the integral curve  $\exp(t\mathbf{X})$  of any tangent vector  $\mathbf{X}$  in  $\mathcal{T}_P(M)$  is simply the geodesic associated with the vector.

The uniqueness of the geodesics establishes a bijection (in fact, a diffeomorphism) between a neighborhood of the origin of  $\mathcal{T}_P(M)$  and a neighborhood of  $P$  in  $M$ . This diffeomorphism can be used to assign coordinates to all points in the vicinity of  $P$ . Recall that a coordinate is a smooth bijection from  $M$  to  $\mathbb{R}^n$ . Now choose a basis for  $\mathcal{T}_P(M)$  and associate the components of  $t\mathbf{X}$  in this basis to the points on the geodesic  $\exp(t\mathbf{X})$ . Specifically, if  $\{a^i\}_{i=1}^n$  are the components of  $\mathbf{X}$  in the chosen basis, then

Riemann normal coordinates

$$x^i(t) = a^i t, \quad i = 1, 2, \dots, n,$$

are the so-called **Riemann normal coordinates** (RNCs) of points on the geodesic of  $\mathbf{X}$ . The geodesic equations in these coordinates become

$$\Gamma^k_{ji} a^i a^j = 0 \Rightarrow \Gamma^k_{ji} = 0$$

because  $\Gamma^k_{ji}$  is symmetric in  $i$  and  $j$ .

**28.5.2. Proposition.** *The connection coefficients at a point  $P \in M$  vanish in the Riemann normal coordinates at  $P$ .*

Using Equation (28.43), we immediately obtain the following:

**28.5.3. Corollary.** *Let  $\mathbf{T}$  be a tensor field on  $M$  with components  $T^{i_1 \dots i_r}_{j_1 \dots j_s}$  with respect to a Riemann normal coordinate system  $\{x^i\}$  at  $P$ . Then*

$$T^{i_1 \dots i_r}_{j_1 \dots j_s; k} = \frac{\partial}{\partial x^k} T^{i_1 \dots i_r}_{j_1 \dots j_s}.$$

<sup>8</sup>We are assuming that through any two neighboring points one can always draw a geodesic. For a proof see [Koba 63, pp. 172–175].

Riemann normal coordinates are very useful in establishing tensor equations. This is because

**28.5.4. Box.** *Two tensors are identical if and only if their components are the same in any coordinate frame.*

Therefore, to show that two tensors fields are equal, we pick an arbitrary point in  $M$ , erect a set of RNCs, and show that the components of the tensors are equal. Since the connection coefficients vanish in an RNC system, and covariant derivatives are the same as ordinary derivatives, tensor manipulations can be simplified considerably. For example, the components of the curvature tensor in RNCs are

$$R^i_{\ jkl} = \frac{\partial \Gamma^i_{\ jl}}{\partial x^k} - \frac{\partial \Gamma^i_{\ jk}}{\partial x^l}. \tag{28.80}$$

This is *not* a tensor relation—the RHS is not a tensor in a general coordinate system. However, if we establish a relation involving the components of the curvature tensor alone, then that relation will hold in all coordinates, i.e., it is a tensor relation. For instance, from the equation above one immediately obtains

$$R^i_{\ jkl} + R^i_{\ ljk} + R^i_{\ klj} = 0.$$

Since this involves only a tensor, it must hold in all coordinate frames. This is the Bianchi identity of Equation (28.24).

**28.5.5. Example.** Differentiate Equation (28.62) with respect to  $x^m$  and evaluate the result in RNC to get

$$R^i_{\ jkl;m} = R^i_{\ jkl,m} = \Gamma^i_{\ jl,km} - \Gamma^i_{\ jk,lm}.$$

second Bianchi identity

From this relation and  $\Gamma^i_{\ jl,km} = \Gamma^i_{\ jl,mk}$ , we obtain the **second Bianchi identity**:

$$R^i_{\ jkl;m} + R^i_{\ jmk;l} + R^i_{\ jlm;k} = 0 \quad \text{and} \quad R^i_{\ j[kl;m]} = 0. \tag{28.81}$$

In Einstein’s general relativity, this identity is the analogue of Maxwell’s pair of homogeneous equations:  $F_{\alpha\beta,\gamma} + F_{\gamma\alpha,\beta} + F_{\beta\gamma,\alpha} = 0$ . ■

### 28.5.2 Newtonian Gravity

The equivalence principle, relating gravity with the curvature of space-time, is not unique to Einstein. What is unique to him is combining that principle with the assumption of *local Lorentz geometry*, i.e., local validity of special relativity. Cartan also used the equivalence principle to reformulate Newtonian gravity in the language of geometry. Rewrite Newton’s second law of motion as

$$\mathbf{F} = m\mathbf{a} \Rightarrow \mathbf{a} = -\nabla\Phi \Rightarrow \frac{\partial^2 x^j}{\partial t^2} + \frac{\partial\Phi}{\partial x^j} = 0,$$

where  $\Phi$  is the gravitational potential (potential energy per unit mass). The Newtonian universal time is a parameter that has two degrees of freedom: Its origin and its unit of measurement are arbitrary. Thus, one can change  $t$  to  $t = a\tau + b$  without changing the physics of gravity. Taking this freedom into account, one can write

$$\frac{d^2 t}{d\tau^2} = 0, \quad \frac{d^2 x^j}{d\tau^2} + \frac{\partial \Phi}{\partial x^j} \left( \frac{dt}{d\tau} \right)^2 = 0. \quad (28.82)$$

Comparing this with the geodesic equation, we can read off the nonzero connection coefficients:

$$\Gamma_{00}^j = \frac{\partial \Phi}{\partial x^j}, \quad j = 1, 2, 3. \quad (28.83)$$

Inserting these in Equation (28.62), we find the nonzero components of Riemann curvature tensor:

$$R_{0k0}^j = -R_{00k}^j = \frac{\partial^2 \Phi}{\partial x^j \partial x^k}. \quad (28.84)$$

Contraction of the two nonzero indices leads to the Laplacian of gravitational potential

$$R_{0j0}^j = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = \nabla^2 \Phi.$$

Therefore, the Poisson equation for gravitational potential can be written in terms of the curvature tensor:

$$R_{00} \equiv R_{0j0}^j = 4\pi G\rho, \quad (28.85)$$

Ricci tensor defined where we have introduced the **Ricci tensor**, defined as

$$R_{ik} \equiv R_{ijk}^j. \quad (28.86)$$

Equations (28.83), (28.84), and (28.85) plus the law of geodesic motion describe the full content of Newtonian gravitational theory in the geometric language of tensors.<sup>9</sup>

It is instructive to discover the relation between curvature and gravity directly from the equation of geodesic deviation as applied to Newtonian gravity. The geodesic equation is the equation of motion:

$$\frac{\partial^2 x^j}{\partial t^2} + \frac{\partial \Phi}{\partial x^j} = 0.$$

<sup>9</sup>The classic and comprehensive book *Gravitation*, by Misner, Thorne, and Wheeler, has a thorough discussion of Newtonian gravity in the language of geometry in Chapter 12 and is highly recommended.

Differentiate this equation with respect to the parameter  $s$ , noting that  $\partial/\partial s = n^i \partial/\partial x^i$ :

$$\frac{\partial}{\partial s} \left( \frac{\partial^2 x^j}{\partial t^2} \right) + \frac{\partial}{\partial s} \left( \frac{\partial \Phi}{\partial x^j} \right) = 0 \Rightarrow \frac{\partial^2}{\partial t^2} \left( \frac{\partial x^j}{\partial s} \right) + n^i \frac{\partial}{\partial x^i} \left( \frac{\partial \Phi}{\partial x^j} \right) = 0.$$

Now note that  $\partial x^j/\partial s = n^i \partial x^j/\partial x^i = n^j$ . So, we obtain

$$\frac{\partial^2 n^j}{\partial t^2} + n^i \frac{\partial^2 \Phi}{\partial x^i \partial x^j} = 0.$$

This is equivalent to Equation (28.78), and one recognizes the second term as the tide-producing (or the curvature) term.

## 28.6 General Theory of Relativity

No treatment of Riemannian geometry is complete without a discussion of the general theory of relativity. That is why we shall devote this last section of the current chapter to a brief exposition of this theory.

We have seen that Newtonian gravity can be translated into the language of differential geometry by identifying the gravitational tidal effects with the curvature of space-time. This straightforward interpretation of Newtonian gravity, in particular the retention of the Euclidean metric and the universality of time, leads to no new physical effect. Furthermore, it is inconsistent with the special theory of relativity, which mixes space and time coordinates via Lorentz transformations. Einstein's general theory of relativity (GTR) combines the equivalence principle (that freely falling objects move on geodesics) with the local validity of the special theory of relativity (that the metric of space-time reduces to the Lorentz-Minkowski metric of the special theory of relativity).

### 28.6.1 Einstein's Equation

Since the metric plays a central role in general relativity, let us study its effect on the curvature tensor. The reader may verify that (Problem 28.23)

$$0 = \mathbf{R}(\mathbf{u}, \mathbf{v})(\mathbf{w} \cdot \mathbf{x}) = [\mathbf{R}(\mathbf{u}, \mathbf{v})\mathbf{w}] \cdot \mathbf{x} + \mathbf{w} \cdot [\mathbf{R}(\mathbf{u}, \mathbf{v})\mathbf{x}].$$

Using this, one can show that

$$R_{ijkl} = -R_{jikl}, \quad \text{where } R_{ijkl} \equiv g_{im} R^m{}_{jkl}, \quad (28.87)$$

which is the first equation of (28.23). Equations (28.23) and (28.24) form a complete set of symmetries of the Riemann curvature tensor. Other symmetries that follow from them are

$$R_{ijkl} = R_{klij} \quad \text{and} \quad R_{[ijkl]} = 0. \quad (28.88)$$



An important tensor that can be constructed out of Riemann curvature tensor and the metric tensor is the **Einstein tensor**  $\mathbf{G}$ , which is related to the Ricci tensor defined in Equation (28.86). To derive the Einstein tensor, first note that the Ricci tensor is symmetric in its indices (see Problem 28.25):

$$R_{ij} = R_{ji}. \quad (28.89)$$

Next, define the **curvature scalar** as

$$R \equiv R^i_i \equiv g^{ij} R_{ij}. \quad (28.90)$$

Now, contract  $i$  with  $m$  in Equation (28.81) and use the antisymmetry of the Riemann tensor in its last two indices to obtain

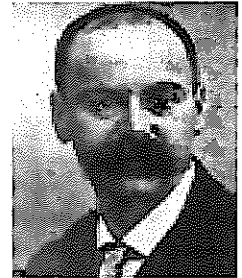
$$R^i_{jkl;i} + R_{jk;l} - R_{jl;k} = 0.$$

Finally, contract  $j$  and  $l$  and use the antisymmetry of the Riemann tensor in its first as well as its last two indices to get  $2R^i_{k;i} - R_{;k} = 0$ , or  $R_{jk;i} - \frac{1}{2}g_{jk}R_{;i} = 0$ . Summarizing the foregoing discussion, we write

$$\nabla \cdot \mathbf{G} = 0, \quad \text{where} \quad G_{ij} \equiv R_{ij} - \frac{1}{2}g_{ij}R. \quad (28.91)$$

**Karl Schwarzschild** (1873–1916) was the eldest of five sons and one daughter born to Moses Martin Schwarzschild and his wife, Henrietta Sabel. His father was a prosperous member of the business community in Frankfurt, with Jewish forbears in that city traced back to the sixteenth century.

From his mother, a vivacious, warm person, Karl undoubtedly inherited his happy, outgoing personality, and from his father, a capacity for sustained hard work. His childhood was spent in comfortable circumstances among a large circle of relatives, whose interests included art and music; he was the first to become a scientist.



After attending a Jewish primary school, Schwarzschild entered the municipal gymnasium in Frankfurt at the age of eleven. His curiosity about the heavens was first manifested then: He saved his allowance and bought lenses to make a telescope. Indulging this interest, his father introduced him to a friend, J. Epstein, a mathematician who had a private observatory. With Epstein's son (later professor of mathematics at the University of Strasbourg), Schwarzschild learned to use a telescope and studied mathematics of a more advanced type than he was getting in school. His precocious mastery of celestial mechanics resulted in two papers on double star orbits, written when he was barely sixteen.

In 1891 Schwarzschild began two years of study at the University of Strasbourg, where Ernst Becker, director of the observatory, guided the development of his skills in practical astronomy—skills that later were to form a solid underpinning for his masterful mathematical abilities.

At age twenty Schwarzschild went to the University of Munich. Three years later, in 1896, he obtained his Ph.D., *summa cum laude*. His dissertation was an application of Poincaré's theory of stable configurations in rotating bodies to several astronomical problems, including tidal deformation in satellites and the validity of Laplace's suggestion as to how the solar system had originated. Before graduating, Schwarzschild also found time to do some practical work with Michelson's interferometer.

At a meeting of the German Astronomical Society in Heidelberg in 1900 he discussed the possibility that space was non-Euclidean. In the same year he published a paper giving a lower limit for the radius of curvature of space as 2500 light years. From 1901 until 1909 he was professor at Göttingen, where he collaborated with Klein, Hilbert, and Minkowski, publishing on electrodynamics and geometrical optics. In 1906, he studied the transport of energy through a star by radiation.

From Göttingen he went to Potsdam, but in 1914 he volunteered for military service. He served in Belgium, France, and Russia. While in Russia he wrote two papers on Einstein's relativity theory and one on Planck's quantum theory. The quantum theory paper explained that the Stark effect could be proved from the postulates of quantum theory.

Schwarzschild's relativity papers give the first exact solution of Einstein's general gravitational equations, giving an understanding of the geometry of space near a point mass. He also made the first study of black holes, showing that bodies of sufficiently large mass would have an escape velocity exceeding the speed of light and so could not be seen. However, he contracted an illness while in Russia and died soon after returning home.

The automatic vanishing of the divergence of the *symmetric* Einstein tensor has an important consequence in the field equation of GTR. It is reminiscent of a similar situation in electromagnetism, in which the vanishing of the divergence of the fields leads to the conservation of the electric charge, the source of electromagnetic fields.<sup>10</sup>

Just as Maxwell's equations are a generalization of the static electricity of Coulomb to a dynamical theory, Einstein's GTR is the generalization of Newtonian static gravity to a dynamical theory. As this generalization ought to agree with the successes of the Newtonian gravity, Equation (28.91) must agree with (28.85). The bold step taken by Einstein was to generalize this relation involving only a single component of the Ricci tensor to a full tensor equation. The natural tensor to be used as the source of gravitation is the **stress energy tensor**<sup>11</sup>

$$T^{\mu\nu} \equiv (\rho + p)u^\mu u^\nu + pg^{\mu\nu}, \quad \text{or} \quad \mathbf{T} = (\rho + p)\mathbf{u} \otimes \mathbf{u} + p\mathbf{g},$$

where the source is treated as a fluid with density  $\rho$ , four-velocity  $\mathbf{u}$ , and pressure  $p$ . So, Einstein suggested the equation  $\mathbf{G} = \kappa\mathbf{T}$  as the generalization of Newton's

<sup>10</sup>It was Maxwell's discovery of the inconsistency of the pre-Maxwellian equations of electromagnetism with charge conservation that prompted him to change not only the fourth equation (to make the entire set of equations consistent with the charge conservation), but also the course of human history.

<sup>11</sup>In GTR, it is customary to use the convention that Greek indices run from 0 to 3, i.e., they include both space and time, while Latin indices encompass only the space components.

Einstein's equation of  
the general theory of  
relativity

universal law of gravitation. Note that  $\nabla \cdot \mathbf{G} = 0$  automatically guarantees mass-energy conservation as in Maxwell's theory of electromagnetism. Problem 28.27 calculates  $\kappa$  to be  $8\pi$  in units in which the universal gravitational constant and the speed of light are set equal to unity. We therefore have

$$\mathbf{G} = 8\pi\mathbf{T}, \quad \text{or} \quad \mathbf{R} - \frac{1}{2}R\mathbf{g} = 8\pi[(\rho + p)\mathbf{u} \otimes \mathbf{u} + p\mathbf{g}]. \quad (28.92)$$

This is **Einstein's equation** of the general theory of relativity.

cosmological  
constant

The Einstein tensor  $\mathbf{G}$  is nearly the only symmetric second-rank tensor made out of the Riemann and metric tensors that is divergence free. The only other tensor with the same properties is  $\mathbf{G} + \Lambda\mathbf{g}$ , where  $\Lambda$  is the so-called **cosmological constant** (see Problem 28.28). When in 1917, Einstein applied his GTR to the universe itself, he found that the universe ought to be expanding. Being a firm believer in Nature, he changed his equation to  $\mathbf{G} + \Lambda\mathbf{g} = 8\pi\mathbf{T}$  to suppress the unobserved prediction of the expansion of the universe. Later, when the expansion was observed by Hubble, Einstein referred to this mutilation of his GTR as "the biggest blunder of my life."

**Aleksandr Aleksandrovich Friedmann** (1888–1925) was born into a musical family—his father, Aleksandr Friedmann, being a composer and his mother, Ludmila Vojáčka, the daughter of the Czech composer Hynek Vojáček.

In 1906 Friedmann graduated from the gymnasium with the gold medal and immediately enrolled in the mathematics section of the department of physics and mathematics of St. Petersburg University. While still a student, he wrote a number of unpublished scientific papers, one of which was awarded a gold medal by the department. After graduation from the university in 1910, Friedmann was retained in the department to prepare for the teaching profession.



In the fall of 1914, Friedmann volunteered for service in an aviation detachment, in which he worked, first on the northern front and later on other fronts, to organize aerologic and aeronavigational services. While at the front, Friedmann often participated in military flights as an aircraft observer. In the summer of 1917 he was appointed a section chief in Russia's first factory for the manufacture of measuring instruments used in aviation; he later became director of the factory. Friedmann had to relinquish this post because of the onset of heart disease. From 1918 until 1920, he was professor in the department of theoretical mechanics of Perm University.

In 1920 he returned to Petrograd and worked at the main physics observatory of the Academy of Sciences, first as head of the mathematics department and later, shortly before his death, as director of the observatory. Friedmann's scientific activity was concentrated in the areas of theoretical meteorology and hydromechanics, where he demonstrated his mathematical talent and his unwavering strife for, and ability to attain, the concrete, practical application of solutions to theoretical problems.

Friedmann made a valuable contribution to Einstein's general theory of relativity. As always, his interest was not limited simply to familiarizing himself with this new field

of science but led to his own remarkable investigations. Friedmann's work on the theory of relativity dealt with the cosmological problem. In his paper "Über die Krümmung des Raumes" (1922), he outlined the fundamental ideas of his cosmology: the supposition concerning the homogeneity of the distribution of matter in space and the consequent homogeneity and isotropy of space-time. This theory is especially important because it leads to a sufficiently correct explanation of the fundamental phenomenon known as the "red shift." Einstein himself thought that the cosmological solution to the equations of a field had to be static and had to lead to a closed model of the universe. Friedmann discarded both conditions and arrived at an independent solution.

Friedmann's interest in the theory of relativity was by no means a passing fancy. In the last years of his life, together with V. K. Frederiks, he began work on a multivolume text on modern physics. The first book, *The World as Space and Time*, is devoted to the theory of relativity, knowledge of which Friedmann considered one of the cornerstones of an education in physics.

In addition to his scientific work, Friedmann taught courses in higher mathematics and theoretical mechanics at various colleges in Petrograd. He found time to create new and original courses, brilliant in their form and exceedingly varied in their content. Friedmann's unique course in theoretical mechanics combined mathematical precision and logical continuity with original procedural and physical trends.

Friedmann died of typhoid fever at the age of thirty-seven. In 1931, he was posthumously awarded the Lenin Prize for his outstanding scientific work.

### 28.6.2 Static Spherically Symmetric Solutions

The general theory of relativity as given in Equation (28.92) has been strikingly successful in predicting the spacetime<sup>12</sup> structure of our universe. It predicts the expansion of the universe, and by time-reversed extrapolation, the big bang cosmology; it predicts the existence of black holes and other final products of stellar collapse; and on a less grandiose scale, it explains the small precession of Mercury, the bending of light in the gravitational field of the Sun, and the gravitational redshift. We shall not discuss the solution of Einstein's equation in any detail. However, due to its simplicity and its use of geometric arguments, we shall consider the solution to Einstein's equation exterior to a static spherically symmetric distribution of mass.

Let us first translate the two adjectives used in the last sentence into a geometric language. Take static first. We call a phenomenon "static" if at different instants it "looks the same." Thus, a static solution of Einstein's equation is a spacetime manifold that "looks the same" for all time. In the language of geometry "looks the same" means isometric, because metric is the essence of the geometry of spacetime. In Euclidean physics, time can be thought of as an axis at each point (moment)

<sup>12</sup>The reader may be surprised to see the two words "space" and "time" juxtaposed with no hyphen; but this is common practice in relativity.

of which one can assign a three-dimensional space corresponding to the “spatial universe” at that moment. In the general theory of relativity, space and time can be mixed, but the character of time as a parameter remains unaltered. Therefore, instead of an axis—a straight line—we pick a curve, a parametric map from the real line to the manifold of space-time. This curve must be *timelike*, so that locally, when curvature is ignored and special relativity becomes a good approximation, we do not violate causality. The curve must also have the property that at each point of it, the space-time manifold has the same metric. Moreover, we need to demand that at each point of this curve, the spatial part of the space-time is orthogonal to the curve.

stationary and static spacetimes; time translation isometries

**28.6.1. Definition.** A spacetime is *stationary* if there exists a one-parameter group of isometries  $F_t$ , called *time translation isometries*, whose Killing vector fields  $\xi$  are timelike for all  $t$ :  $g(\xi, \xi) > 0$ . If in addition, there exists a spacelike hypersurface  $\Sigma$  that is orthogonal to orbits (curves) of the isometries, we say that the spacetime is *static*.

Killing parameter

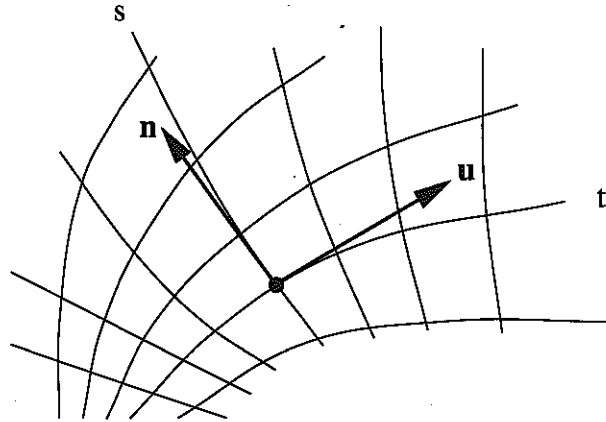
We can simplify the solution to Einstein’s equation by invoking the symmetry of spacetime discussed above in our choice of coordinates. Let  $P$  be a point of the spacetime manifold located in a neighborhood of some spacelike hypersurface  $\Sigma$  as shown in Figure 28.2. Through  $P$  passes a single orbit of the isometry, which starts at a point  $Q$  of  $\Sigma$ . Let  $t$ , the so-called **Killing parameter**, stand for the parameter corresponding to the point  $P$  with  $t = 0$  being the parameter of  $Q$ . On the spacelike hypersurface  $\Sigma$ , choose arbitrary coordinates  $\{x^i\}$  for  $Q$ . Assign the coordinates  $(t \equiv x^0, x^1, x^2, x^3)$  to  $P$ . Since  $F_t$  does not change the metric,  $\Sigma_t \equiv F_t \Sigma$ , the translation of  $\Sigma$  by  $F_t$ , is also orthogonal to the orbit of the isometry. Moreover, the components of the metric in this coordinate system cannot be dependent on the Killing parameter  $t$ . Thus, in this coordinate system, the spacetime metric takes the form

$$g = g_{00} dt \otimes dt - \sum_{i,j=1}^3 g_{ij} dx^i \otimes dx^j. \tag{28.93}$$

spherically symmetric spacetimes

**28.6.2. Definition.** A spacetime is *spherically symmetric* if its isometry group contains a subgroup isomorphic to  $SO(3)$  and the orbits of this group are two-dimensional spheres.

In other words, if we think of isometries as the action of some abstract group, then this group must contain  $SO(3)$  as a subgroup. Since  $SO(3)$  is isomorphic to the group of rotations, we conclude that the metric should be rotationally invariant. The time-translation Killing vector field  $\xi$  must be orthogonal to the orbits of  $SO(3)$ , because otherwise the generators of  $SO(3)$  can change the projection of  $\xi$  on the spheres and destroy the rotational invariance. Therefore, the 2-dimensional spheres must lie entirely in the hypersurfaces  $\Sigma_t$ . Now, we can write down a



**Figure 28.2** The coordinates appropriate for a stationary spacetime.

static spherically symmetric metric in terms of appropriate coordinates as follows. Choose the spherical coordinates  $(\theta, \varphi)$  for the 2-spheres, and write the metric of this sphere as

$$g_2 = r^2 d\theta \otimes d\theta + r^2 \sin^2 \theta d\varphi \otimes d\varphi,$$

where  $r$  is the “radius” of the 2-sphere. Choose the third spatial coordinate to be orthogonal to this sphere, i.e.,  $r$ . Rotational symmetry now implies that the components of the metric must be independent of  $\theta$  and  $\varphi$ . The final form of the metric based entirely on the assumed symmetries is

$$g = f(r) dt \otimes dt - h(r) dr \otimes dr - r^2(d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi). \tag{28.94}$$

We have reduced the problem of finding ten unknown functions  $\{g_{\mu\nu}\}_{\mu,\nu=0}^3$  of four variables  $\{x^\mu\}_{\mu=0}^3$  to that of two functions  $f$  and  $h$  of one variable. The remaining task is to calculate the Ricci tensor corresponding to Equation (28.94), substitute it in Einstein’s equation (with the RHS equal to zero), and solve the resulting differential equation for  $f$  and  $h$ . We shall not pursue this further here, and we refer the reader to textbooks on general relativity (see, for example, [Wald 84, pp. 121–124]). The final result is the so-called **Schwarzschild solution**, which is

Schwarzschild  
solution of Einstein's  
equation

$$g = \left(1 - \frac{2M}{r}\right) dt \otimes dt - \left(1 - \frac{2M}{r}\right)^{-1} dr \otimes dr - r^2(d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi), \tag{28.95}$$

where  $M$  is the total mass of the gravitating body, and the natural units of GTR, in which  $G = 1 = c$ , have been used.

Schwarzschild radius

A remarkable feature of the Schwarzschild solution is that the metric components have singularities at  $r = 2M$  and at  $r = 0$ . It turns out that the first singularity is due to the choice of coordinates (analogous to the singularity at  $r = 0, \theta = 0, \pi$ , and  $\varphi = 0$  in spherical coordinates of  $\mathbb{R}^3$ ), while the second is a true singularity of spacetime. The first singularity occurs at the so-called **Schwarzschild radius** whose numerical value is given by

$$r_S = \frac{2GM}{c^2} \approx 3 \frac{M}{M_\odot} \text{ km,}$$

where  $M_\odot = 2 \times 10^{30}$  kg is the mass of the Sun. Therefore, for an ordinary body such as the Earth, planets, and a typical star, the Schwarzschild radius is well inside the body where the Schwarzschild solution is not applicable.

If we relax the assumption of staticity, we get the following theorem (for a proof, see [Misn 73, p. 843]):

**28.6.3. Theorem.** (Birkhoff’s theorem) *The Schwarzschild solution is the only spherically symmetric solution of Einstein’s equation in vacuum.*

A corollary to this theorem is that

**28.6.4. Box.** *All spherically symmetric solutions of Einstein’s equation in vacuum are static.*

This is analogous to the fact that the Coulomb solution is the only spherically symmetric solution to Maxwell’s equations in vacuum. It can be interpreted as the statement that in gravity, as in electromagnetism, there is no monopole “radiation.”

### 28.6.3 Schwarzschild Geodesics

With the metric available to us, we can, in principle, solve the geodesic equations [Equation (28.65)] and obtain the trajectories of freely falling particles. However, a more elegant way is to make further use of the symmetries to eliminate variables. In particular, Proposition 28.4.4 is extremely useful in this endeavor. Consider first  $\mathbf{g}(\partial_\theta, \mathbf{u})$  where  $\mathbf{u}$  is the 4-velocity (tangent to the geodesic). In the coordinates we are using, this yields

$$\mathbf{g}(\partial_\theta, \mathbf{u}) = \mathbf{g}(\partial_\theta, \dot{x}^\mu \partial_\mu) = \dot{x}^\mu \underbrace{\mathbf{g}(\partial_\theta, \partial_\mu)}_{r^2 \delta_{\theta\mu}} = r^2 \dot{\theta}.$$

This quantity (because of Proposition 28.4.4 and the fact that  $\partial_\theta$  is a Killing vector field) is a constant of the motion, and its initial value will be an attribute of the particle during its entire motion. We assign zero to this constant, i.e., we assume that initially  $\dot{\theta} = 0$ . This is possible, because by rotating our spacetime—an allowed

operation due to rotational symmetry—we take the equatorial plane  $\theta = \pi/2$  to be the initial plane of the motion. Then the motion will be confined to this plane, because  $\dot{\theta} = 0$  for all time.

For the parameter of the geodesic equation, choose proper time  $\tau$  if the geodesic is timelike (massive particles), and any (affine) parameter if the geodesic is null (massless particles such as photons). Then  $g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = \kappa$ , where

$$\kappa = \begin{cases} 1 & \text{for timelike geodesics,} \\ 0 & \text{for null geodesics.} \end{cases} \quad (28.96)$$

In terms of our chosen coordinates (with  $\theta = \pi/2$ ), we have

$$\kappa = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = (1 - 2M/r)\dot{t}^2 - (1 - 2M/r)^{-1}\dot{r}^2 - r^2\dot{\phi}^2. \quad (28.97)$$

Next, we apply Proposition 28.4.4 to the time translation Killing vector and write

$$E = g_{\mu\nu}\dot{x}^\mu\xi^\nu = (1 - 2M/r)\dot{t}, \quad (28.98)$$

where  $E$  is a constant of the motion and  $\xi = \partial_t$ . In the case of massive particles, as  $r \rightarrow \infty$ , i.e., as we approach special relativity,  $E$  becomes  $\dot{t}$ , which is the rest energy of a particle of unit mass.<sup>13</sup> Therefore, it is natural to interpret  $E$  for finite  $r$  as the total energy (including gravitational potential energy) per unit mass of a particle on a geodesic.

Finally, the other rotational Killing vector field  $\partial_\phi$  gives another constant of motion,

$$L = \mathbf{g}(\partial_\phi, \mathbf{u}) = r^2\dot{\phi}, \quad (28.99)$$

which can be interpreted as the angular momentum of the particle. This reduces to Kepler's second law: Equal areas are swept out in equal times, in the limit of Newtonian (or weak) gravity. However, in strong gravitational fields, spacetime is not Euclidean, and Equation (28.99) cannot be interpreted as "areas swept out." Nevertheless, it is interesting that the "form" of Kepler's second law does not change even in strong fields.

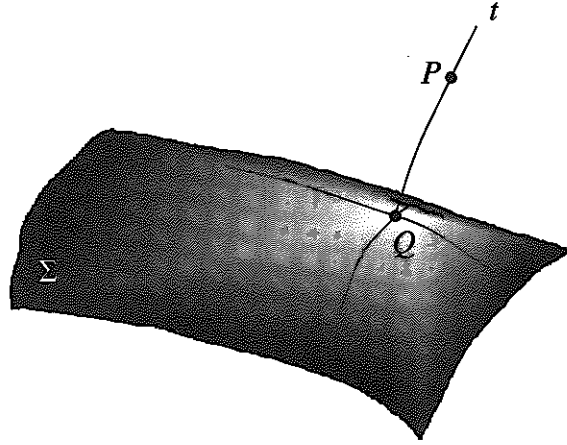
Solving for  $\dot{t}$  and  $\dot{\phi}$  from (28.98) and (28.99) and inserting the result in (28.97), we obtain

$$\frac{1}{2}\dot{r}^2 + \frac{1}{2}\left(1 - \frac{2M}{r}\right)\left(\frac{L^2}{r^2} + \kappa\right) = \frac{1}{2}E^2. \quad (28.100)$$

It follows from this equation that the radial motion of a particle on a geodesic is the same as that of a unit mass particle of energy  $E^2/2$  in ordinary one-dimensional

<sup>13</sup>Recall that the 4-momentum of special relativity is  $p^\mu = m\dot{x}^\mu$ .





**Figure 28.3** The effective potential  $V(r)$  for a massive particle with  $L^2 = 5M^2$ .

nonrelativistic mechanics moving in an effective potential

$$V(r) = \frac{1}{2} \left( 1 - \frac{2M}{r} \right) \left( \frac{L^2}{r^2} + \kappa \right) = \frac{1}{2} \kappa - \kappa \frac{M}{r} + \frac{L^2}{2r^2} - \frac{ML^2}{r^3}. \quad (28.101)$$

Once we solve Equation (28.100) for the radial motion in this effective potential, we can find the angular motion and the time coordinate change from (28.98) and (28.99). The new feature provided by GTR is that in the radial equation of motion, in addition to the “Newtonian term”  $-\kappa M/r$  and the “centrifugal barrier”  $L^2/2r^2$ , we have the new term  $-ML^2/r^3$ , which, a small correction for large  $r$ , will dominate over the centrifugal barrier term for small  $r$ .

Let us consider first the massive particle case,  $\kappa = 1$ . The extrema of the effective potential are given by

$$Mr^2 - L^2r + 3ML^2 = 0 \Rightarrow R_{\pm} = \frac{L^2 \pm \sqrt{L^4 - 12L^2M^2}}{2M}. \quad (28.102)$$

Thus, if  $L^2 < 12M^2$ , no extrema exist (see Figure 28.3), and a particle heading toward the center of attraction, will fall directly to the Schwarzschild radius  $r = 2M$ , the zero of the effective potential, and finally into the spacetime singularity  $r = 0$ . For  $L^2 > 12M^2$ , the reader may check that  $R_+$  is a minimum of  $V(r)$ , while  $R_-$  is a maximum (Figure 28.4). It follows that stable (unstable) circular orbits exist at the radius  $r = R_+$  ( $r = R_-$ ). In the Newtonian limit of  $M \ll L$ , we get  $R_+ \approx L^2/M$ , which agrees with the calculation in Newtonian gravity (Problem 28.30). Furthermore, Equation (28.102) puts a restriction of  $R_+ > 6M$  on  $R_+$  and  $3M < R_- < 6M$  on  $R_-$ . This places the planets of the Sun safely in the region of stable circular orbits.

If a massive particle is displaced slightly from its stable equilibrium radius  $R_+$ , it will oscillate radially with a frequency  $\omega_r$  given by<sup>14</sup>

$$\omega_r^2 = \left. \frac{d^2V}{dr^2} \right|_{r=R_+} = \frac{M(R_+ - 6M)}{R_+^4 - 3MR_+^3}.$$

On the other hand, the orbital frequency is given by Equation (28.99),

$$\omega_\phi^2 = \frac{L^2}{R_+^4} = \frac{M}{R_+^3 - 3MR_+^2},$$

where  $L^2$  has been calculated from (28.102) and inserted in this equation. In the Newtonian limit of  $M \ll R_+$ , we have  $\omega_\phi \approx \omega_r \approx M/R_+^3$ . If  $\omega_\phi = \omega_r$ , the particle will return to a given value of  $r$  in exactly one orbital period and the orbit will close. The difference between  $\omega_\phi$  and  $\omega_r$  in GTR means that the orbit will precess at a rate of

$$\omega_p = \omega_\phi - \omega_r = (1 - \omega_r/\omega_\phi)\omega_\phi = -[(1 - 6M/R_+)^{1/2} - 1]\omega_\phi,$$

which in the limit of  $M \ll R_+$  reduces to

$$\omega_p \approx \frac{3M^{3/2}}{R_+^{5/2}} = \frac{3(GM)^{3/2}}{c^2 R_+^{5/2}}.$$

If we include the eccentricity  $e$  and denote the semimajor axis of the elliptical orbit by  $a$ , then the formula above becomes (see [Misn 73, p. 1110])

$$\omega_p \approx \frac{3(GM)^{3/2}}{c^2(1 - e^2)a^{5/2}}. \quad (28.103)$$

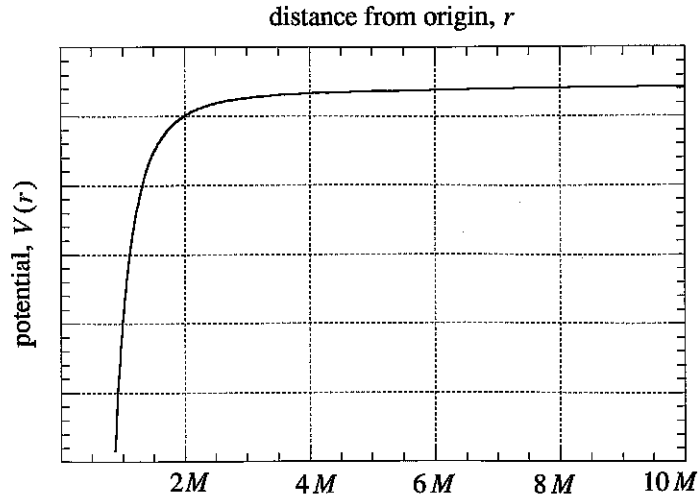
Due to its proximity to the Sun, Mercury shows the largest precession frequency, which, after taking into account all other effects such as perturbations due to other planets, is 43 seconds of arc per century. This residual precession rate had been observed prior to the formulation of GTR and had been an unexplained mystery. Its explanation was one of the most dramatic early successes of GTR.

We now consider the null geodesics. With  $\kappa = 0$  in Equation (28.101), the effective potential becomes

$$V(r) = \frac{L^2}{2r^3}(r - 2M), \quad (28.104)$$

which has a *maximum* at  $r = R_{\max} \equiv 3M$ , as shown in Figure 28.5. It follows that in GTR, unstable circular orbits of photons exist at  $r = 3M$ , and that strong gravity has significant effect on the propagation of light.

<sup>14</sup>In the Taylor expansion of any potential  $V(r)$  about the equilibrium position  $r_0$  of a particle (of unit mass), it is the second derivative term that resembles Hooke's potential,  $\frac{1}{2}kx^2$  with  $k = (d^2V/dr^2)_{r_0}$ .



**Figure 28.4** The effective potential  $V(r)$  for a massive particle with  $L^2 = 20M^2$ .

The minimum energy required to overcome the potential barrier (and avoid falling into the infinitely deep potential well) is given by

$$\frac{1}{2}E^2 = V(R_{\max}) = \frac{L^2}{54M^2} \Rightarrow \frac{L^2}{E^2} = 27M^2.$$

In flat spacetime,  $L/E$  is precisely the impact parameter  $b$  of a photon, i.e., the distance of closest approach to the origin. Thus the Schwarzschild geometry will capture any photon sent toward it if the impact parameter is less than  $b_c \equiv \sqrt{27}M$ . Hence, the cross section for photon capture is

$$\sigma \equiv \pi b_c^2 = 27\pi M^2.$$

To analyze the bending of light, we write Equation (28.100) as

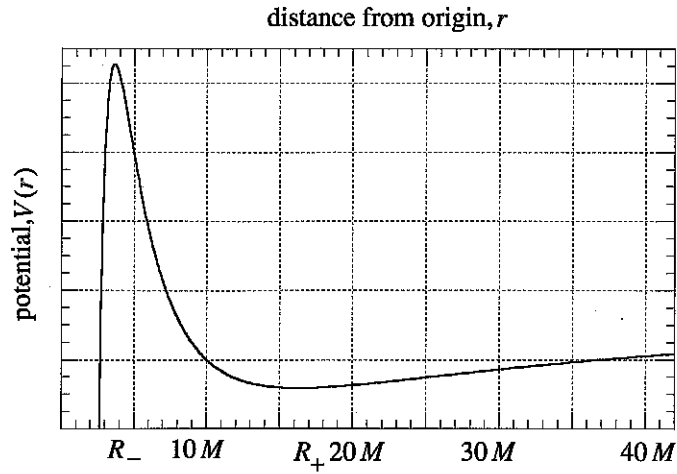
$$\frac{1}{2} \left( \frac{dr}{d\varphi} \right)^2 \dot{\varphi}^2 + V(r) = \frac{1}{2}E^2.$$

Substituting for  $\dot{\varphi}$  from Equation (28.99) and writing the resulting DE in terms of a new variable  $u = M/r$ , we obtain

$$\left( \frac{du}{d\varphi} \right)^2 + u^2(1 - 2u) = \frac{b^2}{M^2}, \quad b \equiv \frac{E}{L},$$

where we used Equation (28.104) for the effective potential. Differentiating this with respect to  $\varphi$ , we finally get the second-order DE,

$$\frac{d^2u}{d\varphi^2} + u = 3u^2. \quad (28.105)$$



**Figure 28.5** The effective potential  $V(r)$  for a massless particle. Note that the shape of the potential is independent of  $L^2$ .

In the large-impact-parameter or small- $\mu$  approximation, we can ignore the second-order term on the RHS and solve for  $\mu$ . This will give the equation of a line in polar coordinates. Substituting this solution on the RHS of Equation (28.105) and solving the resulting equation yields the deviation from a straight line with a deflection angle of

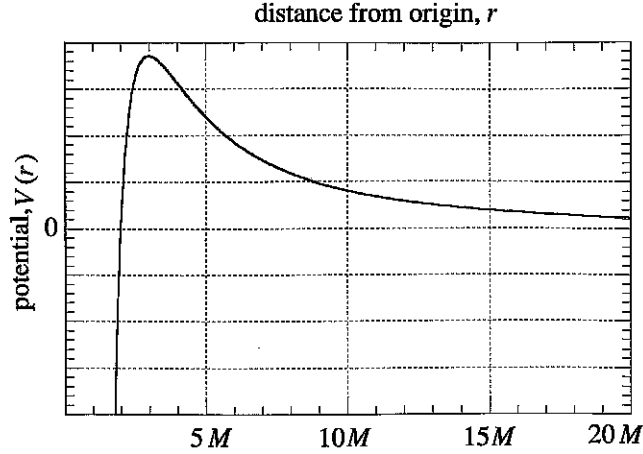
$$\Delta\varphi \approx \frac{4M}{b} = \frac{4GM}{bc^2}, \tag{28.106}$$

where we have restored the  $G$ 's and the  $c$ 's in the last step.

For a light ray grazing the Sun,  $b = R_\odot = 7 \times 10^8$  m and  $M = M_\odot = 2 \times 10^{30}$  kg, so that Equation (28.106) predicts a deflection of 1.747 seconds of arc. This bending of starlight passing near the Sun has been observed many times, beginning with the 1919 expedition led by Eddington. Because of the intrinsic difficulty of such measurements, these observations confirm GTR only to within 10% accuracy. However, the bending of radio waves emitted by quasars has been measured to an accuracy of 1%, and the result has been shown to be in agreement with Equation (28.106) to within this accuracy. The last topic we want to discuss, a beautiful illustration of Proposition 28.4.4 is the **gravitational redshift**. Let  $O_1$  and  $O_2$  be two static observers (by which we mean that they each move on an integral curve of the Killing vector field  $\xi$ ). It follows that the 4-velocities  $\mathbf{u}_1$  and  $\mathbf{u}_2$  of the observers are proportional to  $\xi$ . Since  $\mathbf{u}_1$  and  $\mathbf{u}_2$  have unit lengths, we have

gravitational redshift  
discussed

$$\mathbf{u}_i = \frac{\xi}{\sqrt{g(\xi, \xi)}}, \quad i = 1, 2.$$



**Figure 28.6** A spacetime diagram depicting the emission of light by observer  $O_1$  and its reception by  $O_2$ .

Suppose  $O_1$  emits a light beam and  $O_2$  receives it (see Figure 28.6). Since light travels on a geodesic, Proposition 28.4.4 gives

$$\mathbf{g}(\mathbf{u}, \xi_1) = \mathbf{g}(\mathbf{u}, \xi_2), \quad (28.107)$$

where  $\mathbf{u}$  is tangent to the light trajectory (or the light signal's 4-velocity). The frequency of light for any observer is the time component of its 4-velocity, and because the 4-velocity of an observer has the form  $(1, 0, 0, 0)$  in the frame of that observer, we can write this invariantly as

$$\omega_i = \mathbf{g}(\mathbf{u}, \mathbf{u}_i) = \frac{\mathbf{g}(\mathbf{u}, \xi_i)}{\sqrt{\mathbf{g}(\xi_i, \xi_i)}}, \quad i = 1, 2.$$

In particular, using Equation (28.107), we obtain

$$\frac{\omega_1}{\omega_2} = \frac{\mathbf{g}(\mathbf{u}, \mathbf{u}_1)}{\mathbf{g}(\mathbf{u}, \mathbf{u}_2)} = \frac{\mathbf{g}(\mathbf{u}, \xi_1)/\sqrt{\mathbf{g}(\xi_1, \xi_1)}}{\mathbf{g}(\mathbf{u}, \xi_2)/\sqrt{\mathbf{g}(\xi_2, \xi_2)}} = \frac{\sqrt{\mathbf{g}(\xi_2, \xi_2)}}{\sqrt{\mathbf{g}(\xi_1, \xi_1)}} = \sqrt{\frac{1 - 2M/r_2}{1 - 2M/r_1}},$$

where we used  $\mathbf{g}(\xi, \xi) = g_{00} = (1 - 2M/r)$  for the Schwarzschild spacetime, and  $r_1$  and  $r_2$  are the radial coordinates of the observers  $O_1$  and  $O_2$ , respectively. In terms of wavelengths, we have

$$\frac{\lambda_1}{\lambda_2} = \sqrt{\frac{1 - 2M/r_1}{1 - 2M/r_2}}. \quad (28.108)$$

It follows from Equation (28.108) that as light moves toward regions of weak gravity ( $r_2 > r_1$ ), the wavelength increases ( $\lambda_2 > \lambda_1$ ), i.e., it will be "red-shifted."

this makes sense, because an increase in distance from the center implies an increase in the gravitational potential energy, and, therefore, a decrease in a photon's energy  $\hbar\omega$ . Pound and Rebka used the Mössbauer effect in 1960 to measure the change in the wavelength of a beam of light as it falls down a tower on the surface of the Earth. They found that, to within the 1% experimental accuracy, the GTR prediction of the gravitational redshift was in agreement with their measurement.

## 28.7 Problems

28.1. Show that  $d\Theta = \Omega \wedge (G\Theta) - \Theta \wedge (G\Omega)$ .

28.2. Let A and B be matrices whose elements are one-forms. Show that  $(A \wedge B)^t = -B^t \wedge A^t$ .

28.3. Write Equation (28.20) in component form and derive Equation (28.24).

28.4. Find  $d\Omega$  if

$$\Omega = \begin{pmatrix} 0 & -\cot\theta\epsilon^\varphi \\ \cot\theta\epsilon^\varphi & 0 \end{pmatrix},$$

where  $(\theta, \varphi)$  are coordinates on the unit sphere  $S^2$ .

28.5. Find the curvature of the two-dimensional space whose arc length is given by  $ds^2 = dx^2 + x^2 dy^2$ .

28.6. Find the curvature of the three-dimensional space whose arc length is given by  $ds^2 = dx^2 + x^2 dy^2 + dz^2$ .

28.7. Find the curvature tensors of the Friedmann and Schwarzschild spaces given in Example 28.2.3.

28.8. (a) Show that in  $\mathbb{R}^3$  the composite operator  $d \circ *$  gives the curl of a vector when the vector is written as components of a two-form.

(b) Similarly, show that  $* \circ d$  is the divergence operator for one-forms.

(c) Use these results and the procedure of Example 28.2.4 to find expressions for the curl and the divergence of a vector in curvilinear coordinates.

28.9. Let  $\mathbf{u} = e_i$ ,  $\mathbf{v} = e_k$ , and  $\omega = \epsilon^j$  in (28.34) to arrive at (28.35).

28.10. Using Equations (28.7) and (28.55) show that  $\mathbf{R}$  is a tensor of type  $(1, 3)$ . In particular, it does not differentiate functions that multiply vectors and the 1-form in its argument.

28.11. Show that

$$\mathbf{R}(\mathbf{u}, \mathbf{v})\mathbf{w} + \mathbf{R}(\mathbf{w}, \mathbf{u})\mathbf{v} + \mathbf{R}(\mathbf{v}, \mathbf{w})\mathbf{u} = 0.$$

Hint: Use the first property of Theorem 28.3.2 to change the vector with respect to which covariant differentiation is being performed. Then use the Jacobi identity for Lie brackets of vectors.

**28.12.** Use (28.55) and (28.11) to show that the components of  $\mathbf{R}$  are precisely those introduced in Equation (28.22).

**28.13.** Prove the statement in Box 28.3.8.

**28.14.** Start with  $d^2x^i/dt^2 = 0$ , the geodesic equations in Cartesian coordinates. Transform these equations to spherical coordinates  $(r, \theta, \varphi)$  using  $x = r \sin \theta \cos \varphi$ ,  $y = r \sin \theta \sin \varphi$ , and  $z = r \cos \theta$ , and the chain rule. From these equations read off the connection coefficients in spherical coordinates [refer to Equation (28.65)]. Now use Equation (28.43) and Definition 28.3.6 to evaluate the divergence of a vector.

**28.15.** Find the geodesics of a manifold whose arc element is  $ds^2 = dx^2 + dy^2 + dz^2$ .

**28.16.** Find the geodesics of the metric  $ds^2 = dx^2 + x^2 dy^2$ .

**28.17.** Find the differential equation for the geodesics of the surface of a sphere of radius  $a$  having the line element  $ds^2 = a^2 d\theta^2 + a^2 \sin^2 \theta d\varphi^2$ . Verify that  $A \cos \varphi + B \sin \varphi + \cot \theta = 0$  is the intersection of a plane passing through the origin and the sphere. Also, show that it is a solution of the differential equation of the geodesics. Hence, the geodesics of a sphere are great circles.

**28.18.** The Riemann normal coordinates are given by  $x^i = a^i t$ . For each set of  $a^i$ , one obtains a different set of geodesics. Thus, we can think of  $a^i$  as the parameters that distinguish among the geodesics.

(a) By keeping all  $a^i$  (and  $t$ ) fixed except the  $j$ th one and using the definition of tangent to a curve, show that  $\mathbf{n}_j = t \partial_j$ , where  $\mathbf{n}_j$  is (one of) the  $\mathbf{n}$ (s) appearing in the equation of geodesic deviation.

(b) Substitute (a) plus  $u^i = \dot{x}^i = a^i$  in Equation (28.79) to show that

$$R^m_{ijk} + R^m_{jik} = 3(\Gamma^m_{ki,j} + \Gamma^m_{kj,i}).$$

Substitute for one of the  $\Gamma$ 's on the RHS using Equation (28.80).

(c) Now use the cyclic property of the lower indices of the curvature tensor to show that

$$\Gamma^m_{ij,k} = -\frac{1}{3}(R^m_{ijk} + R^m_{jik}).$$

**28.19.** Let  $\psi : M \rightarrow N$  be isometric at  $P \in M$ . Show that  $\psi_*$  is necessarily injective. Hint: Look at the null space of  $\psi_*$ .

**28.20.** (a) Show that the covariant and Lie derivatives, when applied to a 1-form, are related by

$$\langle L_{\mathbf{u}}\omega, \mathbf{v} \rangle = \langle \nabla_{\mathbf{u}}\omega, \mathbf{v} \rangle + \langle \omega, \nabla_{\mathbf{v}}\mathbf{u} \rangle.$$

(b) Use this to derive the identity

$$(L_{\mathbf{u}}\omega)_i = (\nabla_{\mathbf{u}}\omega)_i + \omega_j (\nabla_{\mathbf{e}_i}\mathbf{u})^j.$$

28.21. Show that Equation (28.69) leads to Equation (28.70).

28.22. Show that a vector field that generates a conformal transformation satisfies

$$X^k \partial_k g_{ij} + \partial_i X^k g_{kj} + \partial_j X^k g_{ki} = -\psi g_{ij}.$$

28.23. (a) Show that  $\mathbf{R}(\mathbf{u}, \mathbf{v})(f) = 0$  for all functions defined on  $M$ .

(b) Using Equation (28.64) show that

$$0 = \mathbf{R}(\mathbf{u}, \mathbf{v})(\mathbf{w} \cdot \mathbf{x}) = \mathbf{w} \cdot [\mathbf{R}(\mathbf{u}, \mathbf{v})\mathbf{x}] + [\mathbf{R}(\mathbf{u}, \mathbf{v})\mathbf{w}] \cdot \mathbf{x}.$$

(c) Conclude from (b) that  $R_{ijkl} = -R_{jikl}$ .

28.24. Use the symmetries of  $R_{ijkl}$  [Equations (28.23) and (28.24)] to show that  $R_{ijkl} = R_{klij}$  and  $R_{[ijkl]} = 0$ .

28.25. Use the symmetry properties of Riemann curvature tensor to show that

(a)  $R^i{}_{ijk} = R_{mij}{}^j = 0$ , and

(b)  $R_{ij} = R_{ji}$ .

(c) Show that  $R^i{}_{jkl;i} + R_{jk;i} - R_{jl;k} = 0$ , and conclude that  $\nabla \cdot \mathbf{G} = 0$ , or, in component form,  $G^i{}_{i;k} = 0$ .

28.26. Show that in an  $n$ -dimensional manifold without metric the number of independent components of the Riemann curvature tensor is

$$\frac{n^3(n-1)}{2} - \frac{n^2(n-1)(n-2)}{6} = \frac{n^2(n^2-1)}{3}.$$

If the manifold has a metric, the number of components reduces to

$$\left[ \frac{n(n-1)}{2} \right]^2 - \frac{n^2(n-1)(n-2)}{6} = \frac{n^2(n^2-1)}{12}.$$

28.27. (a) Take the trace of both sides of  $\mathbf{R} - \frac{1}{2}R\mathbf{g} = \kappa\mathbf{T}$  to obtain  $R = -\kappa T^\mu{}_\mu \equiv -\kappa T$ .

(b) Use (a) to obtain  $R_{00} = \frac{1}{2}\kappa(T_{00} + T^j{}_j)$ .

(c) Now use the fact that in Newtonian limit  $T_{ij} \ll T_{00} \approx \rho$  to conclude that agreement of Einstein's and Newton's gravity demands that  $\kappa = 8\pi$  in units in which the universal gravitational constant is unity.

28.28. (a) Show that the most general second-rank symmetric tensor  $E_{ij}$  constructed from the metric and Riemann curvature tensors that is linear in the curvature tensor is

$$E_{ij} = aR_{ij} + bg_{ij}R + \Lambda g_{ij},$$

where  $a$ ,  $b$ , and  $\Lambda$  are constants.

(b) Show that  $E_{ij}$  has a vanishing divergence if and only if  $b = -\frac{1}{2}a$ .

(c) Show that in addition,  $E_{ij}$  vanishes in flat space-time if and only if  $\Lambda = 0$ .



**28.29.** Show that  $R_+$  and  $R_-$ , as given by Equation (28.102) are, respectively, a minimum and a maximum of  $V(r)$ .

**28.30.** Use  $F = ma$  to show that in a circular orbit of radius  $R$ , we have  $L^2 = GM R$ .

**28.31.** Show that  $R_+ > 6M$  and  $3M < R_- < 6M$ , where  $R_{\pm}$  are given by Equation (28.102).

**28.32.** Calculate the energy of a circular orbit using Equation (28.100), and show that

$$E(R) = \frac{R - 2M}{\sqrt{R^2 - 3MR}},$$

where  $R = R_{\pm}$

**28.33.** Show that the radial frequency of oscillation of a massive particle in a stable orbit of radius  $R_+$  is given by

$$\omega_r^2 = \frac{M(R_+ - 6M)}{R_+^4 - 3MR_+}.$$

**28.34.** Derive Equation (28.106) from Equation (28.105).

## Additional Reading

1. Misner, C., Thorne, K., and Wheeler, J. *Gravitation*, Freeman, 1973. A classic text on index-free differential geometry. This is the definitive text on Einstein's general theory of relativity.
2. Nakahara, M. *Geometry, Topology, and Physics*, Adam Hilger, 1990. A very readable book for graduate students of physics with a lot of emphasis on gauge theories and topological concepts. The book has a standard introduction to manifolds and a good discussion of differential geometry, especially Killing vector fields.
3. Wald, R. *General Relativity*, University of Chicago Press, 1984. A more up-to-date book on differential geometry and general relativity.

# Lie Groups and Differential Equations

Lie groups and Lie algebras, because of their manifold—and therefore, differentiability—structure, find very natural applications in areas of physics and mathematics in which symmetry and differentiability play important roles. Lie himself started the subject by analyzing the symmetry of differential equations in the hope that a systematic method of solving them could be discovered. Later, Emmy Noether applied the same idea to variational problems involving symmetries and obtained one of the most beautiful pieces of mathematical physics: the relation between symmetries and conservation laws. More recently, generalizing the gauge invariance of electromagnetism, Yang and Mills have considered nonabelian gauge theories in which gauge invariance is governed by a nonabelian Lie group. Such theories have been successfully built for three of the four fundamental interactions: electromagnetism, weak nuclear, and strong nuclear. Furthermore, it has been possible to cast the fourth interaction, gravity—as described by Einstein’s general theory of relativity—in a language very similar to the other three interactions with the promise of unifying all four interactions into a single force. This chapter is devoted to a treatment of the first topic, application of Lie groups to DEs. The second topic, the calculus of variations and conservation laws, will be discussed in the next chapter. The third topic, that of gauge theories, although of fundamental importance to the development of physics and our understanding of the universe, is, at this stage, too specialized to be covered in this book.

## 29.1 Symmetries of Algebraic Equations

The symmetry group of a system of DEs is a transformation group that acts on both the independent and dependent variables and transforms solutions of the system to

other solutions. In order to understand this symmetry group, we shall first tackle the simpler question of the symmetries of a system of algebraic equations.

**29.1.1. Definition.** Let  $G$  be a local Lie group of transformations acting on a manifold  $M$ . A subset  $S \subset M$  is called *G-invariant* and  $G$  is called a *symmetry group* of  $S$  if whenever  $g \cdot P$  is defined for  $P \in S$  and  $g \in G$ , then  $g \cdot P \in S$ .

*G*-invariance and symmetry group defined

**29.1.2. Example.** Let  $M = \mathbb{R}^2$ .

(a) Let  $G = \mathbb{R}^+$  be the abelian multiplicative group of real numbers. Let it act on  $M$  componentwise:  $r \cdot (x, y) = (rx, ry)$ . Then any line going through the origin is a  $G$ -invariant subset of  $\mathbb{R}^2$ .

(b) If  $G = SO(2)$  and it acts on  $M$  as usual, then any circle is a  $G$ -invariant subset of  $\mathbb{R}^2$ . ■

system of algebraic equations and their symmetry group

A system of algebraic equations is a system of equations

$$F_\nu(x) = 0, \quad \nu = 1, 2, \dots, n,$$

in which  $F_\nu : M \rightarrow \mathbb{R}$  is smooth. A *solution* is a point  $x \in M$  such that  $F_\nu(x) = 0$  for  $\nu = 1, \dots, n$ . The *solution set* of the system is the collection of all solutions. A Lie group  $G$  is called a *symmetry group of the system* if the solution set is  $G$ -invariant.

invariant map

**29.1.3. Definition.** Let  $G$  be a local Lie group of transformations acting on a manifold  $M$ . A map  $F : M \rightarrow N$ , where  $N$  is another manifold, is called a *G-invariant map* if for all  $P \in M$  and all  $g \in G$  such that  $g \cdot P$  is defined,  $F(g \cdot P) = F(P)$ . A real-valued  $G$ -invariant function is called simply an *invariant*.

The crucial property of Lie group theory is that locally the group and its algebra “look alike.” This allows the complicated nonlinear conditions of invariance of subsets and functions to be replaced by the simpler linear conditions of invariance under infinitesimal actions. From Definition 27.1.25, we obtain the following proposition.

**29.1.4. Proposition.** Let  $G$  be a local group of transformations acting on a manifold  $M$ . A smooth real-valued function  $f : M \rightarrow \mathbb{R}$  is  $G$ -invariant if and only if

$$\xi_M|_P(f) = 0 \quad \text{for all } P \in M \tag{29.1}$$

and for every infinitesimal generator  $\xi \in \mathfrak{g}$ .

**29.1.5. Example.** The infinitesimal generator for  $SO(2)$  is  $\xi_M = x\partial_y - y\partial_x$ . Any function of the form  $f(x^2 + y^2)$  is an  $SO(2)$ -invariant. To see this, we apply Proposition 29.1.4:

$$(x\partial_y - y\partial_x)f(x^2 + y^2) = x(2y)f' - y(2x)f' = 0,$$

where  $f'$  is the derivative of  $f$ . ■

The criterion for the invariance of the solution set of a system of equations is a little more complicated, because now we are not dealing with functions themselves, but with their solutions. The following theorem gives such a criterion (for a proof, see [Olive 86, pp. 82–83]):

**29.1.6. Theorem.** *Let  $G$  be a local Lie group of transformations acting on an  $m$ -dimensional manifold  $M$ . Let  $F : M \rightarrow \mathbb{R}^n$ ,  $n \leq m$ , define a system of algebraic equations  $\{F_\nu(x) = 0\}_{\nu=1}^n$ , and assume that the Jacobian matrix  $(\partial F_\nu / \partial x^k)$  is of rank  $n$  at every solution  $x$  of the system. Then  $G$  is a symmetry group of the system if and only if*

$$\xi_M|_x(F_\nu) = 0 \quad \forall \nu \text{ whenever } F_\nu(x) = 0 \quad (29.2)$$

for every infinitesimal generator  $\xi \in \mathfrak{g}$ .

Note that Equation (29.2) is required to hold *only for solutions  $x$  of the system*.

**29.1.7. Example.** Let  $M = \mathbb{R}^2$  and  $G = SO(2)$ . Consider  $F : M \rightarrow \mathbb{R}$  defined by  $F(x, y) = x^2 + y^2 - 1$ . The Jacobian matrix is simply the gradient,

$$(\partial F / \partial x, \partial F / \partial y) = (2x, 2y),$$

and is of rank 1 for all points of the solution set, because it never vanishes at the points where  $F(x, y) = 0$ , i.e., the unit circle. It follows from Theorem 29.1.6 that  $G$  is a symmetry group of the equation  $F(x, y) = 0$  if and only if  $\xi_M|_r(F) = 0$  whenever  $r \in S^1$ . But

$$\xi_M|_r(F) = (x\partial_y - y\partial_x)F|_r = 2xy - 2yx = 0.$$

This is a proof of the obvious fact that  $SO(2)$  takes points of  $S^1$  to other points of  $S^1$ .

As a less trivial example, consider the function  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  given by

$$F(x, y) = x^2y^2 + y^4 + 2x^2 + y^2 - 2.$$

The infinitesimal action of the group yields

$$\xi_M(F) = (x\partial_y - y\partial_x)F = 2x^3y + 2xy^3 - 2xy = 2xy(x^2 + y^2 - 1).$$

The reader may check that  $\xi_M(F) = 0$  whenever  $F(x, y) = 0$ . The Jacobian matrix of the “system” of equations is the gradient

$$\nabla F = (2xy^2 + 4x, 2x^2y + 4y^3 + 2y),$$

which vanishes only when  $x = 0 = y$ , which does not belong to the solution set. Therefore, the rank of the Jacobian matrix is 1. We conclude that the solution set of  $F(x, y) = 0$  is a rotationally invariant subset of  $\mathbb{R}^2$ . Indeed, we have

$$F(x, y) = x^2y^2 + y^4 + 2x^2 + y^2 - 2 = (y^2 + 2)(x^2 + y^2 - 1),$$

and the solution set is just the unit circle. Note that although the solution set of  $F(x, y) = 0$  is  $G$ -invariant, the function itself is not. ■

We now discuss how to find invariants of a given group action. Start with a one-parameter group and write

$$\mathbf{v} \equiv \xi_M = X^i \frac{\partial}{\partial x^i}$$

for the infinitesimal generator of the group in some local coordinates. A local invariant  $F(x)$  of the group is a solution of the linear, homogeneous PDE

$$\mathbf{v}(F) = X^1(x) \frac{\partial F}{\partial x^1} + \dots + X^n(x) \frac{\partial F}{\partial x^n} = 0. \tag{29.3}$$

It follows that the gradient of  $F$  is perpendicular to the vector  $\mathbf{v}$ . Since the gradient of  $F$  is the normal to the hypersurface of constant  $F$ , we may consider the solution of Equation (29.3) as a surface  $F(x) = c$  whose normal is perpendicular to  $\mathbf{v}$ . Each normal determines one hypersurface, and since there are  $n - 1$  linearly independent vectors perpendicular to  $\mathbf{v}$ , there must be  $n - 1$  different hypersurfaces that solve (29.3). Let us write these hypersurfaces as

$$F^j(x^1, \dots, x^n) = c^j, \quad j = 1, 2, \dots, n - 1, \tag{29.4}$$

and note that

$$\Delta F^j \approx \sum_{i=1}^n \frac{\partial F^j}{\partial x^i} \Delta x^i = 0, \quad j = 1, 2, \dots, n - 1.$$

A solution to this equation is suggested by (29.3):

$$\Delta x^i = \alpha X^i \Rightarrow \frac{\Delta x^i}{X^i} = \alpha.$$

characteristic system  
of a PDE

For  $\Delta x^i \rightarrow dx^i$ , we obtain the following set of ODEs, called the **characteristic system** of the original PDE,

$$\frac{dx^1}{X^1(x)} = \frac{dx^2}{X^2(x)} = \dots = \frac{dx^n}{X^n(x)}, \tag{29.5}$$

whose solutions determine  $\{F^j(x)\}_{j=1}^{n-1}$ . To find these solutions,

**29.1.8. Box.** Take the equalities of (29.5) one at a time, solve the first order DE, write the solution in the form of (29.4), and read off the functions.

The reader may check that any function of the  $F^j$ 's is also a solution of the PDE. In fact, it can be shown that *any solution* of the PDE is a function of these  $F^j$ 's (see [Olve 86, pp. 86–90]).

**29.1.9. Example.** Once again, let us consider  $SO(2)$ , whose infinitesimal generator is  $-y\partial_x + x\partial_y$ . The characteristic "system" of equations is

$$\frac{dx}{-y} = \frac{dy}{x} \Rightarrow x dx + y dy = 0 \Rightarrow x^2 + y^2 = c.$$

Thus,  $F(x, y) = x^2 + y^2$ , or any function thereof, is an invariant of the rotation group in two dimensions.

As a less trivial example, consider the vector field

$$\mathbf{v} = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} + \sqrt{a^2 - z^2} \frac{\partial}{\partial z}$$

where  $a$  is a constant. The characteristic system of ODEs is

$$\frac{dx}{-y} = \frac{dy}{x} = \frac{dz}{\sqrt{a^2 - z^2}}.$$

The first equation was solved above, giving the invariant  $F_1(x, y, z) = \sqrt{x^2 + y^2} = r$ . To find the other invariant, solve for  $x$  and substitute in the second equation to obtain

$$\frac{dy}{\sqrt{r^2 - y^2}} = \frac{dz}{\sqrt{a^2 - z^2}}.$$

The solution to this DE is

$$\underbrace{\arcsin \frac{y}{r}}_{\alpha} = \underbrace{\arcsin \frac{z}{a}}_{\beta} + C \Rightarrow \arcsin \frac{y}{r} - \arcsin \frac{z}{a} = C.$$

Hence,  $F_2(x, y, z) = \arcsin(y/r) - \arcsin(z/a)$  is a second invariant. By taking the sine of  $F_2$ , we can come up with an invariant that is algebraic (rather than trigonometric) in form:

$$\begin{aligned} s &= \sin F_2 = \sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta \\ &= \frac{y}{r} \sqrt{1 - \frac{z^2}{a^2}} - \sqrt{1 - \frac{y^2}{r^2}} \frac{z}{a} = \frac{y\sqrt{a^2 - z^2} - xz}{ra}. \end{aligned}$$

Any function of  $r$  and  $s$  is also an invariant. ■

When the dimension of the Lie group is larger than one, the computation of the invariants can be very complicated. If  $\{\mathbf{v}_k\}_{k=1}^r$  form a basis for the infinitesimal generators, then the invariants are the joint solutions of the system of first order PDEs

$$\mathbf{v}_k(F) = \sum_{j=1}^n X_k^j(x) \frac{\partial F}{\partial x^j}, \quad k = 1, \dots, r.$$

To find such a solution, one solves the first equation and finds all its invariants. Since any function of these invariants is also an invariant, it is natural to express  $F$  as a function of the invariants of  $\mathbf{v}_1$ . One then writes the remaining equations in terms of these new variables and proceeds inductively.

**29.1.10. Example.** Consider the vector fields

$$\mathbf{u} = -y\partial_x + x\partial_y,$$

$$\mathbf{v} = \left( \frac{x^3z + xy^2z + a^3x}{\sqrt{x^2 + y^2}} \right) \partial_x + \left( \frac{x^2yz + y^3z + a^3y}{\sqrt{x^2 + y^2}} \right) \partial_y - \left( \sqrt{x^2 + y^2}z^2 + b^3 \right) \partial_z$$

where  $a$  and  $b$  are constants. The invariants of  $\mathbf{u}$  are functions of  $r = \sqrt{x^2 + y^2}$  and  $z$ . If we are to have a nontrivial solution, the invariant of  $\mathbf{v}$  as well as its PDE should be expressible in terms of  $r = \sqrt{x^2 + y^2}$  and  $z$ . The reader may verify that

$$\mathbf{v}(F) = (r^2z + a^3) \frac{\partial F}{\partial r} - (rz^2 + b^3) \frac{\partial F}{\partial z} = 0$$

with the characteristic equation

$$\frac{dr}{r^2z + a^3} = - \frac{dz}{rz^2 + b^3}.$$

This is an exact first-order DE whose solutions are given by

$$\frac{1}{2}r^2z^2 + a^3z + b^3r = c$$

with  $c$  an arbitrary constant. Therefore,  $F = \frac{1}{2}r^2z^2 + a^3z + b^3r$ , or

$$F(x, y, z) = \frac{1}{2}(x^2 + y^2)z^2 + a^3z + b^3\sqrt{x^2 + y^2},$$

or a function thereof, is the single invariant of this group. ■

## 29.2 Symmetry Groups of Differential Equations

Let  $\mathcal{S}$  be a system of partial differential equations involving  $p$  independent variables  $x = (x^1, \dots, x^p)$ , and  $q$  dependent variables  $u = (u^1, \dots, u^q)$ . The solutions of the system are of the form  $u = f(x)$ , or, in component form,  $u^\alpha = f(x^1, \dots, x^p)$ ,  $\alpha = 1, \dots, q$ . Let  $X = \mathbb{R}^p$  and  $U = \mathbb{R}^q$  be the spaces of independent and dependent variables with coordinates  $\{x^i\}$  and  $\{u^\alpha\}$ , respectively. Roughly speaking, a symmetry group of the system  $\mathcal{S}$  will be a local group of transformations that map solutions of  $\mathcal{S}$  into solutions of  $\mathcal{S}$ .

**Marius Sophus Lie** (1842–1899) was the youngest son of a Lutheran pastor in Norway. He studied mathematics and science at Christiania (which became Kristiania, then Oslo in 1925) University where he attended Sylow’s lectures on group theory. There followed a few years when he could not decide what career to follow. He tutored privately after his graduation and even dabbled a bit in astronomy and mechanics.

A turning point came in 1868 when he read papers on geometry by Poncelet and Plücker from which originated the inspiration in the topic of creating geometries by using elements other than points in space, and provided the seed for the rest of Lie’s career, prompting him to call himself a student of Plücker, even though the two had never met. Lie’s first

publication won him a scholarship to work in Berlin, where he met Klein, who had also been influenced by Plücker's papers. The two had quite different styles—Lie always pursuing the broadest generalization, while Klein could become absorbed in a charming special case—but collaborated effectively for many years. However, in 1892 the lifelong friendship between Lie and Klein broke down, and the following year Lie publicly attacked Klein, saying, "I am no pupil of Klein, nor is the opposite the case, although this might be closer to the truth." Lie and Klein spent a summer in Paris, then parted for some time before resuming their collaboration in Germany. While in Paris, Lie discovered the *contact transformation*, which, for instance, maps lines into spheres. During the Franco-Prussian War, Lie decided to hike to Italy. On the way, however, he was arrested as a German spy and his mathematics notes were assumed to be coded messages. Only after the intervention of Darboux was Lie released, and he decided to return to Christiania. In 1871 Lie became an assistant at Christiania and obtained his doctorate.

After a short stay in Germany, he again returned to Christiania University, where a chair of mathematics was created for him. Several years later Lie succeeded Klein at Leipzig, where he was stricken with a condition, then called neurasthenia, resulting in fatigue and memory loss and once thought to result from exhaustion of the nervous system. Although treatment in a mental hospital nominally restored his health, the once robust and happy Lie became ill-tempered and suspicious, despite the recognition he received for his work. To lure him back to Norway, his friends at Christiania created another special chair for him, and Lie returned in the fall of 1898.



He died of anemia a few months later. Lie had started examining partial differential equations, hoping that he could find a theory that was analogous to Galois's theory of equations. He examined his contact transformations considering how they affected a process due to Jacobi of generating further solutions from a given one. This led to combining the transformations in a way that Lie called a group, but is today called a *Lie algebra*. At this point he left his original intention of examining partial differential equations and examined Lie algebras. Killing was to examine Lie algebras quite independently of Lie, and Cartan was to publish the classification of semisimple Lie algebras in 1900. Much of the work on transformation groups for which Lie is best known was collected with the aid of a postdoctoral student sent to Christiania by Klein in 1884. The student, F. Engel, remained nine months with Lie and was instrumental in the production of the three volume work *Theorie der Transformationsgruppen*, which appeared between 1888 and 1893. A similar effort to collect Lie's work in contact transformations and partial differential equations was sidetracked as Lie's coworker, F. Hausdorff, pursued other topics.

The transformation groups now known as **Lie groups** provided a very fertile area for research for decades to come, although perhaps not at first. When Killing tried to classify the simple Lie groups, Lie considered his efforts so poor that he admonished one of his departing students with these words: "Farewell, and if ever you meet that s.o.b., kill him." Lie's work was continued (somewhat in isolation) by Cartan, but it was the papers of Weyl in the early 1920s that sparked the renewal of strong interest in Lie groups. Much of the foundation of the quantum theory of fundamental processes is built on Lie groups. In 1939,



Wigner showed that application of Lie algebras to the Lorentz transformation required that all particles have the intrinsic properties of mass and spin.

transform the *graph* of a function to find the function's transform!

To make precise the above statement, we have to clarify the meaning of the action of  $G$  on a function  $u = f(x)$ . We start with identifying the function  $f$  (i.e., a map) with its graph (see Chapter 0),

$$\Gamma_f \equiv \{(x, f(x)) \mid x \in \Omega\} \subset X \times U,$$

where  $\Omega \subset X$  is the domain of definition of  $f$ . If the action of  $g \in G$  on  $\Gamma_f$  is defined, then the transform of  $\Gamma_f$  by  $g$  is

$$g \cdot \Gamma_f = \{(\tilde{x}, \tilde{u}) = g \cdot (x, u) \mid (x, u) \in \Gamma_f\}.$$

transform of a function by a group element

In general,  $g \cdot \Gamma_f$  may not represent the graph of a function—in fact, it may not be even a function at all. However, by choosing  $g$  close to the identity of  $G$  and shrinking the size of  $\Omega$ , we can ensure that  $g \cdot \Gamma_f = \Gamma_{\tilde{f}}$ , i.e., that  $g \cdot \Gamma_f$  is indeed the graph of a function  $\tilde{u} = \tilde{f}(\tilde{x})$ . We write  $\tilde{f} = g \cdot f$  and call  $\tilde{f}$  the **transform** of  $f$  by  $g$ .

**29.2.1. Example.** Let  $X = \mathbb{R} = U$ , so that we are dealing with an ODE. Let  $G = SO(2)$  be the rotation group acting on  $X \times U = \mathbb{R}^2$ . The action is given by

$$(\tilde{x}, \tilde{u}) = \theta \cdot (x, u) = (x \cos \theta - u \sin \theta, x \sin \theta + u \cos \theta). \tag{29.6}$$

If  $u = f(x)$  is a function, the group  $SO(2)$  acts on its graph  $\Gamma_f$  by rotating it. This process can lead to a rotated graph  $\theta \cdot \Gamma_f$ , which may not be the graph of a single-valued function. However, if we restrict the interval of definition of  $f$ , and make  $\theta$  small enough, then  $\theta \cdot \Gamma_f$  will be the graph of a well-defined function  $\tilde{u} = \tilde{f}(\tilde{x})$  with  $\Gamma_{\tilde{f}} = \theta \cdot \Gamma_f$ . If we substitute  $f(x)$  for  $u$ , we obtain

$$(\tilde{x}, \tilde{u}) = \theta \cdot (x, f(x)) = (x \cos \theta - f(x) \sin \theta, x \sin \theta + f(x) \cos \theta),$$

or

$$\begin{aligned} \tilde{x} &= x \cos \theta - f(x) \sin \theta, \\ \tilde{u} &= x \sin \theta + f(x) \cos \theta. \end{aligned} \tag{29.7}$$

Eliminating  $x$  from these two equations yields  $\tilde{u}$  in terms of  $\tilde{x}$ , from which the function  $\tilde{f}$  can be deduced.

As a specific example, consider  $f(x) = kx^2$ . Then, the first equation of (29.7) gives

$$(k \sin \theta)x^2 - \cos \theta x + \tilde{x} = 0 \Rightarrow x = \frac{\cos \theta - \sqrt{\cos^2 \theta - 4k\tilde{x} \sin \theta}}{2k \sin \theta},$$

where we kept the root of the quadratic equation that gives a finite answer in the limit  $\theta \rightarrow 0$ . Inserting this in the second equation of (29.7) and simplifying yields

$$\tilde{u} = \tilde{f}(\tilde{x}) = \frac{\cos \theta - \sqrt{\cos^2 \theta - 4k\tilde{x} \sin \theta}}{2k \sin^2 \theta} - \tilde{x} \cot \theta.$$

We write this as

$$\tilde{f}(x) \equiv (\theta \cdot f)(x) = \frac{\cos \theta - \sqrt{\cos^2 \theta - 4kx \sin \theta}}{2k \sin^2 \theta} - x \cot \theta.$$

This equation defines the function  $\tilde{f} = \theta \cdot f$ . ■

The foregoing example illustrates the general procedure for finding the transformed function  $\tilde{f} = g \cdot f$ :

**29.2.2. Box.** *If the rule of transformation of  $g \in G$  is given by*

$$(\tilde{x}, \tilde{u}) = g \cdot (x, u) = (\Psi_g(x, u), \Phi_g(x, u)),$$

*then the graph  $\Gamma_{\tilde{f}} = g \cdot \Gamma_f$  of  $g \cdot f$  is given parametrically by*

$$\tilde{x} = \Psi_g(x, f(x)), \quad \tilde{u} = \Phi_g(x, f(x)). \tag{29.8}$$

In principle, we can solve the first equation for  $x$  in terms of  $\tilde{x}$  and substitute in the second equation to find  $\tilde{u}$  in terms  $\tilde{x}$ , and consequently  $\tilde{f}$ .

For some special but important cases, the transformed functions can be obtained explicitly. If  $G$  is **projectable**, i.e., if the action of  $G$  on  $x$  does not depend on  $u$ , then Equation (29.8) takes the special form  $\tilde{x} = \Psi_g(x)$  and  $\tilde{u} = \Phi_g(x, f(x))$  in which  $\Psi_g$  is a diffeomorphism of  $X$  with inverse  $\Psi_{g^{-1}}$ . If  $\Gamma_f$  is the graph of a function  $f$ , then its transform  $g \cdot \Gamma$  is *always* the graph of some function. In fact,

$$\tilde{u} = \tilde{f}(\tilde{x}) = \Phi_g(x, f(x)) = \Phi_g(\Psi_{g^{-1}}(\tilde{x}), f(\Psi_{g^{-1}}(\tilde{x}))). \tag{29.9}$$

In particular, if  $G$  transforms *only the independent variables*, then

$$\tilde{u} = \tilde{f}(\tilde{x}) = f(x) = f(\Psi_{g^{-1}}(\tilde{x})) \Rightarrow \tilde{f} = f \circ \Psi_{g^{-1}}. \tag{29.10}$$

For example, if  $G$  is the group of translations  $x \mapsto x + a$ , then the transform of  $f$  will be defined by  $\tilde{f}(x) = f(x - a)$ .

symmetry group of a system of DEs

**29.2.3. Definition.** A **symmetry group** of a system of DEs  $\mathcal{S}$  is a local group of transformations  $G$  acting on an open subset  $M$  of  $X \times U$  with the property that whenever  $u = f(x)$  is a solution of  $\mathcal{S}$  and  $\tilde{f} \equiv g \cdot f$  is defined for  $g \in G$ , then  $u = \tilde{f}(x)$  is also a solution of  $\mathcal{S}$ .

The importance of knowing the symmetry group of a system of DEs lies in the property that from one solution we may be able to obtain a family of other solutions by applying the group elements to the given solution. To find such symmetry groups, we have to be able to “prolong” the action of a group to derivatives of the dependent variables as well. This is obvious because to test a symmetry, we have to substitute not only the transformed function  $u = \tilde{f}(x)$ , but also its derivatives in the DE to verify that it satisfies the DE.

### 29.2.1 Prolongation of Functions

Given a function  $f : X \rightarrow \mathbb{R}$ , there are

$$p_k \equiv \binom{p+k-1}{k}$$

different derivatives of order  $k$  of  $f$ . We use the multi-index notation

$$\partial_J f(x) \equiv \frac{\partial^k f(x)}{\partial x^{j_1} \partial x^{j_2} \dots \partial x^{j_k}}$$

for these derivatives, where  $J = (j_1, \dots, j_k)$  is an *unordered*  $k$ -tuple of nonnegative integers with  $1 \leq j_k \leq p$  (see also Section 21.1). The *order* of the multi-index  $J$ , denoted by  $|J|$ , is the sum of its components and indicates the order of differentiation. So, in the derivative above,  $|J| = j_1 + \dots + j_k = k$ . For a smooth map  $f : X \rightarrow U$ , we have  $f(x) = (f^1(x), \dots, f^q(x))$ , so that we need  $q \cdot p_k$  numbers  $\partial_J f^\alpha(x)$  to represent all  $k$ -th order derivatives of all components of  $f$ .

To geometrize the treatment of DEs (and thus facilitate the study of their invariance), we need to construct a space in which derivatives of all orders up to a certain number  $n$  participate. Since derivatives need functions to act on, we arrive at the space of functions whose derivatives share certain common properties. To be specific, we make the following definition.

**29.2.4. Definition.** *Let  $f$  and  $h$  be functions defined on a neighborhood of a point  $a \in X$  with values in  $U$ . We say that  $f$  and  $h$  are  **$n$ -equivalent** at  $a$  if  $\partial_J f^\alpha(a) = \partial_J h^\alpha(a)$  for all  $\alpha$  and  $J$  corresponding to all derivatives up to  $n$ th order. The collection of all  $U$ -valued functions defined on a neighborhood of  $a$  will be denoted by  $\Gamma_a(X \times U)$ , and all functions  $n$ -equivalent to  $f$  at  $a$  by  $j_a^n f$ .*

A convenient representative of such equivalent functions is the Taylor polynomial of order  $n$  (the terms in the Taylor series up to  $n$ th order) about  $a$ . Now collect all  $j_a^n f$  for all  $a$  and  $f$ , and denote the result by  $J^n(X \times U)$ , so that

$$J^n(X \times U) \equiv \{j_a^n f \mid \forall a \in X \text{ and } \forall f \in \Gamma_a(X \times U)\}. \tag{29.11}$$

$J^n(X \times U)$  is called the  **$n$ th prolongation** of  $U$ , or the  **$n$ th jet space** of  $U$ . It turns out that  $J^n(X \times U)$  is a manifold (see [Saun 89, pp. 98 and 199]).

**29.2.5. Theorem.**  *$J^n(X \times U)$  is a manifold with natural coordinate functions  $(x^i, u^\alpha, u_j^\alpha)$  defined by*

$$x^i(j_a^n f) = a^i, \quad u^\alpha(j_a^n f) = f^\alpha(a), \quad u_j^\alpha(j_a^n f) = \partial_j f^\alpha(a).$$

Note that the “points” of  $J^n(X \times U)$  are  $U$ -valued functions!

The natural coordinate functions allow us to identify the space of the derivatives with various powers of  $\mathbb{R}$ . Let  $U_k \equiv \mathbb{R}^{qp_k}$  denote the set of coordinates  $u_j^\alpha$ , and let

$U^{(n)} \equiv U \times U_1 \times \cdots \times U_n$  be the Cartesian product space<sup>1</sup> whose coordinates represent all the derivatives  $u_j^\alpha$  of all orders from 0 to  $n$ . The dimension of  $U^{(n)}$  is

$$q + qp_1 + \cdots + qp_n = q \binom{p+n}{n} \equiv qp^{(n)}.$$

A typical point in  $U^{(n)}$  is denoted by  $u^{(n)}$ , which has  $qp^{(n)}$  different components  $\{u_j^\alpha\}_{\alpha=1}^q$ , where  $J$  runs over all unordered multi-indices  $J = (j_1, \dots, j_k)$  with  $1 \leq j_k \leq p$  and  $0 \leq k \leq n$ . By convention,  $k = 0$  refers to no derivatives at all, in which case we set  $J = 0$  and  $u_0^\alpha \equiv u^\alpha$ . The  $n$ th jet space  $J^n(X \times U)$  can now be identified with  $X \times U^{(n)}$ . From now on, we shall use  $X \times U^{(n)}$  in place of  $J^n(X \times U)$ .

**29.2.6. Example.** Let  $p = 3$  and  $u = 1$ , i.e.,  $X = \mathbb{R}^3$  and  $U = \mathbb{R}$ . The coordinates of  $X$  are  $(x, y, z)$  and that of  $U$  is  $u$ . The coordinates of  $U_1$  are  $(u_x, u_y, u_z)$ , where the subscript denotes the variable of differentiation. Similarly, the coordinates of  $U_2$  are

$$(u_{xx}, u_{xy}, u_{xz}, u_{yy}, u_{yz}, u_{zz})$$

and those of  $U^{(2)} \equiv U \times U_1 \times U_2$  are

$$(u; u_x, u_y, u_z; u_{xx}, u_{xy}, u_{xz}, u_{yy}, u_{yz}, u_{zz}),$$

which shows that  $U^{(2)}$  is 10-dimensional. ■

prolongation of a  
function

**29.2.7. Definition.** Given a smooth map  $f : X \supset \Omega \rightarrow U$ , we define a map  $\text{pr}^{(n)} f : \Omega \rightarrow U^{(n)}$  whose components  $(\text{pr}^{(n)} f)_J^\alpha$  are given by

$$(\text{pr}^{(n)} f)_J^\alpha(x) \equiv \partial_J f^\alpha(x).$$

This map is called the  $n$ th **prolongation** of  $f$ .

Thus, for each  $x \in X$ ,  $\text{pr}^{(n)} f(x)$  is a vector in  $\mathbb{R}^{qp^{(n)}}$  whose components are the values of  $f$  and all its derivatives up to order  $n$  at the point  $x$ . For example, in the case of  $p = 3, q = 1$  discussed above,  $\text{pr}^{(2)} f(x, y, z)$  has components

$$\left( f; \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}; \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial x \partial z}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial y \partial z}, \frac{\partial^2 f}{\partial z^2} \right).$$

When the underlying space is an open subset  $M$  of  $X \times U$ , its corresponding jet space is

$$M^{(n)} \equiv M \times U_1 \times \cdots \times U_n,$$

which is a subspace of  $X \times U^{(n)} \cong J^n(X \times U)$ . If the graph of  $f : X \rightarrow U$  lies in  $M$ , the graph of  $\text{pr}^{(n)} f$  lies in  $M^{(n)}$ .

<sup>1</sup>Note that  $U$ , identified with the space of zeroth derivative, is a factor in  $U^{(n)}$ .

Prolongation allows us to turn a system of DEs into a system of *algebraic* equations: Given a system of  $l$  DEs

$$\Delta_\nu (\{x^i\}, \{u^\alpha\}, \{\partial_i u^\alpha\}, \{\partial_i \partial_j u^\alpha\}, \dots, \{\partial_{i_1} \dots \partial_{i_n} u^\alpha\}) = 0, \quad \nu = 1, \dots, l,$$

one can define a map  $\Delta : M^{(n)} \rightarrow \mathbb{R}^l$  and identify the system of DEs with

$$\mathcal{S}_\Delta \equiv \{(x, u^{(n)}) \in M^{(n)} \mid \Delta(x, u^{(n)}) = 0\}.$$

By identifying the system of DEs with the subset  $\mathcal{S}_\Delta$  of the jet space, we have translated the abstract relations among the derivatives of  $u$  into a geometrical object  $\mathcal{S}_\Delta$ , which is more amenable to symmetry operations.

**29.2.8. Definition.** Let  $\Omega$  be a subset of  $X$  and  $f : \Omega \rightarrow U$  a smooth map. Then  $f$  is called a *solution* of the system of DEs  $\mathcal{S}_\Delta$  if

$$\Delta(x, \text{pr}^{(n)} f(x)) = 0 \quad \forall x \in \Omega.$$

Just as we identified a function with its graph, we can identify the solution of a system of DEs with the graph of its prolongation  $\text{pr}^{(n)} f$ . This graph, which is denoted by  $\Gamma_f^{(n)}$ , will clearly be a subset of  $\mathcal{S}_\Delta$ :

$$\Gamma_f^{(n)} \equiv \{(x, \text{pr}^{(n)} f(x))\} \subset \mathcal{S}_\Delta.$$

**29.2.9. Box.** An  $n$ th order system of differential equations is taken to be a subset  $\mathcal{S}_\Delta$  of the jet space  $J^n(X \times U)$ , and a solution to be a smooth map  $f : \Omega \rightarrow J^n(X \times U)$  the graph of whose  $n$ th prolongation  $\text{pr}^{(n)} f$  is contained in  $\mathcal{S}_\Delta$ .

**29.2.10. Example.** Consider Laplace's equation

$$\nabla^2 u = u_{xx} + u_{yy} + u_{zz} = 0$$

with  $p = 3$ ,  $q = 1$ , and  $n = 2$ . The total jet space is the 13-dimensional Euclidean space  $X \times U^{(2)}$ , whose coordinates are taken to be

$$(x, y, z; u; u_x, u_y, u_z; u_{xx}, u_{xy}, u_{xz}, u_{yy}, u_{yz}, u_{zz}).$$

In this 13-dimensional Euclidean space, Laplace's equation defines a 12-dimensional subspace  $\mathcal{S}_\Delta$  consisting of all points in the jet space whose eighth, eleventh, and thirteenth coordinates add up to zero. A solution  $f : \mathbb{R}^3 \supset \Omega \rightarrow U \subset \mathbb{R}$  must satisfy

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0 \quad \forall (x, y, z) \in \Omega.$$

This is the same as requiring the graph  $\Gamma_f^{(2)}$  to lie in  $S_\Delta$ . For example, if

$$f(\mathbf{r}) \equiv f(x, y, z) = x^3yz - y^3xz + y^3 - 3yz^2,$$

then, collecting each section of  $\text{pr}^{(2)}f$  with fixed  $\alpha$  (see Definition 29.2.7) separately, we have

$$(\text{pr}^{(2)}f)^1(\mathbf{r}) = x^3yz - y^3xz + y^3 - 3yz^2,$$

$$(\text{pr}^{(2)}f)^2(\mathbf{r}) = (3x^2yz - y^3z, x^3z - 3y^2xz + 3y^2 - 3z^2, x^3y - y^3x - 6yz),$$

$$(\text{pr}^{(2)}f)^3(\mathbf{r}) = (6xyz, 3x^2z - 3y^2z, 3x^2y - y^3, -6xyz + 6y, x^3 - 3xy^2 - 6z, -6y)$$

which lies in  $S_\Delta$  because the sum of the eighth, the eleventh, and the thirteenth coordinates of  $(x, y, z; \text{pr}^{(2)}f(x, y, z))$  is  $6xyz - 6xyz + 6y - 6y = 0$ . ■

### 29.2.2 Prolongation of Groups

*n*th prolongation of a group action

Suppose  $G$  is a group of transformations acting on  $M \subset X \times U$ . It is possible to prolong this action to the  $n$ -jet space  $M^{(n)}$ . The resulting group that acts on  $M^{(n)}$  is called the  $n$ th **prolongation** of  $G$  and denoted by  $\text{pr}^{(n)}G$  with group elements  $\text{pr}^{(n)}g$ , for  $g \in G$ . This prolongation is defined naturally: The derivatives of a function  $f$  with respect to  $x$  are transformed into derivatives of  $\tilde{f} = g \cdot f$  with respect to  $\tilde{x}$ . More precisely,

$$\boxed{\text{pr}^{(n)}g \cdot (j_x^n f) \equiv (\tilde{x}, \text{pr}^{(n)}\tilde{f}(\tilde{x})) \equiv (\tilde{x}, \text{pr}^{(n)}(g \cdot f)(\tilde{x}))}. \tag{29.12}$$

For  $n = 0$ , Equation (29.12) reduces to the action of  $G$  on  $M$  as given by Equation (29.8). That the outcome of the action of the prolongation of  $G$  in the defining equation (29.12) is independent of the representative function  $f$  follows from the chain rule and the fact that only derivatives up to the  $n$ th order are involved. The following example illustrates this.

**29.2.11. Example.** Let  $X, U$ , and  $G$  be as in Example 29.2.1. In this case, we have

$$\text{pr}^{(1)}\theta \cdot (j_x^1 f) \equiv \text{pr}^{(1)}\theta \cdot (x, u, u_1) \equiv (\tilde{x}, \tilde{u}, \tilde{u}_1).$$

We calculated  $\tilde{x}$  and  $\tilde{u}$  in that example. They are

$$\begin{aligned} \tilde{x} &= x \cos \theta - u \sin \theta, \\ \tilde{u} &= x \sin \theta + u \cos \theta. \end{aligned} \tag{29.13}$$

To find  $\tilde{u}_1$ , we need to differentiate the second equation with respect to  $\tilde{x}$  and express the result in terms of the original variables. Thus

$$\tilde{u}_1 \equiv \frac{d\tilde{u}}{d\tilde{x}} = \frac{d\tilde{u}}{dx} \frac{dx}{d\tilde{x}} = \left( \sin \theta + \frac{du}{dx} \cos \theta \right) \frac{dx}{d\tilde{x}} = (\sin \theta + u_1 \cos \theta) \frac{dx}{d\tilde{x}};$$

$dx/d\tilde{x}$  is obtained by differentiating the first equation of (29.13):

$$1 = \frac{dx}{d\tilde{x}} \cos \theta - \frac{du}{d\tilde{x}} \sin \theta = \frac{dx}{d\tilde{x}} \cos \theta - \frac{du}{dx} \frac{dx}{d\tilde{x}} \sin \theta = (\cos \theta - u_1 \sin \theta) \frac{dx}{d\tilde{x}},$$

or

$$\frac{dx}{d\bar{x}} = \frac{1}{\cos \theta - u_1 \sin \theta}.$$

It therefore follows that

$$\bar{u}_1 = \frac{\sin \theta + u_1 \cos \theta}{\cos \theta - u_1 \sin \theta}$$

and

$$\text{pr}^{(1)}\theta \cdot (x, u, u_1) = \left( x \cos \theta - u \sin \theta, x \sin \theta + u \cos \theta, \frac{\sin \theta + u_1 \cos \theta}{\cos \theta - u_1 \sin \theta} \right).$$

We note that the RHS involves derivatives up to order one. Therefore, the transformation is independent of the representative function. So, if we had chosen  $j_x^1 h$  where  $h$  is 1-equivalent to  $f$ , we would have obtained the same result. This holds for derivatives of all orders. Therefore, the prolongation of the action of the group  $G$  is well-defined. ■

In many cases, it is convenient to choose the  $n$ th-order Taylor polynomial as the representative of the class of  $n$ -equivalent functions, and, if possible, write the transformed function  $\bar{f}$  explicitly in terms of  $\bar{x}$ , and differentiate it to obtain the transformed derivatives (see Problem 29.3).

Example 29.2.11 illustrates an important property of the prolongation of  $G$ . We note that the first prolongation  $\text{pr}^{(1)}G$  acts on the original coordinates  $(x, u)$  in exactly the same way that  $G$  does. This holds in general:

**29.2.12. Box.** *The effect of the  $n$ th prolongation  $\text{pr}^{(n)}G$  to derivatives up to order  $m \leq n$  is exactly the same as the effect of  $\text{pr}^{(m)}G$ . If we already know the action of the  $m$ th-order prolonged group  $\text{pr}^{(m)}G$ , then to compute  $\text{pr}^{(n)}G$  we need only find how the derivatives  $u_j^\alpha$  of order higher than  $m$  transform, because the lower-order action is already determined.*

### 29.2.3 Prolongation of Vector Fields

The geometrization of a system of DEs makes it possible to use the machinery of differentiable manifolds, Lie groups, and Lie algebras to unravel the symmetries of the system. At the heart of this machinery are the infinitesimal transformations, which are directly connected to vector fields. Therefore, it is necessary to find out how a vector field defined in  $M \subset X \times U$  is prolonged. The most natural way to prolong a vector field is to prolong its integral curve—which is a one-parameter group of transformations of  $M$ —to a curve in  $M^{(n)}$  and then calculate the tangent to the latter curve.

*n*th prolongation of a vector field

**29.2.13. Definition.** Let  $M$  be an open subset of  $X \times U$  and  $\mathbf{X} \in \mathfrak{X}(M)$ . The *n*th prolongation of  $\mathbf{X}$ , denoted by  $\text{pr}^{(n)}\mathbf{X}$ , is a vector field on the *n*th jet space  $M^{(n)}$  defined by

$$\text{pr}^{(n)}\mathbf{X}|_{(x, u^{(n)})} = \left. \frac{d}{dt} \text{pr}^{(n)}[\exp(t\mathbf{X})] \cdot (x, u^{(n)}) \right|_{t=0}$$

for any  $(x, u^{(n)}) \in M^{(n)}$ .

Since  $(x, u^{(n)}) \in M^{(n)}$  form a coordinate system on  $M^{(n)}$ , any vector field in  $M^{(n)}$  can be written as a linear combination of  $\partial/\partial x^i$  and  $\partial/\partial u_j^\alpha$  with coefficients being, in general, functions of all coordinates  $x^i$  and  $u_j^\alpha$ . For a prolonged vector, however, we have

$$\text{pr}^{(n)}\mathbf{X} = \sum_{i=1}^p X^i \frac{\partial}{\partial x^i} + \sum_{\alpha=1}^q \sum_J X_J^\alpha \frac{\partial}{\partial u_J^\alpha}, \tag{29.14}$$

where  $X^i$  and  $X_0^\alpha$  are functions only of  $x^k$  and  $u$ . This is due to the remark made in Box 29.2.12. For the same reason, the coefficients  $X_J^\alpha$  corresponding to derivatives of order  $m$  will be independent of coordinates  $u_j^\alpha$  that involve derivatives of order higher than  $m$ . Thus, it is possible to construct various prolongations of a given vector field recursively.

**29.2.14. Example.** Let us consider our recurrent example of  $X \cong U \cong \mathbb{R}$ ,  $G = SO(2)$ . Given the infinitesimal generator  $\mathbf{v} = -u\partial_x + x\partial_u$ , one can solve the DE of its integral curve to obtain<sup>2</sup>

$$\exp(t\mathbf{v})(x, u) = (x \cos t - u \sin t, x \sin t + u \cos t).$$

Example 29.2.11 calculated the first prolongation of  $SO(2)$ . So

$$\text{pr}^{(1)} \exp(t\mathbf{v}) \cdot (x, u, u_1) = \left( x \cos t - u \sin t, x \sin t + u \cos t, \frac{\sin t + u_1 \cos t}{\cos t - u_1 \sin t} \right).$$

Differentiating the components with respect to  $t$  at  $t = 0$  gives

$$\begin{aligned} \left. \frac{\partial}{\partial t} (x \cos t - u \sin t) \right|_{t=0} &= -u, \\ \left. \frac{\partial}{\partial t} (x \sin t + u \cos t) \right|_{t=0} &= x, \\ \left. \frac{\partial}{\partial t} \left( \frac{\sin t + u_1 \cos t}{\cos t - u_1 \sin t} \right) \right|_{t=0} &= 1 + u_1^2. \end{aligned}$$

Therefore,

$$\text{pr}^{(1)}\mathbf{v} = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u} + (1 + u_1^2) \frac{\partial}{\partial u_1}.$$

Note that the first two terms in  $\text{pr}^{(1)}\mathbf{v}$  are the same as those in  $\mathbf{v}$  itself, in agreement with Box 29.2.12. ■

<sup>2</sup>One can, of course, also write the finite group element directly.



### 29.3 The Central Theorems

We are now in a position to state the first central theorem of the application of Lie groups to the solution of DEs. This theorem is the exact replica of Theorem 29.1.6 in the language of prolongations:

**29.3.1. Theorem.** *Let*

$$\Delta_v(x, u^{(n)}) = 0, \quad v = 1, \dots, l,$$

*be a system of DEs defined on  $M \subset X \times U$  whose Jacobian matrix*

$$J_\Delta(x, u^{(n)}) \equiv \left( \frac{\partial \Delta_v}{\partial x^i}, \frac{\partial \Delta_v}{\partial u_j^\alpha} \right)$$

*has rank  $l$  for all  $(x, u^{(n)}) \in S_\Delta$ . If  $G$  is a local Lie group of transformations acting on  $M$ , and*

$$\text{pr}^{(n)} \xi_M|_{(x, u^{(n)})} \Delta_v = 0, \quad v = 1, \dots, l, \quad \text{whenever } \Delta(x, u^{(n)}) = 0$$

*for every infinitesimal generator  $\xi$  of  $G$ , then  $G$  is the symmetry group of the system.*

**29.3.2. Example.** Consider the first order (so,  $n = 1$ ) ordinary DE

$$\Delta(x, u, u_1) = (u - x^3 - u^2x)u_1 + x + x^2u + u^3 = 0,$$

so that  $X \cong \mathbb{R}$  and  $U \cong \mathbb{R}$ . We first note that

$$\begin{aligned} J_\Delta &= \left( \frac{\partial \Delta}{\partial x}, \frac{\partial \Delta}{\partial u}, \frac{\partial \Delta}{\partial u_1} \right) \\ &= ((-3x^2 - u^2)u_1 + 1 + 2xu, (1 - 2ux)u_1 + x^2 + 3u^2, u - x^3 - u^2x), \end{aligned}$$

which is of rank 1 everywhere. Now let us apply the first prolongation of the generator of  $SO(2)$ —calculated in Example 29.2.14—to  $\Delta$ . We have

$$\begin{aligned} \text{pr}^{(1)}v(\Delta) &= -u \frac{\partial \Delta}{\partial x} + x \frac{\partial \Delta}{\partial u} + (1 + u_1^2) \frac{\partial \Delta}{\partial u_1} \\ &= -u[(-3x^2 - u^2)u_1 + 1 + 2xu] + x[(1 - 2ux)u_1 + x^2 + 3u^2] \\ &\quad + (1 + u_1^2)(u - x^3 - u^2x) \\ &= u_1[(u - x^3 - u^2x)u_1 + x + x^2u + u^3] = u_1 \Delta. \end{aligned}$$

It follows that  $\text{pr}^{(1)}v(\Delta) = 0$  whenever  $\Delta = 0$ , and that  $SO(2)$  is a symmetry group of the DE. Thus, rotations will change solutions of the DE into other solutions. In fact, the reader may verify that in polar coordinates, the DE can be written in the incredibly simple form

$$\frac{dr}{d\theta} = r^3,$$

and the symmetry of the DE conveys the fact that adding a constant to  $\theta$  does not change the polar form of the DE. ■

Theorem 29.3.1 reduces the invariance of a system of DEs to a criterion involving the prolongation of the infinitesimal generators of the symmetry group. The urgent task in front of us is therefore to construct an explicit formula for the prolongation of a vector field. In order to gain insight into this construction, we first look at the simpler case of  $U \cong \mathbb{R}$  and a group  $G$  that transforms only the independent variables. Furthermore, we restrict ourselves to the first prolongation. An infinitesimal generator of such a group will be of the form

$$v = \sum_{i=1}^p X^i(x) \frac{\partial}{\partial x^i},$$

which is assumed to act on the space  $M \subset X \times U$ . The integral curve of this vector field is  $\exp(tv)$  which acts on points of  $M$  as follows:

$$(\tilde{x}, \tilde{u}) = \exp(tv) \cdot (x, u) \equiv (\Psi_t(x), u) \equiv (\Psi_t^i(x), u).$$

By the construction of the integral curves in general, we have

$$\left. \frac{d\Psi_t^i(x)}{dt} \right|_{t=0} = X^i(x). \tag{29.15}$$

Denote the coordinates of the first jet space  $M^{(1)}$  by  $(x^i, u, u_k)$ , where  $u_k(j_x^1 f) \equiv \partial f / \partial x^k$ . By the definition of the action of the prolonged group,

$$pr^{(1)} \exp(tv) \cdot (j_x^1 f) = (\Psi_t(x), u, \tilde{u}_j), \tag{29.16}$$

where  $u = f(x)$  and  $\tilde{u}_j = \partial \tilde{f} / \partial \tilde{x}^j$ . Once we find  $\tilde{u}_j$ , we can differentiate Equation (29.16) with respect to  $t$  at  $t = 0$  to obtain the prolonged vector field. Using Equation (29.10) and commas to indicate differentiation,<sup>3</sup> we obtain

$$\tilde{u}_j = \tilde{f}_{,j}(\tilde{x}) = (f \circ \Psi_{-t})_{,j}(\tilde{x}) = \sum_{i=1}^p f_{,i}(\underbrace{\Psi_{-t}(\tilde{x})}_{=x}) \Psi_{-t,i,j}^i(\tilde{x}) = \sum_{i=1}^p u_i \Psi_{-t,i,j}^i(\tilde{x}).$$

Since  $v$  does not have any component along  $\partial/\partial u$ , its prolongation will not have such components either. The components  $U_j$  along  $\partial/\partial u_j$  are obtained by differ-

<sup>3</sup>We have found it exceedingly convenient to use commas to indicate differentiation with respect to the  $x$ 's. The alternative, i.e., the use of partials, makes it almost impossible to find one's way in the maze of derivatives involving  $x^i$ ,  $\tilde{x}^j$ , and  $t$ , with  $\tilde{x}^j$  depending on  $t$ . The reader will recall that the index after the comma is to be thought of as a "position holder," and the argument as a substitution. Thus, for example,

$$\tilde{f}_{,j}(\tilde{x}) \equiv \left. \frac{\partial \tilde{f}(r_1, \dots, r_p)}{\partial r_j} \right|_{r=\tilde{x}} \equiv \left. \frac{\partial \tilde{f}(s_1, \dots, s_p)}{\partial s_j} \right|_{s=\tilde{x}} \equiv \left. \frac{\partial \tilde{f}(\heartsuit_1, \dots, \heartsuit_p)}{\partial \heartsuit_j} \right|_{\heartsuit=\tilde{x}}$$

entiating  $\tilde{u}_j$  with respect to  $t$ :

$$\begin{aligned} U_j(x, u, u_j) &= \sum_{i=1}^p u_i \frac{d}{dt} \Psi_{-t,j}^i(\tilde{x}) \Big|_{t=0} \\ &= \sum_{i=1}^p u_i \frac{\partial \Psi_{-t,j}^i(\tilde{x})}{\partial t} \Big|_{t=0} + \sum_{i=1}^p u_i \sum_{k=1}^p \left( \Psi_{-t,jk}^i(\tilde{x}) \frac{\partial \tilde{x}^k}{\partial t} \right) \Big|_{t=0}. \end{aligned} \quad (29.17)$$

The derivative of the first term in the sum can be evaluated as follows:

$$\begin{aligned} \frac{\partial \Psi_{-t,j}^i(\tilde{x}(t))}{\partial t} \Big|_{t=0} &= - \frac{\partial \Psi_{s,j}^i(\tilde{x}(-s))}{\partial s} \Big|_{s=0} = - \frac{\partial \Psi_{s,j}^i(s, \tilde{x}(-s))}{\partial s} \Big|_{s=0} \\ &= - \frac{\partial \Psi_{s,j}^i(s, \tilde{x}(0))}{\partial s} \Big|_{s=0} = - \frac{\partial \Psi_{s,j}^i(s, x)}{\partial s} \Big|_{s=0} \\ &= - \frac{\partial \Psi_{s,j}^i(x)}{\partial s} \Big|_{s=0} = - \left( \frac{\partial \Psi_s^i}{\partial s} \Big|_{s=0} \right)_{,j} (x) = -X_{,j}^i(x), \end{aligned}$$

where we have emphasized the dependence of  $\tilde{x}$  on  $t$  (or  $s$ ), treated  $s$  as the first independent variable, and in the second line substituted  $s = 0$  in all  $\tilde{x}$ 's before differentiation. This is possible because we are taking the *partial* derivative with respect to the first variable holding all others constant. The derivative of the second term in Equation (29.17) can be calculated similarly:

$$\Psi_{-t,jk}^i(\tilde{x}(t)) \Big|_{t=0} = \Psi_{0,jk}^i(\tilde{x}(0)) = \frac{\partial^2 x^i}{\partial x^j \partial x^k} = 0$$

because  $\Psi_0^i = x^i$ . We therefore have

$$U_j(x, u, u_k) = - \sum_{i=1}^p \frac{\partial X^i}{\partial x^j} u_i \quad (29.18)$$

and

$$\text{pr}^{(1)}\mathbf{v} = \sum_{i=1}^p \left( X^i \frac{\partial}{\partial x^i} + U_i \frac{\partial}{\partial u_i} \right) = \mathbf{v} - \sum_{i,k=1}^p \frac{\partial X^k}{\partial x^i} u_k \frac{\partial}{\partial u_i}. \quad (29.19)$$

It is also instructive to consider the case in which  $U$  is still  $\mathbb{R}$ , but  $G$  acts only on the dependent variable. Then  $\mathbf{v} = U(x, u)\partial_u$ , and

$$(\tilde{x}, \tilde{u}) = (x, \Phi_t(x, u)), \quad \text{with} \quad U(x, u) = \frac{d\Phi_t(x, u)}{dt} \Big|_{t=0}.$$

The reader may check that in this case, the prolongation of  $\mathbf{v}$  is given by

$$\text{pr}^{(1)}\mathbf{v} = \mathbf{v} + \sum_{j=1}^p U_j(x, u^{(1)}) \frac{\partial}{\partial u_j}, \quad U_j(x, u^{(1)}) = \frac{\partial U}{\partial x^j} + u_j \frac{\partial U}{\partial u}. \quad (29.20)$$

The second equation in (29.20) can also be written as

$$U_j(x, \text{pr}^{(1)} f(x)) = \frac{\partial U}{\partial x^j} + \frac{\partial U}{\partial u} \frac{\partial f}{\partial x^j} = \frac{\partial}{\partial x^j} [U(x, f(x))].$$

In other words,  $U_j(x, u^{(1)})$  is obtained from  $U(x, u)$  by differentiation with respect to  $x^j$ , while treating  $u$  as a function of  $x$ . This leads us to the definition of the total derivative.

**29.3.3. Definition.** Let  $S : M^{(n)} \rightarrow \mathbb{R}$  be a smooth function of  $x, u$ , and all total derivative derivatives of  $u$  up to  $n$ th order. The **total derivative** of  $S$  with respect to  $x^i$ , denoted by  $D_i S$ , is a smooth function  $D_i S : M^{(n+1)} \rightarrow \mathbb{R}$  defined by

$$D_i S(j_x^{n+1} f) = \frac{\partial}{\partial x^i} [S(x, \text{pr}^{(n)} f(x))];$$

i.e.,  $D_i S$  is obtained from  $S$  by differentiating  $S$  with respect to  $x^i$ , treating  $u$  and all the  $u^\alpha$ 's as functions of  $x$ .

The following proposition, whose proof is a straightforward application of the chain rule, gives the explicit formula for calculating the total derivative:

**29.3.4. Proposition.** The  $i$ th total derivative of  $S : M^{(n)} \rightarrow \mathbb{R}$  is of the form

$$D_i S = \frac{\partial S}{\partial x^i} + \sum_{\alpha=1}^q \sum_J u_{J,i}^\alpha \frac{\partial S}{\partial u_J^\alpha}$$

where, for  $J = (j_1, \dots, j_k)$ ,

$$u_{J,i}^\alpha \equiv \frac{\partial u_J^\alpha}{\partial x^i} = \frac{\partial^{k+1} u^\alpha}{\partial x^i \partial x^{j_1} \dots \partial x^{j_k}}$$

and the sum over  $J$  includes derivatives of all orders from 0 to  $n$ .

An immediate consequence of this proposition is

$$D_i u_J^\alpha = u_{J,i}^\alpha = \frac{\partial u_J^\alpha}{\partial x^i} \quad \forall J, \alpha,$$

$$D_i (ST) = T D_i S + S D_i T. \tag{29.21}$$

Higher-order total derivatives are defined in analogy with partial derivatives: If  $I$  is a multi-index of the form  $I = (i_1, \dots, i_k)$ , then the  $I$ th total derivative is

$$D_I S = D_{i_1} D_{i_2} \dots D_{i_k} S. \tag{29.22}$$

As in the case of partial derivatives, the order of differentiation is immaterial.

We are now ready to state the second central theorem of the application of Lie groups to the solution of DEs (for a proof, see [Olvé 86, pp. 113–115]).

**29.3.5. Theorem.** *Let*

$$\mathbf{v} = \sum_{i=1}^p X^i(x, u) \frac{\partial}{\partial x^i} + \sum_{\alpha=1}^q U^\alpha(x, u) \frac{\partial}{\partial u^\alpha}$$

be a vector field on an open subset  $M \subset X \times U$ . The  $n$ th prolongation of  $\mathbf{v}$ , i.e.,  $\text{pr}^{(n)}\mathbf{v} \in \mathcal{X}(M^{(n)})$ , is

$$\text{pr}^{(n)}\mathbf{v} = \mathbf{v} + \sum_{\alpha=1}^q \sum_J U_J^\alpha(x, u^{(n)}) \frac{\partial}{\partial u_J^\alpha},$$

where for  $J = (j_1, \dots, j_k)$ , the inner sum extends over  $1 \leq |J| \leq n$  and the coefficients  $U_J^\alpha$  are given by

$$U_J^\alpha(x, u^{(n)}) = D_J \left( U^\alpha - \sum_{i=1}^p X^i \frac{\partial u^\alpha}{\partial x^i} \right) + \sum_{i=1}^p X^i \frac{\partial u_J^\alpha}{\partial x^i}$$

and the higher-order derivative  $D_J$  is as given in Equation (29.22).

**29.3.6. Example.** Let  $p = 2$ ,  $q = 1$ , and consider the case in which  $G$  acts only on the independent variables  $(x, y)$ . The general vector field for this situation is

$$\mathbf{v} = \xi(x, y) \frac{\partial}{\partial x} + \eta(x, y) \frac{\partial}{\partial y}.$$

We are interested in the first prolongation of this vector field. Thus,  $n = 1$ ,  $X^1 = \xi$ ,  $X^2 = \eta$ , and  $J$  has only one component, which we denote by  $j$  (also written as  $x$  or  $y$ ). Theorem 29.3.5 gives

$$U_J^\alpha \equiv U_j = -D_j \sum_{i=1}^2 X^i \frac{\partial u}{\partial x^i} + \sum_{i=1}^2 X^i \frac{\partial u_j}{\partial x^i} = - \sum_{i=1}^2 \frac{\partial X^i}{\partial x^j} \frac{\partial u}{\partial x^i},$$

and using the notation  $u_x = \partial u / \partial x$  and  $u_y = \partial u / \partial y$ , we obtain

$$\text{pr}^{(1)}\mathbf{v} = \mathbf{v} + U_x \frac{\partial}{\partial u_x} + U_y \frac{\partial}{\partial u_y}, \quad (29.23)$$

where

$$U_x = -u_x \frac{\partial \xi}{\partial x} - u_y \frac{\partial \eta}{\partial x} \quad \text{and} \quad U_y = -u_x \frac{\partial \xi}{\partial y} - u_y \frac{\partial \eta}{\partial y}.$$

In particular, if  $G = SO(2)$ , so that  $\mathbf{v} = -y\partial_x + x\partial_y$ , then  $U_x = -u_y$  and  $U_y = u_x$ . It then follows that

$$\text{pr}^{(1)}\mathbf{v} = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} - u_y \frac{\partial}{\partial u_x} + u_x \frac{\partial}{\partial u_y}. \quad \blacksquare$$

**29.3.7. Example.** Let  $p = 1, q = 1$ , and  $G = SO(2)$ . The general vector field for this situation is

$$\mathbf{v} = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u}.$$

For the first prolongation of this vector field  $n = 1, X^1 = -u$ , and  $J$  has only one component, which we denote by  $x$ . Theorem 29.3.5 gives

$$U_j^\alpha \equiv U_x = D_x(x - X^1 u_x) + X^1 u_{xx} = 1 - \frac{\partial X^1}{\partial x} u_x = 1 + u_x^2.$$

It follows that

$$\text{pr}^{(1)}\mathbf{v} = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u} + (1 + u_x^2) \frac{\partial}{\partial u_x},$$

which is the result obtained in Example 29.2.14.

The second prolongation can be obtained as well. Once again we use Theorem 29.3.5 with obvious change of notation:

$$U_{xx} = D_x D_x(x - X^1 u_x) + X^1 u_{xxx} = D_x(1 + u_x^2 + u u_{xx}) - u u_{xxx} = 3u_x u_{xx}.$$

Then

$$\text{pr}^{(2)}\mathbf{v} = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u} + (1 + u_x^2) \frac{\partial}{\partial u_x} + 3u_x u_{xx} \frac{\partial}{\partial u_{xx}}.$$

Using Theorem 29.3.1, we note that the DE  $u_{xx} = 0$  has  $SO(2)$  as a symmetry group, because with  $\Delta(x, u, u_x, u_{xx}) \equiv u_{xx}$ ,

$$\text{pr}^{(2)}\mathbf{v}(\Delta) = -u \frac{\partial \Delta}{\partial x} + x \frac{\partial \Delta}{\partial u} + (1 + u_x^2) \frac{\partial \Delta}{\partial u_x} + 3u_x u_{xx} \frac{\partial \Delta}{\partial u_{xx}} = 3u_x u_{xx},$$

which vanishes whenever  $\Delta(x, u, u_x, u_{xx})$  vanishes. This is the statement that rotations take straight lines to straight lines. ■

## 29.4 Application to Some Known PDEs

We have all the tools at our disposal to compute (in principle) the most general symmetry group of almost any system of PDEs. The coefficients  $U_j^\alpha$  of the prolonged vector field  $\text{pr}^{(n)}\mathbf{v}$  will be functions of the partial derivatives of the coefficients  $X^i$  and  $U^\alpha$  of  $\mathbf{v}$  with respect to both  $x$  and  $u$ . The infinitesimal criterion of invariance as given in Theorem 29.3.1 will involve  $x, u$ , and the derivatives of  $u$  with respect to  $x$ , as well as  $X^i$  and  $U^\alpha$  and their partial derivatives with respect to  $x$  and  $u$ . Using the system of PDEs, we can obtain some of the derivatives of the  $u$ 's in terms of the others. Substituting these relations in the equation of infinitesimal criterion, we get an equation involving  $u$ 's and powers of its derivatives that are to be treated as independent. We then equate the coefficients of these powers of partial derivatives of  $u$  to zero. This will result in a large number of elementary

defining equations  
for the symmetry  
group of a system of  
PDEs

PDEs for the coefficient functions  $X^i$  and  $U^\alpha$  of  $\mathbf{v}$ , called the **defining equations** for the symmetry group of the given system of PDEs. In most applications, these defining equations can be solved, and the general solution will determine the most general infinitesimal symmetry of the system. The symmetry group itself can then be calculated by exponentiation of the vector fields, i.e., by finding their integral curves. In the remaining part of this section, we construct the symmetry groups of the heat and the wave equations.

### 29.4.1 The Heat Equation

The one-dimensional heat equation  $u_t = u_{xx}$  corresponds to  $p = 2, q = 1$ , and  $n = 2$ . So it is determined by the vanishing of  $\Delta(x, t, u^{(2)}) = u_t - u_{xx}$ . The most general infinitesimal generator of symmetry appropriate for this equation can be written as

$$\mathbf{v} = \xi(x, t, u) \frac{\partial}{\partial x} + \tau(x, t, u) \frac{\partial}{\partial t} + \phi(x, t, u) \frac{\partial}{\partial u}, \tag{29.24}$$

which, as the reader may check (see Problem 29.11), has a second prolongation of the form

$$\text{pr}^{(2)}\mathbf{v} = \mathbf{v} + \phi^x \frac{\partial}{\partial u_x} + \phi^t \frac{\partial}{\partial u_t} + \phi^{xx} \frac{\partial}{\partial u_{xx}} + \phi^{xt} \frac{\partial}{\partial u_{xt}} + \phi^{tt} \frac{\partial}{\partial u_{tt}},$$

where, for example,

$$\begin{aligned} \phi^t &= \phi_t - \xi_t u_x + (\phi_u - \tau_t) u_t - \xi_u u_x u_t - \tau_u u_t^2 \\ \phi^{xx} &= \phi_{xx} + (2\phi_{xu} - \xi_{xx}) u_x - \tau_{xx} u_t + (\phi_{uu} - 2\xi_{xu}) u_x^2 \\ &\quad - 2\tau_{xu} u_x u_t - \xi_{uu} u_x^3 - \tau_{uu} u_x^2 u_t \\ &\quad + (\phi_u - 2\xi_x) u_{xx} - 2\tau_x u_{xt} - 3\xi_u u_x u_{xx} - \tau_u u_t u_{xx} - 2\tau_u u_x u_{xt}, \end{aligned} \tag{29.25}$$

and subscripts indicate partial derivatives. Theorem 29.3.1 now gives

$$\text{pr}^{(2)}\mathbf{v}(\Delta) = \phi^t - \phi^{xx} = 0 \quad \text{whenever} \quad u_t = u_{xx} \tag{29.26}$$

as the infinitesimal criterion. Substituting (29.25) in (29.26), replacing  $u_t$  with  $u_{xx}$  in the resulting equation, and equating to zero the coefficients of the monomials in derivatives of  $u$ , we obtain a number of equations involving  $\xi, \tau$ , and  $\phi$ . These equations as well as the monomials of which they are coefficients are given in Table 29.1. Complicated as the defining equations may look, they are fairly easy to solve. From (d) and (f) we conclude that  $\tau$  is a function of  $t$  only. Then (c) shows that  $\xi$  is independent of  $u$ , and (e) gives  $2\xi_x = \tau_t$ , or  $\xi(x, t) = \frac{1}{2}\tau_t x + \eta(t)$ , for some arbitrary function  $\eta$ . From (h) and the fact that  $\xi$  is independent of  $u$  we get  $\phi_{uu} = 0$ , or

$$\phi(x, t, u) = \alpha(x, t)u + \beta(x, t)$$

Monomial	Coefficient Equation	
$u_{xx}^2$	$0 = 0$	(a)
$u_x^2 u_{xx}$	$\tau_{uu} = 0$	(b)
$u_x u_{xx}$	$2\xi_u + 2\tau_{xu} = 0$	(c)
$u_x u_{xt}$	$2\tau_u = 0$	(d)
$u_{xx}$	$2\xi_x + \tau_{xx} - \tau_t = 0$	(e)
$u_{xt}$	$2\tau_x = 0$	(f)
$u_x^3$	$\xi_{uu} = 0$	(g)
$u_x^2$	$2\xi_{xu} - \phi_{uu} = 0$	(h)
$u_x$	$\xi_{xx} - 2\phi_{xu} - \xi_t = 0$	(i)
1	$\phi_t - \phi_{xx} = 0$	(j)

**Table 29.1** The defining equations of the heat equation and the monomials that give rise to them.

for some as yet undetermined functions  $\alpha$  and  $\beta$ . Since  $\xi$  is linear in  $x$ ,  $\xi_{xx} = 0$ , and (i) yields  $\xi_t = -2\phi_{xu} = -2\alpha_x$ , or

$$\frac{\partial \alpha}{\partial x} = -\frac{1}{2}\xi_t = -\frac{1}{4}\tau_{tt}x - \frac{1}{2}\eta_t \Rightarrow \alpha(x, t) = -\frac{1}{8}\tau_{tt}x^2 - \frac{1}{2}\eta_t x + \rho(t).$$

Finally, with  $\phi_t = \alpha_t u + \beta_t$  and  $\phi_{xx} = \alpha_{xx}u + \beta_{xx}$  (recall that when taking partial derivatives,  $u$  is considered independent of  $x$  and  $t$ ), the last defining equation (j) gives  $\alpha_t = \alpha_{xx}$  and  $\beta_t = \beta_{xx}$ , i.e., that  $\alpha$  and  $\beta$  are to satisfy the heat equation. Substituting  $\alpha$  in the heat equation, we obtain

$$-\frac{1}{8}\tau_{ttt}x^2 - \frac{1}{2}\eta_{tt}x + \rho_t = -\frac{1}{4}\tau_{tt},$$

which must hold for all  $x$  and  $t$ . Therefore,

$$\tau_{ttt} = 0, \quad \eta_{tt} = 0, \quad \rho_t = -\frac{1}{4}\tau_{tt}.$$

These equations have the solution

$$\tau = c_1 t^2 + c_2 t + c_3, \quad \rho = -\frac{1}{2}c_1 t + c_4, \quad \eta = c_5 t + c_6.$$

It follows that  $\alpha(x, t) = -\frac{1}{4}c_1 x^2 - \frac{1}{2}c_5 x - \frac{1}{2}c_1 t + c_4$  and

$$\xi(x, t) = \frac{1}{2}(2c_1 t + c_2)x + c_5 t + c_6,$$

$$\tau(t) = c_1 t^2 + c_2 t + c_3,$$

$$\phi(x, t, u) = \left(-\frac{1}{4}c_1 x^2 - \frac{1}{2}c_5 x - \frac{1}{2}c_1 t + c_4\right)u + \beta(x, t).$$



Inserting in Equation (29.24) yields

$$\begin{aligned} \mathbf{v} &= \left[ \frac{1}{2}(2c_1t + c_2)x + c_5t + c_6 \right] \frac{\partial}{\partial x} + (c_1t^2 + c_2t + c_3) \frac{\partial}{\partial t} \\ &\quad + \left[ \left(-\frac{1}{4}c_1x^2 - \frac{1}{2}c_5x - \frac{1}{2}c_1t + c_4\right)u + \beta(x, t) \right] \frac{\partial}{\partial u} \\ &= c_1 \left[ xt \frac{\partial}{\partial x} + t^2 \frac{\partial}{\partial t} - \frac{1}{4}(2t + x^2)u \frac{\partial}{\partial u} \right] + c_2 \left( \frac{1}{2}x \frac{\partial}{\partial x} + t \frac{\partial}{\partial t} \right) \\ &\quad + c_3 \frac{\partial}{\partial t} + c_4u \frac{\partial}{\partial u} + c_5 \left( t \frac{\partial}{\partial x} - \frac{1}{2}xu \frac{\partial}{\partial u} \right) + c_6 \frac{\partial}{\partial x} + \beta(x, t) \frac{\partial}{\partial u}. \end{aligned}$$

Thus the Lie algebra of the infinitesimal symmetries of the heat equation is spanned by the six vector fields

$$\begin{aligned} \mathbf{v}_1 &= \partial_x, & \mathbf{v}_2 &= \partial_t, & \mathbf{v}_3 &= u\partial_u, & \mathbf{v}_4 &= x\partial_x + 2t\partial_t, \\ \mathbf{v}_5 &= 2t\partial_x - xu\partial_u, & \mathbf{v}_6 &= 4tx\partial_x + 4t^2\partial_t - (x^2 + 2t)u\partial_u \end{aligned} \quad (29.27)$$

and the infinite-dimensional subalgebra

$$\mathbf{v}_\beta = \beta(x, t)\partial_u,$$

where  $\beta$  is an arbitrary solution of the heat equation.

The one-parameter groups  $G_i$  generated by the  $\mathbf{v}_i$  can be found by solving the appropriate DEs for the integral curves. We show a sample calculation and leave the rest of the computation to the reader. Consider  $\mathbf{v}_5$ , whose integral curve is given by the set of DEs

$$\frac{dx}{ds} = 2t, \quad \frac{dt}{ds} = 0, \quad \frac{du}{ds} = -xu.$$

The second equation shows that  $t$  is not affected by the group. So,  $t = t_0$ , where  $t_0$  is the initial value of  $t$ . The first equation now gives

$$\frac{dx}{ds} = 2t_0 \Rightarrow x = 2t_0s + x_0,$$

and the last equation yields

$$\frac{du}{ds} = -(2t_0s + x_0)u \Rightarrow \frac{du}{u} = -(2t_0s + x_0)ds \Rightarrow u = u_0e^{-t_0s^2 - x_0s}.$$

Changing the transformed coordinates to  $\tilde{x}^i$  and removing the subscript from the initial coordinates, we can write

$$\exp(\mathbf{v}_5s) \cdot (x, t, u) = (\tilde{x}, \tilde{t}, \tilde{u}) = (x + 2ts, t, ue^{-sx - s^2t}).$$

Table 29.2 gives the result of the action of  $\exp(\mathbf{v}_i s)$  to  $(x, t, u)$ .

Group Element	Transformed Coordinates $(\tilde{x}, \tilde{t}, \tilde{u})$
$G_1 = \exp(v_1 s)$	$(x + s, t, u)$
$G_2 = \exp(v_2 s)$	$(x, t + s, u)$
$G_3 = \exp(v_3 s)$	$(x, t, e^s u)$
$G_4 = \exp(v_4 s)$	$(e^s x, e^{2s} t, u)$
$G_5 = \exp(v_5 s)$	$(x + 2ts, t, ue^{-sx-s^2t})$
$G_6 = \exp(v_6 s)$	$\left( \frac{x}{1-4st}, \frac{t}{1-4st}, u\sqrt{1-4st} \exp\left[ \frac{-sx^2}{1-4st} \right] \right)$
$G_\beta = \exp(v_\beta s)$	$(x, t, u + s\beta(x, t))$

**Table 29.2** The transformations caused by the symmetry group of the heat equation.

The symmetry groups  $G_1$  and  $G_2$  reflect the invariance of the heat equation under space and time translations.  $G_3$  and  $G_\beta$  demonstrate the linearity of the heat equation: We can multiply solutions by constants and add solutions to get new solutions. The scaling symmetry is contained in  $G_4$ , which shows that if you scale time by the square of the scaling of  $x$ , you obtain a new solution.  $G_5$  is a Galilean boost to a moving frame. Finally,  $G_6$  is a transformation that cannot be obtained from any physical principle. Since each group  $G_i$  is a one-parameter group of symmetries, if  $f$  is a solution of the heat equation, so are the functions  $f_i \equiv G_i \cdot f$  for all  $i$ . These functions can be obtained from Equation (29.8). As an illustration, we find  $f_6$ . First note that for  $u = f(x, t)$ , we have

$$\tilde{x} = \frac{x}{1-4st}, \quad \tilde{t} = \frac{t}{1-4st},$$

$$\tilde{u} \equiv \tilde{f}(\tilde{x}, \tilde{t}) = f(x, t)\sqrt{1-4st} \exp\left[ \frac{-sx^2}{1-4st} \right].$$

Next solve the first two equations above for  $x$  and  $t$  in terms of  $\tilde{x}$  and  $\tilde{t}$ :

$$x = \frac{\tilde{x}}{1+4s\tilde{t}}, \quad t = \frac{\tilde{t}}{1+4s\tilde{t}}.$$

Finally, substitute in the last equation to get

$$\tilde{f}(\tilde{x}, \tilde{t}) = f\left( \frac{\tilde{x}}{1+4s\tilde{t}}, \frac{\tilde{t}}{1+4s\tilde{t}} \right) \sqrt{\frac{1}{1+4s\tilde{t}}} \exp\left[ \frac{-s\tilde{x}^2}{1+4s\tilde{t}} \right],$$

or, changing  $\tilde{x}$  to  $x$  and  $\tilde{t}$  to  $t$ ,

$$f_6(x, t) = \frac{1}{\sqrt{1+4st}} \exp\left[ \frac{-sx^2}{1+4st} \right] f\left( \frac{x}{1+4st}, \frac{t}{1+4st} \right).$$

The other transformed functions can be obtained similarly. We simply list these functions:

$$\begin{aligned} f_1(x, t) &= f(x - s, t), & f_2(x, t) &= f(x, t - s), \\ f_3(x, t) &= e^s f(x, t), & f_4(x, t) &= f(e^{-s}x, e^{-2s}t), \\ f_5(x, t) &= e^{-sx+s^2t} f(x - 2st, t), & f_6(x, t) &= f(x, t) + s\beta(x, t), \\ f_6(x, t) &= \frac{1}{\sqrt{1+4st}} \exp\left[\frac{-sx^2}{1+4st}\right] f\left(\frac{x}{1+4st}, \frac{t}{1+4st}\right). \end{aligned} \quad (29.28)$$

We can find the *fundamental solution* to the heat equation very simply as follows. Let  $f(x, t)$  be the trivial constant solution  $c$ . Then

$$u = f_6(x, t) = \frac{c}{\sqrt{1+4st}} e^{-sx^2/(1+4st)}$$

is also a solution. Now choose  $c = \sqrt{s/\pi}$  and translate  $t$  to  $t - 1/4s$  (an allowed operation due to the invariance of the heat equation under time translation  $G_2$ ). The result is

$$u = \frac{1}{\sqrt{4\pi t}} e^{-x^2/4t},$$

which is the fundamental solution of the heat equation [see (22.45)].

### 29.4.2 The Wave Equation

As the next example of the application of Lie groups to differential equations, we consider the wave equation in two dimensions. This equation is written as

$$u_{tt} - u_{xx} - u_{yy} = 0, \quad \text{or} \quad \eta^{ij} u_{ij} = 0 \quad \text{and} \quad \Delta = \eta^{ij} u_{ij}, \quad (29.29)$$

where  $\eta = \text{diag}(1, -1, -1)$ , and subscripts indicate derivatives with respect to coordinate functions  $x^1 = t$ ,  $x^2 = x$ , and  $x^3 = y$ . With  $p = 3$  and  $q = 1$ , a typical generator of symmetry will be of the form

$$\mathbf{v} = \sum_{i=1}^3 X^{(i)} \frac{\partial}{\partial x^i} + U \frac{\partial}{\partial u}, \quad (29.30)$$

where  $\{X^{(i)}\}_{i=1}^3$  and  $U$  are functions of  $t, x, y$ , and  $u$  to be determined. The second prolongation of such a vector field is

$$\text{pr}^{(2)}\mathbf{v} = \mathbf{v} + \sum_{i=1}^3 U^{(i)}(x, u^{(2)}) \frac{\partial}{\partial u_i} + \sum_{i,j=1}^3 U^{(ij)}(x, u^{(2)}) \frac{\partial}{\partial u_{ij}},$$

where by Theorem 29.3.5,

$$\begin{aligned}
 U^{(i)} &= D_i \left( U - \sum_{k=1}^3 X^{(k)} u_k \right) + \sum_{k=1}^3 X^{(k)} u_{ik} = D_i U - \sum_{k=1}^3 (D_i X^{(k)}) u_k, \\
 U^{(ij)} &= D_i D_j \left( U - \sum_{k=1}^3 X^{(k)} u_k \right) + \sum_{k=1}^3 X^{(k)} u_{ijk} \\
 &= D_i D_j U - \sum_{k=1}^3 [(D_i D_j X^{(k)}) u_k + u_{ik} (D_j X^{(k)}) + u_{jk} (D_i X^{(k)})],
 \end{aligned}$$

and we have used Equation (29.21). Using (29.21) further, the reader may show that

$$\begin{aligned}
 U^{(ij)} &= U_{ij} + u_k (\delta_{jk} U_{iu} + \delta_{ik} U_{ju} - X_{ij}^{(k)}) \\
 &\quad + u_l u_k (\delta_{il} \delta_{jk} U_{uu} - X_{iu}^{(k)} \delta_{jl} - X_{ju}^{(k)} \delta_{il}) \\
 &\quad + u_{kl} (\delta_{il} \delta_{jk} U_u - X_i^{(k)} \delta_{jl} - X_j^{(k)} \delta_{il}) \\
 &\quad - u_k u_{lm} (X_u^{(k)} \delta_{il} \delta_{jm} + X_u^{(m)} \delta_{il} \delta_{jk} + X_u^{(m)} \delta_{ik} \delta_{jl}) - u_i u_j u_k X_{uu}^{(k)},
 \end{aligned} \tag{29.31}$$

where a sum over repeated indices is understood.

Applying  $\text{pr}^{(2)} \nu$  to  $\Delta$ , we obtain the infinitesimal criterion

$$U^{(tt)} = U^{(xx)} + U^{(yy)} \quad \text{or} \quad \eta^{ij} U^{(ij)} = 0.$$

Multiplying Equation (29.31) by  $\eta^{ij}$  and setting the result equal to zero yields

$$\begin{aligned}
 0 &= \eta^{ij} U^{(ij)} = \eta^{ij} U_{ij} + u_k (2\eta^{ik} U_{iu} - \eta^{ij} X_{ij}^{(k)}) + u_l u_k (\eta^{kl} U_{uu} - 2X_{iu}^{(k)} \eta^{il}) \\
 &\quad - 2u_{kl} X_i^{(k)} \eta^{il} - 2u_k u_{lm} X_u^{(m)} \eta^{kl} - u_i u_j u_k X_{uu}^{(k)} \eta^{ij},
 \end{aligned} \tag{29.32}$$

where we have used the wave equation,  $\eta^{kl} u_{kl} = 0$ . Equation (29.32) must hold for all derivatives of  $u$  and powers thereof (treated as independent) modulo the wave equation. Therefore, the coefficients of such “monomials” must vanish. For example, since all the terms involving  $u_k u_{lm}$  are independent (even after substituting  $u_{xx} + u_{yy}$  for  $u_{tt}$ ), we have to conclude that  $X_u^{(m)} = 0$  for all  $m$ , i.e., that  $X^{(i)}$  are independent of  $u$ . Setting the coefficient of  $u_k u_l$  equal to zero and noting that  $X_{iu}^{(k)} = \partial X_u^{(k)} / \partial x^i = 0$  yields

$$U_{uu} = 0 \Rightarrow U(x, y, t, u) = \alpha(x, y, t)u + \beta(x, y, t). \tag{29.33}$$

Let us concentrate on the functions  $X^{(i)}$ . These are related via the term linear in  $u_{kl}$ . After inserting the wave equation in this term, we get

$$\begin{aligned}
 u_{kl} X_i^{(k)} \eta^{il} &= u_{12} (X_1^{(2)} - X_2^{(1)}) + u_{13} (X_1^{(3)} - X_3^{(1)}) - u_{23} (X_2^{(3)} + X_3^{(2)}) \\
 &\quad + u_{22} (X_1^{(1)} - X_2^{(2)}) + u_{33} (X_1^{(1)} - X_3^{(3)}).
 \end{aligned}$$

The  $u_{ij}$  in this equation are all independent; so, we can set their coefficients equal to zero:

$$X_1^{(2)} = X_2^{(1)}, \quad X_1^{(3)} = X_3^{(1)}, \quad X_2^{(3)} + X_3^{(2)} = 0, \quad X_1^{(1)} = X_2^{(2)} = X_3^{(3)}. \quad (29.34)$$

The reader may verify that these relations imply that  $X_{jkl}^{(i)} = 0$  for any  $i, j, k$ , and  $l$ . For example,

$$\begin{aligned} X_{222}^{(2)} &= X_{122}^{(1)} = \underbrace{X_{322}^{(3)}}_{\equiv X_{223}^{(3)}} = -\underbrace{X_{323}^{(2)}}_{\equiv X_{233}^{(2)}} = -\underbrace{X_{133}^{(1)}}_{\equiv X_{313}^{(1)}} = -\underbrace{X_{113}^{(3)}}_{\equiv X_{311}^{(3)}} \\ &= -\underbrace{X_{211}^{(2)}}_{\equiv X_{121}^{(2)}} = -\underbrace{X_{221}^{(1)}}_{\equiv X_{122}^{(1)}} = -X_{222}^{(2)}. \end{aligned} \quad (29.35)$$

So, the first link of this chain is equal to its negative. Therefore, all the third derivatives in the chain of Equation (29.35) vanish. It follows that all  $X^{(i)}$ 's are mixed polynomials of at most degree two. Writing the most general such polynomials for the three functions  $X^{(1)}$ ,  $X^{(2)}$ , and  $X^{(3)}$  and having them satisfy Equation (29.34) yields

$$\begin{aligned} X^{(1)} &= a_1 + a_4x + a_7y + a_5t + a_8(x^2 + y^2 + t^2) + 2a_9xt + 2a_{10}yt, \\ X^{(2)} &= a_2 + a_5x - a_6y + a_4t + a_9(x^2 - y^2 + t^2) + 2a_{10}xy + 2a_8xt, \\ X^{(3)} &= a_3 + a_6x + a_5y + a_7t + a_{10}(-x^2 + y^2 + t^2) + 2a_9xy + 2a_8yt. \end{aligned} \quad (29.36)$$

Setting the coefficient of  $u_k$  and  $\eta^{ij}U_{ij}$  equal to zero and using Equation (29.33) gives

$$\begin{aligned} 2\alpha_x &= X_{xx}^{(2)} + X_{yy}^{(2)} - X_{tt}^{(2)}, & 2\alpha_y &= X_{xx}^{(3)} + X_{yy}^{(3)} - X_{tt}^{(3)}, \\ 2\alpha_t &= X_{tt}^{(1)} - X_{xx}^{(1)} - X_{yy}^{(1)}, & \beta_{tt} - \beta_{xx} - \beta_{yy} &= 0. \end{aligned}$$

It follows that  $\beta$  is any solution of the wave equation, and

$$\alpha(x, y, t) = a_{11} - a_8t - a_9x - a_{10}y.$$

By inserting the expressions found for  $X^{(i)}$  and  $U$  in (29.30) and writing the result in the form  $\sum_i a_i \mathbf{v}_i$ , we discover that the generators of the symmetry group consist of the ten vector fields given in Table 29.3 as well as the vector fields

$$u\partial_u, \quad \mathbf{v}_\beta = \beta(x, y, t)\partial_u$$

for  $\beta$  an arbitrary solution of the wave equation. The ten vector fields of Table 29.3 comprise the generators of the conformal group in three dimensions whose generalization to  $m$  dimensions was studied in Section 28.4.

Infinitesimal generator	Transformation
$v_1 = \partial_t, \quad v_2 = \partial_x, \quad v_3 = \partial_y$	Translation
$v_4 = x\partial_t + t\partial_x$ $v_6 = -y\partial_x + x\partial_y$ $v_7 = y\partial_t + t\partial_y$	Rotation/Boost
$v_5 = t\partial_t + x\partial_x + y\partial_y$	Dilatation
$v_8 = (t^2 + x^2 + y^2)\partial_t + 2xt\partial_x + 2yt\partial_y - tu\partial_u$ $v_9 = 2xt\partial_t + (t^2 + x^2 - y^2)\partial_x + 2xy\partial_y - xu\partial_u$ $v_{10} = 2yt\partial_t + 2xy\partial_x + (t^2 - x^2 + y^2)\partial_y - yu\partial_u$	Inversions

**Table 29.3** The generators of the conformal group for  $\mathbb{R}^3$ , part of the symmetry group of the wave equation in two dimensions.

## 29.5 Application to ODEs

The theory of Lie groups finds one of its most rewarding applications in the integration of ODEs. Lie’s fundamental observation was that if one could come up with a sufficiently large group of symmetries of a system of ODEs, then one could integrate the system. In this section we outline the general technique of solving ODEs once we know their symmetries. The following proposition will be useful (see [Warn 83, p. 40]):

**29.5.1. Proposition.** *Let  $M$  be an  $n$ -dimensional manifold and  $v \in \mathfrak{X}(M)$ . Assume that  $v|_P \neq 0$  for some  $P \in M$ . Then there exists a local chart, i.e., local set of coordinate functions,  $(w^1, \dots, w^n)$  at  $P$  such that  $v = \partial/\partial w^1$ .*

### 29.5.1 First-Order ODEs

The most general first-order ODE can be written as

$$\frac{du}{dx} = F(x, u) \Rightarrow \Delta(x, u, u_x) \equiv u_x - F(x, u) = 0. \tag{29.37}$$

A typical infinitesimal generator of the symmetry group of this equation is<sup>4</sup>  $\mathbf{v} = X\partial_x + U\partial_u$ , whose prolongation is

$$\text{pr}^{(1)} = \mathbf{v} + U^{(x)} \frac{\partial}{\partial u_x}, \quad \text{where} \quad U^{(x)} \equiv U_x + (U_u - X_x)u_x - X_u u_x^2, \tag{29.38}$$

as the reader may verify. The infinitesimal criterion for the one-parameter group of transformations  $G$  to be a symmetry group of Equation (29.37) is  $\text{pr}^{(1)}\mathbf{v}(\Delta) = 0$ , or

$$\frac{\partial U}{\partial x} + \left( \frac{\partial U}{\partial u} - \frac{\partial X}{\partial x} \right) F - \frac{\partial X}{\partial u} F^2 = X \frac{\partial F}{\partial x} + U \frac{\partial F}{\partial u}. \tag{29.39}$$

Any solution  $(X, U)$  of this equation generates a 1-parameter group of transformations. The problem is that a *systematic* procedure for solving (29.39) is *more difficult* than solving the original equation. However, in most cases, one can *guess* a symmetry transformation (based on physical, or other, grounds), and that makes Lie’s method worthwhile.

Suppose we have found a symmetry group  $G$  with infinitesimal generator  $\mathbf{v}$  that does not vanish at  $P \in M \subset X \times U$ . Based on Proposition 29.5.1, we can introduce new coordinates

$$w = \xi(x, u), \quad y = \eta(x, u) \tag{29.40}$$

in a neighborhood of  $P$  such that  $\mathbf{v} = \partial/\partial w$ , whose prolongation is also  $\partial/\partial w$  [see (29.38)]. This transforms the DE of (29.37) into<sup>5</sup>  $\tilde{\Delta}(y, w, w_y) = 0$ , and the infinitesimal criterion into

$$\text{pr}^{(1)}\mathbf{v}(\tilde{\Delta}) = \frac{\partial \tilde{\Delta}}{\partial w} = 0.$$

It follows that  $\tilde{\Delta}$  is independent of  $w$ . The transformed DE is therefore  $\tilde{\Delta}(y, w_y) = 0$ , whose normal form, obtained by implicitly solving for  $dw/dy$ , is

$$\frac{dw}{dy} = H(y) \Rightarrow w = \int_a^y H(t) dt - w(a)$$

for some function  $H$  of  $y$  alone and some convenient point  $y = a$ . Substituting this expression of  $w$  as a function of  $y$  in Equation (29.40) and eliminating  $y$  between the two equations yields  $u$  as a function of  $x$ .

Thus our task is to find the change of variables (29.40). For this, we use  $\mathbf{v}(w) = 1$  and  $\mathbf{v}(y) = 0$ , and express them in terms of  $x$  and  $u$ :

$$\begin{aligned} \mathbf{v}(w) = \mathbf{v}(\xi) &= X \frac{\partial \xi}{\partial x} + U \frac{\partial \xi}{\partial u} = 1, \\ \mathbf{v}(y) = \mathbf{v}(\eta) &= X \frac{\partial \eta}{\partial x} + U \frac{\partial \eta}{\partial u} = 0. \end{aligned} \tag{29.41}$$

<sup>4</sup>The reader is warned against the unfortunate coincidence of notation:  $X$  and  $U$  represent *both* the components of the infinitesimal generator and the spaces of independent and dependent variables!

<sup>5</sup>Here we are choosing  $w$  to be the dependent variable. This choice is a freedom that is always available to us.

The second equation says that  $\eta$  is an invariant of the group generated by  $\mathbf{v}$ . We therefore use the associated characteristic ODE [see (29.3) and (29.5)] to find  $y$  (or  $\eta$ ):

$$\frac{dx}{X(x, u)} = \frac{du}{U(x, u)}. \quad (29.42)$$

To find  $w$  (or  $\xi$ ), we introduce  $\chi(x, u, v) = v - \xi(x, u)$  and note that an equivalent relation containing the same information as the first equation in (29.41) is

$$X \frac{\partial \chi}{\partial x} + U \frac{\partial \chi}{\partial u} + \frac{\partial \chi}{\partial v} = 0,$$

which has the characteristic ODE

$$\frac{dx}{X(x, u)} = \frac{du}{U(x, u)} = \frac{dv}{1}, \quad (29.43)$$

for which we seek a solution of the form  $v - \xi(x, u) = c$  to read off  $\xi(x, u)$ .

The reader may wonder whether it is sane to go through so much trouble only to replace the original single ODE with *two* ODEs such as (29.42) and (29.43)! The answer is that *in practice*, the latter two DEs are much easier to solve than the original ODE.

**29.5.2. Example.** The homogeneous FODE  $du/dx = F(u/x)$  is invariant under the scaling transformation  $(x, u) \mapsto (sx, su)$  whose infinitesimal generator is  $\mathbf{v} = x\partial_x + u\partial_u$ . The first prolongation of this vector is the same as the vector itself (reader, verify!).

To find the new coordinates  $w$  and  $y$ , first use Equation (29.42) with  $X(x, u) = x$  and  $U(x, u) = u$ :

$$\frac{dx}{x} = \frac{du}{u} \Rightarrow \frac{u}{x} = c_1 \Rightarrow y = \frac{u}{x} \quad (\text{see Box 29.1.8}).$$

Next, we note that (29.43) yields

$$\frac{dx}{x} = \frac{du}{u} = dv \Rightarrow \ln u = v + \ln c_2 \Rightarrow v = \ln(u/c_2).$$

Substituting from the previous equation, we obtain

$$v = \ln(c_1 x / c_2) = \ln x + \underbrace{\ln(c_1 / c_2)}_{=c} \Rightarrow v - \ln x = c \Rightarrow w = \ln x.$$

The chain rule gives  $du/dx = (1 + yw_y)/w_y$ , so that the DE becomes

$$\frac{1 + yw_y}{w_y} = F(y) \Rightarrow \frac{dw}{dy} = \frac{1}{F(y) - y},$$

which can be integrated to give  $w = H(y)$  or  $\ln x = H(y) = H(u/x)$ , which defines  $u$  as an implicit function of  $x$ . ■



## 29.5.2 Higher-Order ODEs

The same argument used in the first order ODEs can be used for higher-order ODEs to reduce their orders.

### 29.5.3 Proposition. *Let*

$$\Delta(x, u^{(n)}) = \Delta(x, u, u_1, \dots, u_n) = 0, \quad u_k \equiv \frac{d^k u}{dx^k},$$

*be an  $n$ th order ODE. If this ODE has a one-parameter symmetry group, then there exist variables  $w = \xi(x, u)$  and  $y = \eta(x, u)$  such that*

$$\Delta(x, u^{(n)}) = \bar{\Delta}\left(y, \frac{dw}{dy}, \dots, \frac{d^n w}{dy^n}\right) = 0,$$

*i.e., in terms of  $w$  and  $y$ , the ODE becomes of  $(n - 1)$ st order in  $w_y$ .*

*Proof.* The proof is exactly the same as in the first-order case. The only difference is that one has to consider  $\text{pr}^{(n)}\mathbf{v}$ , where  $\mathbf{v} = \partial/\partial w$ . But Problem 29.7 shows that  $\text{pr}^{(n)}\mathbf{v} = \mathbf{v}$ , as in the first-order case.  $\square$

**29.5.4. Example.** Consider a second-order DE  $\Delta(u, u_x, u_{xx}) = 0$ , which does not depend on  $x$  explicitly. The fact that  $\partial\Delta/\partial x = 0$  suggests  $w = x$ . So, we switch the dependent and independent variables and write  $w = x$ , and  $y = u$ . Then, using the chain rule, we get

$$\frac{du}{dx} = \frac{1}{w_y}, \quad \frac{d^2u}{dx^2} = -\frac{w_{yy}}{w_y^3}.$$

Substituting in the original DE, we obtain

$$\bar{\Delta}(y, w_y, w_{yy}) \equiv \Delta\left(y, \frac{1}{w_y}, -\frac{w_{yy}}{w_y^3}\right) = 0,$$

which is of first order in  $w_y$ .  $\blacksquare$

**29.5.5. Example.** The order of the SOLDE  $u_{xx} + p(x)u_x + q(x)u = 0$  can be reduced by noting that the DE is invariant under the scaling transformation  $(x, u) \mapsto (x, su)$ , whose infinitesimal generator is  $\mathbf{v} = u\partial_u$ . With this vector field, Equations (29.42) and (29.43) give

$$\frac{dx}{0} = \frac{du}{u}, \quad \frac{dx}{0} = \frac{du}{u} = dv.$$

For the first equation to make sense, we have to have

$$dx = 0 \Rightarrow x = c_1 \Rightarrow y = x \quad (\text{by Box 29.1.8}).$$

The second equation in  $u$  gives

$$v = \ln u + c \Rightarrow v - \ln u = c \Rightarrow w = \ln u \Rightarrow u = e^w.$$

Using the chain rule, we obtain

$$u_x = \frac{dw}{dx}e^w = \frac{dw}{dy}e^w \quad \text{and} \quad u_{xx} = (w_{yy} + w_y^2)e^w.$$

Riccati equation By inserting this in the original DE and writing  $z = w_y$ , we obtain

$$\frac{dz}{dy} = -z^2 - p(y)z - q(y),$$

which is the well-known first-order Riccati equation. ■

### 29.5.3 DEs with Multiparameter Symmetries

We have seen that 1-parameter symmetries reduce the order of an ODE by 1. It is natural to suspect that an  $r$ -parameter symmetry will reduce the order by  $r$ . Although this suspicion is correct, it turns out that in general, one cannot reconstruct the solution of the original equation from those of the reduced  $(n - r)$ th-order equation. (See [Olve 86, pp. 148–158] for a thorough discussion of this problem.) However, the special, but important, case of second-order DEs is an exception. The deep reason behind this is the exceptional structure of 2-dimensional Lie algebras given in Box 27.2.5. We cannot afford to go into details of the reasoning, but simply quote the following important theorem.

**29.5.6. Theorem.** *Let  $\Delta(x, u^{(n)}) = 0$  be an  $n$ th-order ODE invariant under a 2-parameter group. Then there is an  $(n - 2)$ nd-order ODE  $\tilde{\Delta}(y, w^{(n-2)}) = 0$  with the property that the general solution to  $\Delta$  can be found by integrating the general solution to  $\tilde{\Delta}$ . In particular, a second-order ODE having a 2-parameter group of symmetries can be solved by integration.*

Let us analyze the case of a second-order ODE in some detail. By Box 27.2.5, the infinitesimal generators  $\mathbf{v}_1$  and  $\mathbf{v}_2$  satisfy the Lie bracket relation

$$[\mathbf{v}_1, \mathbf{v}_2] = c\mathbf{v}_1, \quad c = 0 \text{ or } 1.$$

We shall treat the abelian case ( $c = 0$ ) and leave the nonabelian case for the reader. To begin with, we use  $s$  and  $t$  for the transformed variables, and at the end replace them with  $y$  and  $w$ .

By Proposition 29.5.1, we can let  $\mathbf{v}_1 = \partial/\partial s$ . Then  $\mathbf{v}_2$  can be expressed as the linear combination

$$\mathbf{v}_2 = \alpha(s, t) \frac{\partial}{\partial s} + \beta(s, t) \frac{\partial}{\partial t}.$$

The commutation relation  $[\mathbf{v}_1, \mathbf{v}_2] = 0$  gives

$$0 = [\partial_s, \alpha \partial_s + \beta \partial_t] = \frac{\partial \alpha}{\partial s} \partial_t + \frac{\partial \beta}{\partial s} \partial_s,$$

showing that  $\alpha$  and  $\beta$  are independent of  $s$ . We want to simplify  $\mathbf{v}_2$  as much as possible without changing  $\mathbf{v}_1$ . A transformation that accomplishes this is  $S = s + h(t)$  and  $T = T(t)$ . Then, by Equation (26.8) we obtain

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{v}_1(S) \frac{\partial}{\partial S} + \mathbf{v}_1(T) \frac{\partial}{\partial T} = \frac{\partial S}{\partial s} \frac{\partial}{\partial S} + \frac{\partial T}{\partial s} \frac{\partial}{\partial T} = \frac{\partial}{\partial s}, \\ \mathbf{v}_2 &= \mathbf{v}_2(S) \frac{\partial}{\partial S} + \mathbf{v}_2(T) \frac{\partial}{\partial T} = \left( \alpha \frac{\partial S}{\partial s} + \beta \frac{\partial S}{\partial t} \right) \frac{\partial}{\partial S} + \left( \alpha \frac{\partial T}{\partial s} + \beta \frac{\partial T}{\partial t} \right) \frac{\partial}{\partial T} \\ &= (\alpha + \beta h') \frac{\partial}{\partial S} + \beta T' \frac{\partial}{\partial T}. \end{aligned}$$

If  $\beta \neq 0$ , we choose  $T' = 1/\beta$  and  $h' = -\alpha/\beta$  to obtain

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \quad \mathbf{v}_2 = \frac{\partial}{\partial t}, \tag{29.44}$$

where we have substituted  $s$  for  $S$  and  $t$  for  $T$ . If  $\beta = 0$ , we choose  $\alpha = T$ , and change the notation from  $S$  to  $s$  and  $T$  to  $t$  to obtain

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \quad \mathbf{v}_2 = t \frac{\partial}{\partial s}. \tag{29.45}$$

The next step is to decide which coordinate is the independent variable, prolong the vector fields, and apply it to the DE to find the infinitesimal criterion. For  $\beta \neq 0$ , the choice is immaterial. So, let  $w = s$  and  $y = t$ . Then the prolongation of  $\mathbf{v}_1$  and  $\mathbf{v}_2$  will be the same as the vectors themselves, and with  $\Delta(y, w, w_y, w_{yy}) \equiv w_{yy} - F(y, w, w_y)$ , the infinitesimal criteria for invariance will be

$$\begin{aligned} 0 &= \text{pr}^{(2)}\mathbf{v}_1(\Delta) = \mathbf{v}_1(\Delta) = \frac{\partial \Delta}{\partial w} = -\frac{\partial F}{\partial w}, \\ 0 &= \text{pr}^{(2)}\mathbf{v}_2(\Delta) = \mathbf{v}_2(\Delta) = \frac{\partial \Delta}{\partial y} = -\frac{\partial F}{\partial y}. \end{aligned}$$

It follows that in the  $(y, w)$  system,  $F$  will be a function of  $w_y$  alone and the DE will be of the form

$$w_{yy} = F(w_y) \Rightarrow \frac{dw_y}{dy} = F(w_y) \Rightarrow \underbrace{\int^{w_y} \frac{dz}{F(z)}}_{\equiv H(w_y)}.$$

The last equation can be solved for  $w_y$  in terms of  $y$  and the result integrated.

For  $\beta = 0$ , choose  $w = t$  and  $y = s$ . Then  $\mathbf{v}_1$  will not prolongate, and as the reader may verify,

$$\text{pr}^{(2)}\mathbf{v}_2 = \mathbf{v}_2 - w_y^2 \frac{\partial}{\partial w_y} - 3w_y w_{yy} \frac{\partial}{\partial w_{yy}} = w \frac{\partial}{\partial y} - w_y^2 \frac{\partial}{\partial w_y} - 3w_y w_{yy} \frac{\partial}{\partial w_{yy}},$$

and the infinitesimal criteria for invariance will be

$$0 = \text{pr}^{(2)}\mathbf{v}_1(\Delta) = \mathbf{v}_1(\Delta) = \frac{\partial \Delta}{\partial y} = -\frac{\partial F}{\partial y},$$

$$0 = \text{pr}^{(2)}\mathbf{v}_2(\Delta) = -w \underbrace{\frac{\partial F}{\partial y}}_{=0} + w_y^2 \frac{\partial F}{\partial w_y} + 3w_y \underbrace{w_{yy}}_{=F}.$$

It follows that in the  $(y, w)$  system,  $F$  will be a function of  $w$  and  $w_y$  and satisfy the DE

$$w_y \frac{\partial F}{\partial w_y} = 3F,$$

whose solution is of the form  $F(w, w_y) = w_y^3 \tilde{F}(w)$ . The original DE now becomes

$$w_{yy} = w_y^3 \tilde{F}(w),$$

for which we use the chain rule  $w_{yy} = w_y \partial w_y / \partial w$  to obtain

$$\frac{dw_y}{dw} = w_y^2 \tilde{F}(w) \Rightarrow -\frac{1}{w_y} = \underbrace{\int^w \tilde{F}(z) dz}_{=H(w)} \Rightarrow \frac{dw}{dy} = -\frac{1}{H(w)},$$

which can be integrated. Had we chosen  $w = s$  and  $y = t$ ,  $F$  would have been a function of  $y$  and the DE would have reduced to  $w_{yy} = F(y)$ , which could be solved by two consecutive integrations. The nonabelian 2-dimensional Lie algebra can be analyzed similarly. The reader may verify that if  $\beta = 0$ , the vector fields can be chosen to be

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \quad \mathbf{v}_2 = s \frac{\partial}{\partial s}, \tag{29.46}$$

leading to the ODE  $w_{yy} = w_y \tilde{F}(y)$ , and if  $\beta \neq 0$ , the vector fields can be chosen to be

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \quad \mathbf{v}_2 = s \frac{\partial}{\partial s} + t \frac{\partial}{\partial t}, \tag{29.47}$$

leading to the ODE  $w_{yy} = \tilde{F}(w_y)/y$ . Both of these ODEs are integrable as in the abelian case.

## 29.6 Problems

**29.1.** Suppose that  $\{F_i\}_{i=1}^n$  are invariants of the PDE (29.3). Show that any function  $f(F_1, F_2, \dots, F_n)$  is also an invariant of the PDE.

29.2. Find the function  $\tilde{f} = \theta \cdot f$  when  $f(x) = ax + b$  and  $\theta$  is the angle of rotation of  $SO(2)$ .

29.3. Use the result of Problem 29.2 to find  $\tilde{u}_1$ . Hint: Note that  $a = u_1$ .

29.4. Transform the DE of Example 29.3.2 from Cartesian to polar coordinates to obtain  $dr/d\theta = r^3$ .

29.5. Using the definition of total derivative, verify Equation (29.21).

29.6. Show that  $SO(2)$  is a symmetry group of the first-order DE

$$\Delta(x, u, u_1) = (u - x)u_1 + x + u = 0$$

and write the same DE in polar coordinates.

29.7. Show that the  $n$ th prolongation of the generator of the  $i$ th translation,  $\partial_i$ , is the same as the original vector.

29.8. Find the first prolongation of the generator of scaling:  $x\partial_x + u\partial_u$ .

29.9. Show that when the group acts only on the single dependent variable  $u$ , the prolongation of  $\mathbf{v} = U\partial_u$  is given by

$$\text{pr}^{(1)}\mathbf{v} = \mathbf{v} + \sum_{j=1}^p U_j \frac{\partial}{\partial u_j}, \quad \text{where } U_j = \frac{\partial U}{\partial x^j} + u_j \frac{\partial U}{\partial u}.$$

29.10. Show that the  $n$ th prolongation of  $\mathbf{v} = X(x, u)\partial_x + U(x, u)\partial_u$  for an ordinary DE of  $n$ th order is

$$\text{pr}^{(n)}\mathbf{v} = \mathbf{v} + \sum_{k=1}^n U^{[k]} \frac{\partial}{\partial u^{(k)}},$$

where

$$u^{(k)} \equiv \frac{\partial^k u}{\partial x^k} \quad \text{and} \quad U^{[k]} = D_x^k(U - Xu_x) + Xu^{(k+1)}.$$

29.11. Compute the second prolongation of the infinitesimal generators of the symmetry group of the heat equation.

29.12. Derive Equations (29.31) and (29.32).

29.13. Using Equation (29.34) show that  $X_{jkl}^{(l)} = 0$  for any  $i, j, k$ , and  $l$ .

Korteweg–de Vries equation

29.14. The Korteweg–de Vries equation is  $u_t + u_{xxx} + uu_x = 0$ . Using the technique employed in computing the symmetries of the heat and wave equations, show that the infinitesimal generators of symmetries of the Korteweg–de Vries equation are

$\mathbf{v}_1 = \partial_x,$	$\mathbf{v}_2 = \partial_t,$	translation
$\mathbf{v}_3 = t\partial_x + \partial_u,$		Galilean boost
$\mathbf{v}_4 = x\partial_x + 3t\partial_t - 2u\partial_u.$		scaling

**29.15.** Suppose  $M(x, u) dx + N(x, u) du = 0$  has a 1-parameter symmetry group with generator  $\mathbf{v} = X\partial_x + U\partial_u$ . Show that the function  $q(x, u) = 1/(XM + UN)$  is an integrating factor.

**29.16.** Show that the second prolongation of  $\mathbf{v} = w\partial_y$  (with  $y$  treated as independent variable) is

$$\text{pr}^{(2)}\mathbf{v} = \mathbf{v} - w_y^2 \frac{\partial}{\partial w_y} - 3w_y w_{yy} \frac{\partial}{\partial w_{yy}}.$$

**29.17.** Go through the case of  $\beta = 0$  in the solution of the second order ODE and, choosing  $w = s$  and  $y = t$ , show that  $F$  will be a function of  $y$  alone and the original DE will reduce to  $w_{yy} = F(y)$ .

**29.18.** Show that in the case of the nonabelian 2-dimensional Lie algebra, (a) the vector fields can be chosen to be

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \quad \mathbf{v}_2 = s \frac{\partial}{\partial s}$$

if  $\beta = 0$ .

(b) Show that these vector fields lead to the ODE  $w_{yy} = w_y \tilde{F}(y)$ .

(c) If  $\beta \neq 0$ , show that the vector fields can be chosen to be

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \quad \mathbf{v}_2 = s \frac{\partial}{\partial s} + t \frac{\partial}{\partial t}.$$

(d) Finally, show that the latter vector fields lead to the ODE  $w_{yy} = \tilde{F}(w_y)/y$ .

## Additional Reading

1. Bluman, G. and Kumei, S. *Symmetries and Differential Equations*, Springer-Verlag, 1989. A readable book on the subject of this chapter with many worked-out examples.
2. Olver, P. *Application of Lie Groups to Differential Equations*, Springer-Verlag, 1986. Our treatment of the subject follows closely that of Olver. This self-contained book, although formal, is very lucid in style with many historical notes. All the concepts introduced are clearly stated and many examples introduced to clarify them.
3. Stephani, H. *Differential Equations: Their solutions using symmetries*, Cambridge University Press, 1989. Written by a physicist, this little book treats solutions of ordinary differential equations in great detail with very little formalism.

# Calculus of Variations, Symmetries, and Conservation Laws

In this chapter we shall start with one of the oldest and most useful branches of mathematical physics, the calculus of variations. After giving the fundamentals and some examples, we shall investigate the consequences of symmetries associated with variational problems. The chapter then ends with Noether's theorem, which connects such symmetries with their associated conservation laws. All vector spaces of relevance in this chapter will be assumed to be real.

## 30.1 The Calculus of Variations

One of the main themes of calculus is the extremal problem: Given a function  $f : \mathbb{R} \supset D \rightarrow \mathbb{R}$ , find the points in the domain  $D$  of  $f$  at which  $f$  attains a maximum or minimum. To locate such points, we find the zeros of the derivative of  $f$ . For multivariable functions,  $f : \mathbb{R}^P \supset \Omega \rightarrow \mathbb{R}$ , the notion of *gradient* generalizes that of the derivative. To find the  $j$ th component of the gradient  $\nabla f$ , we calculate the difference  $\Delta f$  between the value of  $f$  at  $(x^1, \dots, x^j + \varepsilon, \dots, x^P)$  and its value at  $(x^1, \dots, x^j, \dots, x^P)$ , divide this difference by  $\varepsilon$ , and take the limit  $\varepsilon \rightarrow 0$ . This is simply partial differentiation, and the  $j$ th component of the gradient is just the  $j$ th partial derivative of  $f$ .

### 30.1.1 Derivative for Hilbert Spaces

To make contact with the subject of this chapter, let us reinterpret the notion of differentiation. The most useful interpretation is geometric. In fact, our first encounter with the derivative is geometrical: We are introduced to the concept through lines tangent to curves. In this language, the derivative of a function  $f : \mathbb{R} \supset \Omega \rightarrow \mathbb{R}$  at  $x_0$  is a line (or function)  $\psi : \Omega \supset \Omega_0 \rightarrow \mathbb{R}$  passing through

$(x_0, f(x_0))$  whose slope is defined to be the derivative of  $f$  at  $x_0$  (see Figure 30.1):

$$\psi(x) = f(x_0) + f'(x_0)(x - x_0).$$

The function  $\psi(x)$  describes a line, but it is not a *linear* function (in the vector-space sense of the word). The requirement of linearity is due to our desire for generalization of differentiation to Hilbert spaces, on which linear maps are the most natural objects. Therefore, we consider the line parallel to  $\psi(x)$  that passes through the origin. Call this  $\phi(x)$ . Then

$$\phi(x) = f'(x_0)x, \tag{30.1}$$

which is indeed a linear function. We identify  $\phi(x)$  as the derivative of  $f$  at  $x_0$ . This identification may appear strange at first but, as we shall see shortly, is the most convenient and useful. Of course, any identification requires a one-to-one correspondence between objects identified. It is clear that indeed there is a one-to-one correspondence between derivatives at points and linear functions with appropriate slopes. Equation (30.1) can be used to geometrize the definition of derivative. First consider

$$f'(x_0) = \frac{\phi(x) - \phi(x_0)}{x - x_0}, \quad \text{and} \quad f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Next note that, contrary to  $f$  which is usually defined only for a subset of the real line,  $\phi$  is defined for all real numbers  $\mathbb{R}$ , and that  $\phi(x - x_0) = \phi(x) - \phi(x_0)$  due to the linearity of  $\phi$ . Thus, we have

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \frac{\phi(x) - \phi(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{\phi(x - x_0)}{x - x_0},$$

or

$$\lim_{x \rightarrow x_0} \frac{|f(x) - f(x_0) - \phi(x - x_0)|}{|x - x_0|} = 0 \tag{30.2}$$

where we have introduced absolute values in anticipation of its analogue—norm. Equation (30.2) is readily generalized to any complete normed vector space (Banach space), and in particular to any Hilbert space:

**30.1.1. Definition.** Let  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be Hilbert spaces with norms  $\|\cdot\|_1$  and  $\|\cdot\|_2$ , respectively. Let  $f : \mathcal{H}_1 \supset \Omega \rightarrow \mathcal{H}_2$  be any map and  $|x_0\rangle \in \Omega$ . Suppose there is a linear map  $\mathbf{T} \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$  with the property that

$$\lim_{\|x - x_0\|_1 \rightarrow 0} \frac{\|f(|x\rangle) - f(|x_0\rangle) - \mathbf{T}(|x\rangle - |x_0\rangle)\|_2}{\|x - x_0\|_1} = 0 \quad \text{for } |x\rangle \in \Omega.$$

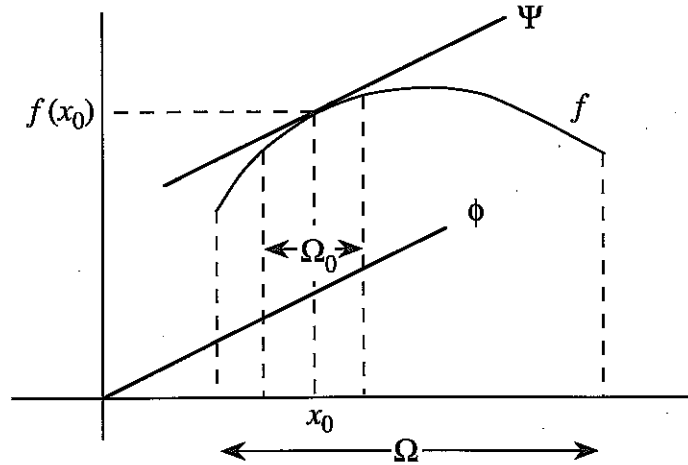
differentiability of a function on a Hilbert space at a point

Then, we say that  $f$  is **differentiable at  $|x_0\rangle$** , and we define the **derivative of  $f$  at  $|x_0\rangle$**  to be  $\mathbf{D}f(x_0) \equiv \mathbf{T}$ . If  $f$  is differentiable at each  $|x\rangle \in \Omega$ , the map

$$\mathbf{D}f : \Omega \rightarrow \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2) \quad \text{given by} \quad \mathbf{D}f(|x\rangle) = \mathbf{D}f(x)$$

is called the **derivative of  $f$** .





**Figure 30.1** The derivative at  $(x_0, f(x_0))$  as a linear function passing through the origin with a slope  $f'(x_0)$ . The function  $f$  is assumed to be defined on a subset  $\Omega$  of the real line.  $\Omega_0$  restricts the  $x$ 's to be close to  $x_0$  to prevent the function from misbehaving (blowing up), and to make sure that the limit in the definition of derivative makes sense.

The reader may verify that if the derivative exists, it is unique.

**30.1.2. Example.** Let  $\mathcal{H}_1 = \mathbb{R}^n$  and  $\mathcal{H}_2 = \mathbb{R}^m$  and  $f : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^m$ . Then for  $|x\rangle \in \Omega$ ,  $\mathbf{D}f(x)$  is a linear map, which can be represented by a matrix in the standard bases of  $\mathbb{R}^n$  and  $\mathbb{R}^m$ . To find this matrix, we need to let  $\mathbf{D}f(x)$  act on the  $j$ th standard basis of  $\mathbb{R}^n$ , i.e., we need to evaluate  $\mathbf{D}f(x)|e_j\rangle$ . This suggests taking  $|y\rangle = |x\rangle + h|e_j\rangle$  (with  $h \rightarrow 0$ ) as the vector appearing in the definition of derivative at  $|x\rangle$ . Then

$$\begin{aligned} & \frac{\|f(|y\rangle) - f(|x\rangle) - \mathbf{D}f(x)(|y\rangle - |x\rangle)\|_2}{\|y - x\|_1} \\ &= \frac{\|f(x^1, \dots, x^j + h, \dots, x^n) - f(x^1, \dots, x^j, \dots, x^n) - h\mathbf{D}f(x)|e_j\rangle\|_2}{|h|} \end{aligned}$$

approaches zero as  $h \rightarrow 0$ , so that the  $i$ th component of the ratio also goes to zero. But the  $i$ th component of  $\mathbf{D}f(x)|e_j\rangle$  is simply  $a_j^i$ , the  $ij$ th component of the matrix of  $\mathbf{D}f(x)$ . Therefore,

$$\lim_{h \rightarrow 0} \frac{|f^i(x^1, \dots, x^j + h, \dots, x^n) - f^i(x^1, \dots, x^j, \dots, x^n) - ha_j^i|}{|h|} = 0,$$

which means that  $a_j^i = \partial f^i / \partial x^j$ . ■

The result of the example above can be stated as follows:

**30.1.3. Box.** For  $f : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^m$ , the matrix of  $\mathbf{D}f(x)$  in the standard basis of  $\mathbb{R}^n$  and  $\mathbb{R}^m$  is the **Jacobian matrix** of  $f$ .

differential and  
gradient of  $f$  at  $|x\rangle$

The case of  $\mathcal{H}_2 = \mathbb{R}$  deserves special attention. Let  $\mathcal{H}$  be a Hilbert space. Then  $\mathbf{D}f(x) \in \mathcal{L}(\mathcal{H}, \mathbb{R}) = \mathcal{H}^*$  is denoted by  $\mathbf{d}f(x)$  and renamed the **differential** of  $f$  at  $|x\rangle$ . Furthermore, through the inner product, one can identify  $\mathbf{d}f : \Omega \rightarrow \mathcal{H}^*$  with another map defined as follows:

**30.1.4. Definition.** Let  $\mathcal{H}$  be a Hilbert space and  $f : \mathcal{H} \supset \Omega \rightarrow \mathbb{R}$ . The **gradient**  $\nabla f$  of  $f$  is the map  $\nabla f : \Omega \rightarrow \mathcal{H}$  defined by

$$\langle \nabla f(x) | a \rangle \equiv \langle \mathbf{d}f(x), a \rangle \quad \forall |x\rangle \in \Omega, |a\rangle \in \mathcal{H}$$

where  $\langle, \rangle$  is the pairing  $\langle, \rangle : \mathcal{H}^* \times \mathcal{H} \rightarrow \mathbb{R}$  of  $\mathcal{H}$  and its dual.

Note that although  $f$  is not an element of  $\mathcal{H}^*$ ,  $\mathbf{d}f(x)$  is, for all points  $|x\rangle \in \Omega$  at which the differential is defined.

**30.1.5. Example.** Consider the function  $f : \mathcal{H} \rightarrow \mathbb{R}$  given by  $f(|x\rangle) = \|x\|^2$ . Since

$$\|y - x\|^2 = \|y\|^2 - \|x\|^2 - 2 \langle x | y - x \rangle$$

and since the derivative is unique, the reader may check that  $\mathbf{d}f(x) | a \rangle = 2 \langle x | a \rangle$ , or  $\nabla f(|x\rangle) = 2 |x\rangle$ . ■

Derivatives could be defined in terms of directions as well:

directional derivative

**30.1.6. Definition.** Let  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be Hilbert spaces. Let  $f : \mathcal{H}_1 \supset \Omega \rightarrow \mathcal{H}_2$  be any map and  $|x\rangle \in \Omega$ . We say that  $f$  has a derivative in the direction  $|a\rangle \in \mathcal{H}_1$  at  $|x\rangle$  if

$$\left. \frac{d}{dt} f(|x\rangle + t |a\rangle) \right|_{t=0}$$

exists. We call this element of  $\mathcal{H}_2$  the **directional derivative of  $f$  in the direction  $|a\rangle \in \mathcal{H}_1$  at  $|x\rangle$** .

The reader may verify that if  $f$  is differentiable at  $|x\rangle$  (in the context of Definition 30.1.1), then the directional derivative of  $f$  in any direction  $|a\rangle$  exists at  $|x\rangle$  and is given by

$$\left. \frac{d}{dt} f(|x\rangle + t |a\rangle) \right|_{t=0} = \mathbf{D}f(x) | a \rangle. \tag{30.3}$$

### 30.1.2 Functional Derivative

We now specialize to the Hilbert space of square-integrable functions  $\mathcal{L}^2(\Omega)$  for some open subset  $\Omega$  of some  $\mathbb{R}^m$ . We need to change our notation somewhat. Let us agree to denote the elements of  $\mathcal{L}^2(\Omega)$  by  $f, u$ , etc. Real-valued functions on  $\mathcal{L}^2(\Omega)$  will be denoted by  $\mathbf{L}, \mathbf{H}$ , etc. The  $m$ -tuples will be denoted by boldface lowercase letters. To summarize,

$$\mathbf{x}, \mathbf{y} \in \mathbb{R}^m, \quad f, u \in \mathcal{L}^2(\Omega) \Rightarrow f, u : \mathbb{R}^m \supset \Omega \rightarrow \mathbb{R},$$

$$\langle f | u \rangle = \int_{\Omega} f(\mathbf{x})u(\mathbf{x}) d^m x, \quad \mathbf{L}, \mathbf{H} : \mathcal{L}^2(\Omega) \rightarrow \mathbb{R}.$$

Furthermore, the evaluation of  $\mathbf{L}$  at  $u$  is denoted by  $\mathbf{L}[u]$ .

functional derivative  
or variational  
derivative

When dealing with the space of functions, the gradient of Definition 30.1.4 is called a **functional derivative** or **variational derivative** and denoted by  $\delta\mathbf{L}/\delta u$ . So

$$\left\langle \frac{\delta\mathbf{L}}{\delta u} \middle| f \right\rangle \equiv \int_{\Omega} \frac{\delta\mathbf{L}}{\delta u}(\mathbf{x})f(\mathbf{x}) d^m x = \left. \frac{d}{dt} \mathbf{L}[u + tf] \right|_{t=0}, \tag{30.4}$$

where we have used Equation (30.3). Note that by definition,  $\delta\mathbf{L}/\delta u$  is an element of the Hilbert space  $\mathcal{L}^2(\Omega)$ ; so, the integral of (30.4) makes sense. Equation (30.4) is frequently used to compute functional derivatives. An immediate consequence of Equation (30.4) is the following important result.

**30.1.7. Proposition.** *Let  $\mathbf{L} : \mathcal{L}^2(\Omega) \rightarrow \mathbb{R}$  for some  $\Omega \subset \mathbb{R}^m$ . If  $\mathbf{L}$  has an extremum at  $u$ , then*

$$\frac{\delta\mathbf{L}}{\delta u} = 0.$$

*Proof.* If  $\mathbf{L}$  has an extremum at  $u$ , then the RHS of (30.4) vanishes for any function  $f$ , in particular, for any orthonormal basis vector  $|e_i\rangle$ . Completeness of a basis now implies that the directional derivative must vanish (see Proposition 5.1.9).  $\square$

Just as in the case of partial derivatives, where some simple relations such as derivative of powers and products can be used to differentiate more complicated expressions, there are some primitive formulas involving functional derivatives that are useful in computing other more complicated expressions. Consider the

evaluation function

**evaluation function**

$$\mathbf{E}_y : \mathcal{L}^2(\Omega) \rightarrow \mathbb{R} \quad \text{given by} \quad \mathbf{E}_y[f] = f(\mathbf{y}).$$

Using Equation (30.4), we can easily compute the functional derivative of  $\mathbf{E}_y$ :

$$\int_{\Omega} \frac{\delta\mathbf{E}_y[u]}{\delta u}(\mathbf{x})f(\mathbf{x}) d^m x = \left. \frac{d}{dt} \mathbf{E}_y[u + tf] \right|_{t=0} = \left. \frac{d}{dt} \{u(\mathbf{y}) + tf(\mathbf{y})\} \right|_{t=0}$$

$$= f(\mathbf{y}) \Rightarrow \frac{\delta\mathbf{E}_y[u]}{\delta u}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{y}). \tag{30.5}$$

It is instructive to compare (30.5) with the similar formula in multivariable calculus, where real-valued functions  $f$  take a vector  $\mathbf{x}$  and give a real number. The analogue of the evaluation function is  $E_i$ , which takes a vector  $\mathbf{x}$  and gives the real number  $x^i$ , the  $i$ th component of  $\mathbf{x}$ . Using the definition of partial derivative, one readily shows that  $\partial E_i / \partial x^j = \delta_{ij}$ , which is (somewhat less precisely) written as  $\partial x^i / \partial x^j = \delta_{ij}$ . The same sort of imprecision is used to rewrite Equation (30.5) as

$$\frac{\delta u(\mathbf{y})}{\delta u(\mathbf{x})} \equiv \frac{\delta u_{\mathbf{y}}}{\delta u_{\mathbf{x}}} = \delta(\mathbf{x} - \mathbf{y}), \quad (30.6)$$

where we have turned the arguments into indices to make the analogy with the discrete case even stronger.

Another useful formula concerns *derivatives* of square-integrable functions. Let  $\mathbf{E}_{\mathbf{y},i}$  denote the evaluation of the derivative of functions with respect to the  $i$ th coordinate:

$$\mathbf{E}_{\mathbf{y},i} : \mathcal{L}^2(\Omega) \rightarrow \mathbb{R} \quad \text{given by} \quad \mathbf{E}_{\mathbf{y},i}(f) = \partial_i f(\mathbf{y}).$$

Then a similar argument as above will show that

$$\frac{\delta \mathbf{E}_{\mathbf{y},i}}{\delta u}(\mathbf{x}) = -\partial_i \delta(\mathbf{x} - \mathbf{y}), \quad \text{or} \quad \frac{\delta \partial_i u(\mathbf{y})}{\delta u(\mathbf{x})} = -\partial_i \delta(\mathbf{x} - \mathbf{y}),$$

and in general,

$$\frac{\delta \partial_{i_1 \dots i_k} u(\mathbf{y})}{\delta u(\mathbf{x})} = (-1)^k \partial_{i_1 \dots i_k} \delta(\mathbf{x} - \mathbf{y}). \quad (30.7)$$

Equation (30.7) holds only if the function  $f$ , the so-called *test function*, vanishes on  $\partial\Omega$ , the boundary of the region of integration. If it does not, then there will be a "surface term" that will complicate matters considerably. Fortunately, in most applications this surface term is *required* to vanish. So, let us adhere to the convention that

**30.1.8. Box.** All test functions  $f(\mathbf{x})$  appearing in the integral of Equation (30.4) are assumed to vanish at the boundary of  $\Omega$ .

For applications, we need to generalize the concept of functions on Hilbert spaces. First, it is necessary to consider maps from a Hilbert space to  $\mathbb{R}^n$ . For simplicity, we confine ourselves to the Hilbert space  $\mathcal{L}^2(\Omega)$ . Such a map  $\mathbf{H} : \mathcal{L}^2(\Omega) \rightarrow D \subset \mathbb{R}^n$ , for some subset  $D$  of  $\mathbb{R}^n$ , can be written in components

$$\mathbf{H} = (\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_n), \quad \text{where} \quad \mathbf{H}_i : \mathcal{L}^2(\Omega) \rightarrow \mathbb{R}, \quad i = 1, \dots, n.$$

Next, we consider an ordinary multivariable function  $L : \mathbb{R}^n \supset D \rightarrow \mathbb{R}$ , and use it to construct a new function on  $\mathcal{L}^2(\Omega)$ , the composite of  $L$  and  $\mathbf{H}$ :

$$L \circ \mathbf{H} : \mathcal{L}^2(\Omega) \rightarrow \mathbb{R}, \quad L \circ \mathbf{H}[u] = L(\mathbf{H}_1[u], \dots, \mathbf{H}_n[u]).$$

Then the functional derivative of  $L \circ \mathbf{H}$  can be obtained using the chain rule and noting that the derivative of  $L$  is the common partial derivative. It follows that

$$\frac{\delta L \circ \mathbf{H}[u]}{\delta u}(\mathbf{x}) = \left\{ \frac{\delta}{\delta u} L(\mathbf{H}_1[u], \dots, \mathbf{H}_n[u]) \right\}(\mathbf{x}) = \sum_{i=1}^n \partial_i L \frac{\delta \mathbf{H}_i}{\delta u}(\mathbf{x}), \tag{30.8}$$

where  $\partial_i L$  is the partial derivative of  $L$  with respect to its  $i$ th argument.

**30.1.9. Example.** Let  $L : (a, b) \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , be a function of three variables the first one of which is defined for the real interval  $(a, b)$ . Let  $\mathbf{H}_i : \mathcal{L}^2(a, b) \rightarrow \mathbb{R}$ ,  $i = 1, 2, 3$ , be defined by

$$\mathbf{H}_1[u] \equiv x, \quad \mathbf{H}_2[u] \equiv \mathbf{E}_x[u] = u(x), \quad \mathbf{H}_3[u] \equiv \mathbf{E}'_x[u] \equiv u'(x),$$

where  $\mathbf{E}_x$  is the evaluation function and  $\mathbf{E}'_x$  evaluates the derivative. It follows that  $L \circ \mathbf{H}[u] = L(x, u(x), u'(x))$ . Then, noting that  $\mathbf{H}_1[u]$  is independent of  $u$ , we have

$$\begin{aligned} \frac{\delta L \circ \mathbf{H}[u]}{\delta u}(y) &= \partial_1 L \frac{\delta \mathbf{H}_1[u]}{\delta u}(y) + \partial_2 L \frac{\delta \mathbf{E}_x[u]}{\delta u}(y) + \partial_3 L \frac{\delta \mathbf{E}'_x[u]}{\delta u}(y) \\ &= 0 + \partial_2 L \delta(y - x) - \partial_3 L \delta'(y - x) = \partial_2 L \delta(x - y) + \partial_3 L \delta'(x - y). \end{aligned}$$

This is normally written as

$$\frac{\delta L(x, u(x), u'(x))}{\delta u}(y) = \frac{\partial L}{\partial u} \delta(x - y) + \frac{\partial L}{\partial u'} \delta'(x - y), \tag{30.9}$$

which is the unintegrated version of the classical Euler-Lagrange equation for a single particle, to which we shall return shortly. ■

A generalization of the example above turns  $L$  into a function on  $\Omega \times \mathbb{R} \times \mathbb{R}^m$  with  $\Omega \subset \mathbb{R}^m$ , so that

$$L(x^1, \dots, x^m, u(\mathbf{x}), \partial_1 u(\mathbf{x}), \dots, \partial_m u(\mathbf{x})) \in \mathbb{R}, \quad \text{with } \mathbf{x} \in \mathbb{R}^m.$$

The functions  $\{\mathbf{H}_i\}_{i=1}^{2m+1}$  are defined as

$$\begin{aligned} \mathbf{H}_i[u] &\equiv x^i && \text{for } i = 1, 2, \dots, m, \\ \mathbf{H}_i[u] &\equiv \mathbf{E}_x[u] = u(\mathbf{x}) && \text{for } i = m + 1, \\ \mathbf{H}_i[u] &\equiv \mathbf{E}_{x,i}[u] = \partial_i u(\mathbf{x}) && \text{for } i = m + 2, \dots, 2m + 1, \end{aligned}$$

and lead to the equation

$$\frac{\delta L \circ \mathbf{H}[u]}{\delta u}(\mathbf{y}) = \partial_{m+1} L \delta(\mathbf{x} - \mathbf{y}) + \sum_{i=m+2}^{2m+1} \partial_i L \partial_i \delta(\mathbf{x} - \mathbf{y}), \tag{30.10}$$

which is the unintegrated version of the classical Euler-Lagrange equation for a field in  $m$  dimensions.

### 30.1.3 Variational Problems

The fundamental theme of the calculus of variations is to find functions that extremize an integral and are fixed on the boundary of the integration region. A prime example is the determination of the equation of the curve of minimum length in the  $xy$ -plane passing through two points  $(x_1, y_1)$  and  $(x_2, y_2)$ . Such a curve, written as  $y = u(x)$ , would minimize the integral

$$\text{int}[u] \equiv \int_{x_1}^{x_2} \sqrt{1 + [u'(x)]^2} dx, \quad u(x_1) = y_1, \quad u(x_2) = y_2. \quad (30.11)$$

Note that  $\text{int}$  takes a function and gives a real number, i.e.—if we restrict our functions to square-integrable ones— $\text{int}$  belongs to  $\mathcal{L}^2(x_1, x_2)$ . This is how contact is established between the calculus of variations and what we have studied so far in this chapter.

To be as general as possible, we allow the integral to contain derivatives up to the  $n$ th order. Then, using the notation of the previous chapter, we consider functions  $L$  on  $M^{(n)} \subset \Omega \times U^{(n)}$ , where we have replaced  $X$  with  $\Omega$ , so that  $M = \mathbb{R}^p \supset \Omega \times U \subset \mathbb{R}^q$ .

*n*th-order variational  
problem; Lagrangian;  
functional

**30.1.10. Definition.** *By an  $n$ th-order variational problem we mean finding the extremum of the real-valued function  $\mathbf{L} : \mathcal{L}^2(\Omega) \rightarrow \mathbb{R}$  given by*

$$\mathbf{L}[u] \equiv \int_{\Omega} L(x, u^{(n)}) d^p x, \quad (30.12)$$

where  $\Omega$  is a subset of  $\mathbb{R}^p \times \mathbb{R}^{qp^{(n)}}$ ,  $L$  is a real-valued function on  $\Omega$ , and  $p^{(n)} = (p + n)!/(n!p!)$ . In this context the function  $L$  is called the **Lagrangian** of the problem, and  $\mathbf{L}$  is called a **functional**.<sup>1</sup>

The solution to the variational problem is given by Proposition 30.1.7, moving the functional derivative inside the integral, and a straightforward (but tedious!) generalization of Equation (30.10) to include derivatives of order higher than one. Due to the presence of the integral, the Dirac delta function and all its derivatives will be integrated out. Before stating the solution of the variational problem, let us introduce a convenient operator.

Euler operator **30.1.11. Definition.** *For  $1 \leq \alpha \leq q$ , the  $\alpha$ th Euler operator is*

$$\mathbb{E}_{\alpha} \equiv \sum_J (-D)_J \frac{\partial}{\partial u_J^{\alpha}}, \quad (30.13)$$

where for  $J = (j_1, \dots, j_k)$ ,

$$(-D)_J \equiv (-1)^k D_J = (-D_{j_1})(-D_{j_2}) \cdots (-D_{j_k}),$$

and the sum extends over all multi-indices  $J = (j_1, \dots, j_k)$ , including  $J = 0$ .

<sup>1</sup>Do not confuse this functional with the linear functional of Chapter 1.

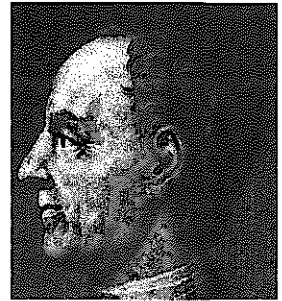
The negative signs are introduced because of the integration by parts involved in the evaluation of the derivatives of the delta function. Although the sum in Equation (30.13) extends over *all* multi-indices, only a finite number of terms in the sum will be nonzero, because any function on which the Euler operator acts depends on a finite number of derivatives.

Euler–Lagrange  
equations

**30.1.12. Theorem.** *If  $u$  is an extremal of the variational problem (30.12), then it must be a solution of the Euler–Lagrange equations*

$$\mathbb{E}_\alpha(L) \equiv \sum_J (-D)_J \frac{\partial L}{\partial u_J^\alpha} = 0, \quad \alpha = 1, \dots, q.$$

Leonhard Euler (1707–1783) was Switzerland’s foremost scientist and one of the three greatest mathematicians of modern times (Gauss and Riemann being the other two). He was perhaps the most prolific author of all time in any field. From 1727 to 1783 his writings poured out in a seemingly endless flood, constantly adding knowledge to every known branch of pure and applied mathematics, and also to many that were not known until he created them. He averaged about 800 printed pages a year throughout his long life, and yet he almost always had something worthwhile to say. The publication of his complete works was started in 1911, and the end is not in sight. This edition was planned to include 887 titles in 72 volumes, but since that time extensive new deposits of previously unknown manuscripts have been unearthed, and it is now estimated that more than 100 large volumes will be required for completion of the project. Euler evidently wrote mathematics with the ease and fluency of a skilled speaker discoursing on subjects with which he is intimately familiar. His writings are models of relaxed clarity. He never condensed, and he reveled in the rich abundance of his ideas and the vast scope of his interests. The French physicist Arago, in speaking of Euler’s incomparable mathematical facility, remarked that “He calculated without apparent effort, as men breathe, or as eagles sustain themselves in the wind.” He suffered total blindness during the last 17 years of his life, but with the aid of his powerful memory and fertile imagination, and with assistants to write his books and scientific papers from dictation, he actually increased his already prodigious output of work.



Euler was a native of Basel and a student of Johann Bernoulli at the University, but he soon outstripped his teacher. He was also a man of broad culture, well versed in the classical languages and literatures (he knew the Aeneid by heart), many modern languages, physiology, medicine, botany, geography, and the entire body of physical science as it was known in his time. His personal life was as placid and uneventful as is possible for a man with 13 children.

Though he was not himself a teacher, Euler has had a deeper influence on the teaching of mathematics than any other person. This came about chiefly through his three great treatises: *Introductio in Analysis Infnitorum* (1748); *Institutiones Calculi Differentialis* (1755); and *Institutiones Calculi Integralis* (1768–1794). There is considerable truth in the old saying

that all elementary and advanced calculus textbooks since 1748 are essentially copies of Euler or copies of copies of Euler. These works summed up and codified the discoveries of his predecessors, and are full of Euler's own ideas. He extended and perfected plane and solid analytic geometry, introduced the analytic approach to trigonometry, and was responsible for the modern treatment of the functions  $\ln x$  and  $e^x$ . He created a consistent theory of logarithms of negative and imaginary numbers, and discovered that  $\ln x$  has an infinite number of values. It was through his work that the symbols  $e$ ,  $\pi$ , and  $i = \sqrt{-1}$  became common currency for all mathematicians, and it was he who linked them together in the astonishing relation  $e^{i\pi} = -1$ . Among his other contributions to standard mathematical notation were  $\sin x$ ,  $\cos x$ , the use of  $f(x)$  for an unspecified function, and the use of  $\sum$  for summation.

His work in all departments of analysis strongly influenced the further development of this subject through the next two centuries. He contributed many important ideas to differential equations, including substantial parts of the theory of second-order linear equations and the method of solution by power series. He gave the first systematic discussion of the calculus of variations, which he founded on his basic differential equation for a minimizing curve. He discovered the integral defining the gamma function and developed many of its applications and special properties. He also worked with Fourier series, encountered the Bessel functions in his study of the vibrations of a stretched circular membrane, and applied Laplace transforms to solve differential equations—all before Fourier, Bessel, and Laplace were born.

E. T. Bell, the well-known historian of mathematics, observed that "One of the most remarkable features of Euler's universal genius was its equal strength in both of the main currents of mathematics, the continuous and the discrete." In the realm of the discrete, he was one of the originators of number theory and made many far-reaching contributions to this subject throughout his life. In addition, the origins of topology—one of the dominant forces in modern mathematics—lie in his solution of the Königsberg bridge problem and his formula  $V - E + F = 2$  connecting the numbers of vertices, edges, and faces of a simple polyhedron.

The distinction between pure and applied mathematics did not exist in Euler's day, and for him the entire physical universe was a convenient object whose diverse phenomena offered scope for his methods of analysis. The foundations of classical mechanics had been laid down by Newton, but Euler was the principal architect. In his treatise of 1736 he was the first to explicitly introduce the concept of a mass-point, or particle, and he was also the first to study the acceleration of a particle moving along any curve and to use the notion of a vector in connection with velocity and acceleration. His continued successes in mathematical physics were so numerous, and his influence was so pervasive, that most of his discoveries are not credited to him at all and are taken for granted in the physics community as part of the natural order of things. However, we do have Euler's angles for the rotation of a rigid body, and the all-important *Euler-Lagrange equation* of variational dynamics.

Euler was the Shakespeare of mathematics—universal, richly detailed, and inexhaustible.

---



For the special case of  $p = q = 1$ , the Euler operator becomes

$$\mathbb{E} = \frac{\partial}{\partial u} + \sum_{j=1}^{\infty} (-D_x)^j \frac{\partial}{\partial u_j} = \frac{\partial}{\partial u} - D_x \frac{\partial}{\partial u_x} + D_x^2 \frac{\partial}{\partial u_{xx}} - \dots,$$

where  $D_x$  is the *total* derivative with respect to  $x$ , and  $u_j$  is the  $j$ th derivative of  $u$  with respect to  $x$ ; and the Euler–Lagrange equation for the variational problem

$$\mathbb{L}[u] \equiv \int_a^b L(x, u^{(n)}) dx$$

becomes

$$\mathbb{E}(L) = \frac{\partial L}{\partial u} + \sum_{j=1}^n (-1)^j D_x^j \frac{\partial L}{\partial u_j} = 0. \tag{30.14}$$

Since  $L$  carries derivatives up to the  $n$ -th order and each  $D_x$  carries one derivative, we conclude that Equation (30.14) is a  $2n$ -th order ODE.

**30.1.13. Example.** The variational problem of Equation (30.11) has a Lagrangian

$$L(u, u^{(n)}) = L(u, u^{(1)}) = \sqrt{1 + u_x^2},$$

which is a function of the first derivative only. So, the Euler–Lagrange equation takes the form

$$0 = -D_x \frac{\partial L}{\partial u_x} = -D_x \left( \frac{u_x}{\sqrt{1 + u_x^2}} \right) = -\frac{d}{dx} \left( \frac{u_x}{\sqrt{1 + u_x^2}} \right) = -\frac{u_{xx}}{(1 + u_x^2)^{3/2}},$$

or  $u_{xx} = 0$ , so that  $u = f(x) = c_1x + c_2$ . The solution to the variational problem is a straight line passing through the two points  $(x_1, y_1)$  and  $(x_2, y_2)$ . ■

The variational problem is a problem involving only the first functional derivative, or the *first variation*. We know from calculus that the first derivative by itself cannot determine the nature of the extremum. To test whether the point in question is maximum or minimum, we need all the second derivatives (see Example 4.7.4). One uses these derivatives to expand the functional in a Taylor series up to the second order. The sign of the second order contribution determines whether the functional is maximum or minimum at the extremal point. In analogy with Example 4.7.4, we expand  $\mathbb{L}[u]$  about  $f$  up to the second-order derivative:

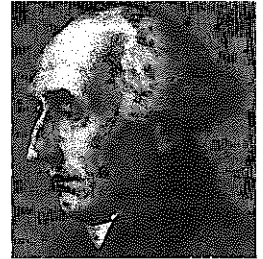
$$\begin{aligned} \mathbb{L}[u] &= \mathbb{L}[f] + \int_{\Omega} d^p y \frac{\delta \mathbb{L}}{\delta u(\mathbf{y})} \Big|_{u=f} (u(\mathbf{y}) - f(\mathbf{y})) \\ &\quad + \frac{1}{2} \int_{\Omega} d^p y \int_{\Omega} d^p y' \frac{\delta^2 \mathbb{L}}{\delta u(\mathbf{y}) \delta u(\mathbf{y}')} \Big|_{u=f} (u(\mathbf{y}) - f(\mathbf{y}))(u(\mathbf{y}') - f(\mathbf{y}')). \end{aligned}$$

The first variation is not sufficient for a full knowledge of the nature of the extremum!

The integrals have replaced the sums of the discrete case of Taylor expansion of the multivariable functions. Since we are interested in comparing  $u$  with the  $f$  that extremizes the functional, the second term vanishes and we get

$$\begin{aligned} \mathbf{L}[u] = \mathbf{L}[f] + \frac{1}{2} \int_{\Omega} d^p y \int_{\Omega} d^p y' \frac{\delta^2 \mathbf{L}}{\delta u(\mathbf{y}) \delta u(\mathbf{y}')} \Big|_{u=f} \\ \cdot [(u(\mathbf{y}) - f(\mathbf{y}))(u(\mathbf{y}') - f(\mathbf{y}'))]. \end{aligned} \quad (30.15)$$

**Joseph Louis Lagrange** (1736–1813) was born Giuseppe Luigi Lagrangia but adopted the French version of his name. He was the eldest of eleven children, most of whom did not reach adulthood. His father destined him for the law—a profession that one of his brothers later pursued—and Lagrange offered no objections. But having begun the study of physics and geometry, he quickly became aware of his talents and henceforth devoted himself to the exact sciences. Attracted first by geometry, at the age of seventeen he turned to analysis, then a rapidly developing field.



In 1755, in a letter to the geometer Giulio da Fagnano, Lagrange speaks of one of Euler's papers published at Lausanne and Geneva in 1744. The same letter shows that as early as the end of 1754 Lagrange had found interesting results in this area, which was to become the *calculus of variations* (a term coined by Euler in 1766). In the same year, Lagrange sent Euler a summary, written in Latin, of the purely analytical method that he used for this type of problem. Euler replied to Lagrange that he was very interested in the technique. Lagrange's merit was likewise recognized in Turin; and he was named, by a royal decree, professor at the Royal Artillery School with an annual salary of 250 crowns—a sum never increased in all the years he remained in his native country. Many years later, in a letter to d'Alembert, Lagrange confirmed that this method of maxima and minima was the first fruit of his studies—he was only nineteen when he devised it—and that he regarded it as his best work in mathematics. In 1756, in a letter to Euler that has been lost, Lagrange, applying the calculus of variations to mechanics, generalized Euler's earlier work on the trajectory described by a material point subject to the influence of central forces to an arbitrary system of bodies, and derived from it a procedure for solving all the problems of dynamics.

In 1757 some young Turin scientists, among them Lagrange, founded a scientific society that was the origin of the Royal Academy of Sciences of Turin. One of the main goals of this society was the publication of a miscellany in French and Latin, *Miscellanea Taurinensia ou Mélanges de Turin*, to which Lagrange contributed fundamentally. These contributions included works on the calculus of variations, probability, vibrating strings, and the principle of least action.

To enter a competition for a prize, in 1763 Lagrange sent to the Paris Academy of Sciences a memoir in which he provided a satisfactory explanation of the translational motion of the moon. In the meantime, the Marquis Caraccioli, ambassador from the kingdom of Naples to the court of Turin, was transferred by his government to London. He took along the young Lagrange, who until then seems never to have left the immediate vicinity of Turin. Lagrange was warmly received in Paris, where he had been preceded by his memoir on lunar

liberation. He may perhaps have been treated too well in the Paris scientific community, where austerity was not a leading virtue. Being of a delicate constitution, Lagrange fell ill and had to interrupt his trip. In the spring of 1765 Lagrange returned to Turin by way of Geneva.

In the autumn of 1765 d'Alembert, who was on excellent terms with Frederick II of Prussia, and familiar with Lagrange's work through *Mélanges de Turin*, suggested to Lagrange that he accept the vacant position in Berlin created by Euler's departure for St. Petersburg. It seems quite likely that Lagrange would gladly have remained in Turin had the court of Turin been willing to improve his material and scientific situation. On 26 April, d'Alembert transmitted to Lagrange the very precise and advantageous propositions of the king of Prussia. Lagrange accepted the proposals of the Prussian king and, not without difficulties, obtained his leave through the intercession of Frederick II with the king of Sardinia. Eleven months after his arrival in Berlin, Lagrange married his cousin Vittoria Conti who died in 1783 after a long illness. With the death of Frederick II in August 1786 he also lost his strongest support in Berlin. Advised of the situation, the princes of Italy zealously competed in attracting him to their courts. In the meantime the French government decided to bring Lagrange to Paris through an advantageous offer. Of all the candidates, Paris was victorious.

Lagrange left Berlin on 18 May 1787 to become *pensionnaire vétérane* of the Paris Academy of Sciences, of which he had been a foreign associate member since 1772. Warmly welcomed in Paris, he experienced a certain lassitude and did not immediately resume his research. Yet he astonished those around him by his extensive knowledge of metaphysics, history, religion, linguistics, medicine, and botany.

In 1792 Lagrange married the daughter of his colleague at the Academy, the astronomer Pierre Charles Le Monnier. This was a troubled period, about a year after the flight of the king and his arrest at Varennes. Nevertheless, on 3 June the royal family signed the marriage contract "as a sign of its agreement to the union." Lagrange had no children from this second marriage, which, like the first, was a happy one.

When the academy was suppressed in 1793, many noted scientists, including Lavoisier, Laplace, and Coulomb were purged from its membership; but Lagrange remained as its chairman. For the next ten years, Lagrange survived the turmoil of the aftermath of the French Revolution, but by March of 1813, he became seriously ill. He died on the morning of 11 April 1813, and three days later his body was carried to the Panthéon. The funeral oration was given by Laplace in the name of the Senate.

A straight line segment is indeed the *shortest* distance between two points.

**30.1.14. Example.** Let us apply Equation (30.15) to the extremal function of Example 30.1.13 to see if the line is truly the *shortest* distance between two points. The first functional derivative, obtained using Equation (30.9), is simply  $\mathbb{E}(L)$ :

$$\frac{\delta L}{\delta u(y)} = \mathbb{E}(L) = -\frac{u_{yy}}{(1+u_y^2)^{3/2}}$$

To find the second variational derivative, we use the basic relations (30.6), (30.7), and the chain rule (30.10):

$$\left. \frac{\delta^2 L}{\delta u(y') \delta u(y)} \right|_{u=f} = -\frac{\delta}{\delta u(y')} \left[ \frac{u_{yy}}{(1+u_y^2)^{3/2}} \right] \Big|_{u=f}$$

$$\begin{aligned}
 &= -\left\{ (1+u_y^2)^{-3/2} \frac{\delta u_{yy}}{\delta u(y')} - u_{yy} \frac{3}{2} (1+u_y^2)^{-5/2} 2u_y \frac{\delta u_y}{\delta u(y')} \right\} \Big|_{u=f} \\
 &= -\frac{\delta''(y-y')}{(1+u_y^2)^{3/2}} \Big|_{u=f} = -\frac{\delta''(y-y')}{(1+c_1^2)^{3/2}},
 \end{aligned}$$

because  $u_{yy} = 0$  and  $u_y = c_1$  when  $u = f$ . Inserting this in Equation (30.15), we obtain

$$\begin{aligned}
 \mathbf{L}[u] &= \mathbf{L}[f] - \frac{1}{2(1+c_1^2)^{3/2}} \int_{x_1}^{x_2} dy \int_{x_1}^{x_2} dy' \delta''(y-y')(u(y) - f(y))(u(y') - f(y')) \\
 &= \mathbf{L}[f] - \frac{1}{2(1+c_1^2)^{3/2}} \int_{x_1}^{x_2} dy (u(y) - f(y)) \frac{d^2}{dy^2} (u(y) - f(y)).
 \end{aligned}$$

The last integral can be integrated by parts, with the result

$$\underbrace{(u(y) - f(y)) \frac{d}{dy} (u(y) - f(y)) \Big|_{x_1}^{x_2}}_{=0 \text{ because } u(x_i) = f(x_i), i = 1, 2} - \int_{x_1}^{x_2} dy \left[ \frac{d}{dy} (u(y) - f(y)) \right]^2.$$

Therefore,

$$\mathbf{L}[u] = \mathbf{L}[f] + \frac{1}{2(1+c_1^2)^{3/2}} \underbrace{\int_{x_1}^{x_2} dy \left[ \frac{d}{dy} (u(y) - f(y)) \right]^2}_{\text{always positive}}.$$

It follows that  $\mathbf{L}[f] < \mathbf{L}[u]$ , i.e., that  $f$  indeed gives the shortest distance. ■

**30.1.15. Example.** In the special theory of relativity, the element of the invariant “length,” or proper time, is given by  $\sqrt{dt^2 - dx^2}$ . Thus, the total proper time between two events  $(t_1, x_1)$  and  $(t_2, x_2)$  is given by

$$\mathbf{L}[x] = \int_{t_1}^{t_2} \sqrt{1 - x_t^2} dt, \quad x_t \equiv \frac{dx}{dt}.$$

The extremum of this variational problem is exactly the same as in the previous example, the only difference being a sign. In fact, the reader may verify that

$$\frac{\delta \mathbf{L}[x]}{\delta x(s)} = \mathbb{E}(L) = \frac{x_{ss}}{(1 - x_s^2)^{3/2}},$$

connection between  
variational problem  
and the twin paradox

and therefore,  $x = f(t) = c_1 t + c_2$  extremizes the functional. The second variational derivative can be obtained as before. It is left for the reader to show that in the case at hand,  $\mathbf{L}[f] > \mathbf{L}[x]$ , i.e., that  $f$  gives the *longest* proper time. Since the function  $f(t) = c_1 t + c_2$  corresponds to an inertial (unaccelerated) observer, we conclude that

**30.1.16. Box.** Accelerated observers measure a shorter proper time between any two events than inertial observers.

This is the content of the famous **twin paradox**, in which the twin who goes to a distant galaxy and comes back (therefore being accelerated) will return younger than her (unaccelerated) twin. ■

### 30.1.4 Divergence and Null Lagrangians

The variational problem integrates a Lagrangian over a region  $\Omega$  of  $\mathbb{R}^p$ . If the Lagrangian happens to be the divergence of a function that vanishes at the boundary of  $\Omega$ , the variational problem becomes trivial, because all functions will extremize the functional. We now study such Lagrangians in more detail.

**30.1.17. Definition.** Let  $\{F_i : M^{(n)} \rightarrow \mathbb{R}\}_{i=1}^p$  be functions on  $M^{(n)}$ , and  $\mathbf{F} = (F_1, \dots, F_p)$ . The *total divergence* of  $\mathbf{F}$  is defined to be<sup>2</sup>

$$\mathbf{D} \cdot \mathbf{F} \equiv \sum_{j=1}^p D_j F_j,$$

where  $D_j$  is the total derivative with respect to  $x^j$ .

Now suppose that the Lagrangian  $L(x, u^{(n)})$  can be written as the divergence of some  $p$ -tuple  $\mathbf{F}$ . Then by the divergence theorem,

$$\mathbf{L}[u] = \int_{\Omega} L(x, u^{(n)}) d^p x = \int_{\Omega} \mathbf{D} \cdot \mathbf{F} d^p x = \int_{\partial\Omega} \mathbf{F} \cdot d\mathbf{a}$$

for any  $u = f(x)$  and any domain  $\Omega$ . It follows that  $\mathbf{L}[f]$  depends on the behavior of  $f$  only at the boundary. Since in a typical problem no variation takes place at the boundary, all functions that satisfy the boundary conditions will be solutions of the variational problem, i.e., they satisfy the Euler–Lagrange equation. Lagrangians that satisfy the Euler–Lagrange equation for all  $u$  and  $x$  are called **null Lagrangians**. It turns out that null Lagrangians are the *only* such solutions of the Euler–Lagrange equation (for a proof, see [Olve 86, pp. 252–253]).

**30.1.18. Theorem.** A function  $L(x, u^{(n)})$  satisfies  $\mathbb{E}(L) \equiv 0$  for all  $x$  and  $u$  if and only if  $L = \mathbf{D} \cdot \mathbf{F}$  for some  $p$ -tuple of functions  $\mathbf{F} = (F_1, \dots, F_p)$  of  $x, u$ , and the derivatives of  $u$ .

In preparation for the investigation of symmetries of the variational problems, we look into the effect of a change of variables on the variational problem and the Euler operator. This is important, because the variational problem should be independent of the variables chosen. Let

$$\tilde{x} = \Psi(x, u), \quad \tilde{u} = \Phi(x, u) \tag{30.16}$$

be any change of variables. Then by prolongation, we also have  $\tilde{u}^{(n)} = \Phi^{(n)}(x, u^{(n)})$  for the derivatives. Substituting  $u = f(x)$  and all its prolongations in terms of the new variables, the functional

$$\mathbf{L}[f] = \int_{\Omega} L(x, \text{pr}^{(n)} f(x)) d^p x$$

<sup>2</sup>The reader need not be concerned about lack of consistency in the location of indices (upper vs. lower), because we are dealing with indexed objects, such as  $F_i$ , which are *not* tensors!

will be transformed into

$$\tilde{\mathbb{L}}[\tilde{f}] = \int_{\tilde{\Omega}} \tilde{L}(\tilde{x}, \text{pr}^{(n)} \tilde{f}(\tilde{x})) d^p \tilde{x},$$

where the transformed domain, defined by

$$\tilde{\Omega} = \{\tilde{x} = \Psi(x, f(x)) \mid x \in \Omega\},$$

will depend not only on the original domain  $\Omega$ , but also on the function  $f$ . The new Lagrangian is then related to the old one by the change of variables formula for multiple integrals:

$$L(x, \text{pr}^{(n)} f(x)) = \tilde{L}(\tilde{x}, \text{pr}^{(n)} \tilde{f}(\tilde{x})) \det J(x, \text{pr}^{(1)} f(x)), \quad (30.17)$$

where  $J$  is the Jacobian matrix of the change of variables induced by the function  $f$ .

Starting with Equations (30.16) and (30.17), one can obtain the transformation formula for the Euler operator stated below. The details can be found in [Olive 86, pp. 254–255].

**30.1.19. Theorem.** *Let  $L(x, u^{(n)})$  and  $\tilde{L}(\tilde{x}, \tilde{u}^{(n)})$  be two Lagrangians related by the change of variable formulas (30.16) and (30.17). Then*

$$\mathbb{E}_\alpha(L) = \sum_{\beta=1}^q F_{\alpha\beta}(x, u^{(1)}) \tilde{\mathbb{E}}_\beta(\tilde{L}), \quad \alpha = 1, \dots, q$$

where  $\tilde{\mathbb{E}}_\beta$  is the Euler operator associated with the new variables, and

$$F_{\alpha\beta} \equiv \det \begin{pmatrix} D_1 \Psi^1 & \dots & D_p \Psi^1 & \partial \Psi^1 / \partial u^\alpha \\ \vdots & & \vdots & \vdots \\ D_1 \Psi^p & \dots & D_p \Psi^p & \partial \Psi^p / \partial u^\alpha \\ D_1 \Phi^\beta & \dots & D_p \Phi^\beta & \partial \Phi^\beta / \partial u^\alpha \end{pmatrix}.$$

## 30.2 Symmetry Groups of Variational Problems

In the theory of fields, as well as in mechanics, condensed matter theory, and statistical mechanics, the starting point is usually a Lagrangian. The variational problem of this Lagrangian gives the classical equations of motion, and its symmetries lead to the important conservation laws.

**30.2.1. Definition.** *A local group of transformations  $G$  acting on  $M \subset \Omega_0 \times U$  is a **variational symmetry group** of the functional*

$$\mathbb{L}[u] = \int_{\Omega_0} L(x, u^{(n)}) d^p x \quad (30.18)$$

variational symmetry  
group

if whenever (the closure of)  $\Omega$  lies in  $\Omega_0$ ,  $f$  is a function over  $\Omega$  whose graph is in  $M$ , and  $g \in G$  is such that  $\tilde{f} = g \cdot f$  is a single-valued function defined over  $\tilde{\Omega}$ , then

$$\int_{\tilde{\Omega}} L(\tilde{x}, \text{pr}^{(n)} \tilde{f}(\tilde{x})) d^p \tilde{x} = \int_{\Omega} L(x, \text{pr}^{(n)} f(x)) d^p x. \tag{30.19}$$

“Symmetry of the Lagrangian” is really the symmetry group of the variational problem!

In the physics community, the symmetry group of the variational problem is (somewhat erroneously) called the **symmetry of the Lagrangian**. Note that if we had used  $\tilde{L}$  in the LHS of Equation (30.19), we would have obtained an identity valid for *all* lagrangians because of Equation (30.17) and the formula for the change in the volume element of integration. Only *symmetric Lagrangians* will satisfy Equation (30.19).

As we have experienced so far, the action of a group can be very complicated and very nonlinear. On the other hand, the *infinitesimal* action simplifies the problem considerably. Fortunately, we have (see [Olve 86, pp. 257–258] for a proof).

**30.2.2. Theorem.** A local group of transformations  $G$  acting on  $M \subset \Omega_0 \times U$  is a variational symmetry group of the functional (30.18) if and only if

$$\text{pr}^{(n)} \mathbf{v}(L) + LD \cdot \mathbf{X} = 0 \tag{30.20}$$

for all  $(x, u^{(n)}) \in M^{(n)}$  and every infinitesimal generator

$$\mathbf{v} = \sum_{i=1}^p X^i(x, u) \frac{\partial}{\partial x^i} + \sum_{\alpha=1}^q U^\alpha(x, u) \frac{\partial}{\partial u^\alpha}$$

of  $G$ , where  $\mathbf{X} \equiv (X^1, \dots, X^p)$ .

**30.2.3. Example.** Consider the case of  $p = 1 = q$ , and assume that the Lagrangian is independent of  $x$  but depends on  $u \in \mathcal{L}^2(a, b)$  and its first derivative. Then the variational problem takes the form

$$\mathbb{L}[u] = \int_a^b L(u^{(1)}) dx \equiv \int_a^b L(u, u_x) dx.$$

Since derivatives are independent of translations, we expect translations to be part of the symmetry group of this variational problem. Let us verify this. The infinitesimal generator of translation is  $\partial_x$ , which is its own prolongation. Therefore, with  $X = 1$  and  $U = 0$ , it follows that

$$\text{pr}^{(1)} \mathbf{v}(L) + LD \cdot \mathbf{X} = \partial_x L + LD_x X = 0 + 0 = 0. \quad \blacksquare$$

**30.2.4. Example.** As a less trivial case, consider the proper time of Example 30.1.15. Lorentz transformations generated by<sup>3</sup>  $\mathbf{v} = u\partial_x + x\partial_u$  are symmetries of that variational

<sup>3</sup>In order to avoid confusion in applying formula (30.20), we use  $x$  (instead of  $t$ ) as the independent variable and  $u$  (instead of  $x$ ) as the dependent variable.

problem. We can verify this by noting that the first prolongation of  $\mathbf{v}$  is, as the reader is urged to verify,

$$\text{pr}^{(1)}\mathbf{v} = \mathbf{v} + (1 - u_x^2) \frac{\partial}{\partial u_x}.$$

Therefore,

$$\text{pr}^{(1)}\mathbf{v}(L) = 0 + 0 + ((1 - u_x^2))^{\frac{1}{2}}(-2u_x) \frac{1}{\sqrt{1 - u_x^2}} = -u_x \sqrt{1 - u_x^2}.$$

On the other hand, since  $X = u$  and  $U = x$ ,

$$LD_x(X) = \sqrt{1 - u_x^2} D_x(u) = \sqrt{1 - u_x^2} u_x,$$

so that Equation (30.20) is satisfied. ■

In the last chapter, we studied the symmetries of the DEs in some detail. This chapter introduces us to a particular DE that arises from a variational problem, namely, the Euler–Lagrange equation. The natural question to ask now is, How does the variational symmetry manifest itself in the Euler–Lagrange equation? Barring some technical difficulties, we note that for any change of variables, if  $u = f(x)$  is an extremal of the variational problem  $L[u]$ , then  $\tilde{u} = \tilde{f}(\tilde{x})$  is an extremal of the variational problem  $\tilde{L}[\tilde{u}]$ . In particular, if the change is achieved by the action of the variational symmetry group,  $(\tilde{x}, \tilde{u}) = g \cdot (x, u)$  for some  $g \in G$ , then  $\tilde{L}[\tilde{u}] = L[\tilde{u}]$ , and  $g \cdot f$  is also an extremal of  $L$ . We thus have

**30.2.5. Theorem.** *If  $G$  is the variational symmetry group of a functional, then  $G$  is also the symmetry group of the associated Euler–Lagrange equations.*

The converse is *not* true! There are symmetry groups of the Euler–Lagrange equations that are not the symmetry group of the variational problem. Problem 30.8 illustrates this for  $p = 3$ ,  $q = 1$ , and the functional

$$L[u] = \frac{1}{2} \iiint (u_t^2 - u_x^2 - u_y^2) dx dy dt, \tag{30.21}$$

whose Euler–Lagrange equation is the wave equation. The reader is asked to show that while the rotations and Lorentz boosts of Table 29.3 are variational symmetries, the dilatations and inversions (special conformal transformations) are not.

We now treat the case of  $p = 1 = q$ , whose Euler–Lagrange equation is an ODE. Recall that the knowledge of a symmetry group of an ODE led to a reduction in the order of that ODE. Let us see what happens in the present case. Suppose  $\mathbf{v} = X\partial_x + U\partial_u$  is the infinitesimal generator of a 1-parameter group of variational symmetries of  $L$ . By an appropriate coordinate transformation from  $(x, u)$  to  $(y, w)$ , as in Section 29.5,  $\mathbf{v}$  will reduce to  $\partial/\partial w$ , whose prolongation is also  $\partial/\partial w$ . In terms of the new coordinates, Equation (30.20) will reduce to

Symmetries of the Euler–Lagrange equations are not necessarily the symmetries of the corresponding variational problem!



$\partial \tilde{L} / \partial w = 0$ ; i.e., the new Lagrangian is independent of  $w$ , and the Euler–Lagrange equation (30.14) becomes

$$0 = \mathbb{E}(L) = \sum_{j=1}^n (-1)^j D_y^j \frac{\partial \tilde{L}}{\partial w_j} = (-D_y) \left[ \sum_{j=0}^{n-1} (-D_y)^j \frac{\partial \tilde{L}}{\partial w_{j+1}} \right]. \tag{30.22}$$

Therefore, the expression in the brackets is some constant  $\lambda$  (because  $D_y$  is a total derivative). Furthermore, if we introduce  $v = w_y$ , the expression in the brackets becomes the Euler–Lagrange equation of the variational problem

$$\hat{\mathbb{L}}[v] = \int \hat{L}(y, v^{(n-1)}) dy, \quad \text{where} \quad \hat{L}(y, v^{(n-1)}) = \tilde{L}(y, w_y, \dots, w_n),$$

and every solution  $w = f(y)$  of the original  $(2n)$ th-order Euler–Lagrange equation corresponds to the  $(2n - 2)$ nd-order equation

$$\hat{\mathbb{E}}(\hat{L}) = \frac{\partial \hat{L}}{\partial y} + \sum_{j=1}^{n-1} (-D_y)^j \frac{\partial \hat{L}}{\partial v_j} = \lambda. \tag{30.23}$$

Moreover, this equation can be written as the Euler–Lagrange equation for

$$\hat{\mathbb{L}}_\lambda[v] = \int [\hat{L}(y, v^{(n-1)}) - \lambda v] dy,$$

Lagrange multiplier

and  $\lambda$  can be thought of as a **Lagrange multiplier**, so that in analogy with the multivariable extremal problem,<sup>4</sup> the extremization of  $\hat{\mathbb{L}}_\lambda[v]$  becomes equivalent to that of  $\hat{\mathbb{L}}[v]$  subject to the constraint  $\int v dy = 0$ . We summarize the foregoing discussion in the following theorem.

**30.2.6. Theorem.** *Let  $p = 1 = q$ , and  $\mathbb{L}[u]$  an  $n$ th-order variational problem with a 1-parameter group of variational symmetries  $G$ . Then there exists a one-parameter family of variational problems  $\hat{\mathbb{L}}_\lambda[v]$  of order  $n - 1$  such that every solution of the Euler–Lagrange equation for  $\mathbb{L}[u]$  can be found by integrating the solutions to the Euler–Lagrange equation for  $\hat{\mathbb{L}}_\lambda[v]$ .*

Thus, we have the following important result:

**30.2.7. Box.** *A 1-parameter variational symmetry of a functional reduces the order of the corresponding Euler–Lagrange equation by two.*

This conclusion is to be contrasted with the symmetry of ODEs, where each 1-parameter group of symmetries reduces the order of the ODE by 1. It follows from

<sup>4</sup>See [Math 70, pp. 331–341] for a discussion of Lagrange multipliers and their use in variational techniques, especially those used in approximating solutions of the Schrödinger equation.

Box 30.2.7 that the ODEs of order  $2n$  derived from a variational problem—the Euler–Lagrange equation—are special.

**30.2.8. Example.** A first-order variational problem with a 1-parameter group of symmetries can be integrated out. By transforming to a new coordinate system, we can always assume that the Lagrangian is independent of the *dependent variable* (see Proposition 29.5.1). The Euler–Lagrange equation in this case becomes

$$0 = \mathbb{E}(L) = \underbrace{\frac{\partial L}{\partial u}}_{=0} - D_x \frac{\partial L}{\partial u_x} \Rightarrow \frac{\partial L}{\partial u_x}(x, u_x) = \lambda.$$

Solving this implicit relation, we get  $u_x = F(x, \lambda)$ , which can be integrated to give  $u$  as a function of  $x$  (and  $\lambda$ ). ■

The procedure can be generalized to  $r$ -parameter symmetry groups, but the order cannot be expected to be reduced by 2 unless the group is abelian. We shall not pursue this matter here, but ask the reader to refer to Problem 30.9.

### 30.3 Conservation Laws and Noether’s Theorem

A conserved physical quantity is generally defined as a quantity whose flux through any arbitrary closed surface is equal to (the negative of) the rate of depletion of the quantity in the volume enclosed. This statement, through the use of the divergence theorem, translates into a relation connecting the time rate of change of the density and the divergence of the current corresponding to the physical quantity. Treating time and space coordinates as independent variables and extending to  $p$  independent variables, we have the following:

current density and conservation law

**30.3.1. Definition.** A *conservation law* for a system of differential equations  $\Delta(x, u^{(n)}) = 0$  is a divergence expression  $\mathbf{D} \cdot \mathbf{J} = 0$  valid for all solutions  $u = f(x)$  of the system. Here,

$$\mathbf{J} \equiv (J_1(x, u^{(n)}), J_2(x, u^{(n)}), \dots, J_p(x, u^{(n)}))$$

is called *current density*.

constant of the motion, or first integral of a system of ODEs

For  $p = 1 = q$ , i.e., for a system of ODEs, a conservation law takes the form  $D_x J(x, u^{(n)}) = 0$  for all solutions  $u = f(x)$  of the system. This requires  $J(x, u^{(n)})$  to be a constant, i.e., that  $J(x, u^{(n)})$  be a **constant of the motion**, or, as it is sometimes called, the *first integral* of the system.

In order to understand conservation laws, we need to get a handle on those conservation laws that are trivially satisfied.

trivial conservation law of the first kind

**30.3.2. Definition.** If the current density  $\mathbf{J}$  itself vanishes for all solutions  $u = f(x)$  of the system  $\Delta(x, u^{(n)}) = 0$ , then  $\mathbf{D} \cdot \mathbf{J} = 0$  is called a *trivial conservation law of the first kind*.

To eliminate this kind of triviality, one solves the system and its prolongations  $\Delta^{(k)}(x, u^{(n)}) = 0$  for some of the variables  $u_j^\alpha$  in terms of the remaining variables and substitutes the latter whenever they occur. For example, one can differentiate the evolution equation  $u_t = F(x, u^{(n)})$ —in which  $u^{(n)}$  have derivatives with respect to  $x$  only—with respect to  $t$  and  $x$  sufficient number of times (this is what is meant by “prolongation” of the system of equations) and solve for all derivatives of  $u$  involving time. Then, in the conservation law, substitute for any such derivatives to obtain a conservation law involving only  $x$  derivatives of  $u$ .

**30.3.3. Example.** The current density  $\mathbf{J}_1 = (\frac{1}{2}u_t^2 + \frac{1}{2}u_x^2, -u_t u_x)$  is easily seen to be conserved for the system of first-order DEs

$$u_t = v_x, \quad u_x = v_t.$$

By eliminating all the time derivatives in  $\mathbf{J}_1$ , we obtain  $\mathbf{J}_2 = (\frac{1}{2}u_x^2 + \frac{1}{2}v_x^2, -u_x v_x)$ , which is also conserved. However, the difference between these two currents,

$$\mathbf{J} = \mathbf{J}_1 - \mathbf{J}_2 = (\frac{1}{2}u_t^2 - \frac{1}{2}v_x^2, u_x v_x - u_t u_x),$$

satisfies a trivial conservation law of the first kind, because the components of  $\mathbf{J}$  vanish on the solutions of the system. ■

trivial conservation law of the second kind; null divergence

**30.3.4. Definition.** If the current density  $\mathbf{J}$  satisfies  $\mathbf{D} \cdot \mathbf{J} = 0$  for all functions  $u = f(x)$ , even if they are not solutions of the system of DEs, the divergence identity is called a **trivial conservation law of the second kind**. In this case  $\mathbf{J}$  is called a **null divergence**.

If we treat  $J_j$  as the components of a  $(p - 1)$ -form  $\omega$ , so that the exterior derivative  $d\omega$  is the divergence of  $\mathbf{J}$  (times a volume element), then the triviality of the conservation law for  $\mathbf{J}$  is equivalent to the fact that  $\omega$  is closed. By the converse of the Poincaré lemma, there must be a  $(p - 2)$ -form  $\eta$  such that  $\omega = d\eta$ . In the context of this chapter, we have the following theorem.

**30.3.5. Theorem.** Suppose  $\mathbf{J} = (J_1(x, u^{(n)}), \dots, J_p(x, u^{(n)}))$  is a  $p$ -tuple of functions on  $X \times U^{(n)}$ . Then  $\mathbf{J}$  is a null divergence if and only if there exist smooth functions  $A_{kj}(x, u^{(n)})$ ,  $j, k = 1, \dots, p$ , antisymmetric in their indices, such that

$$J_k = \sum_{j=1}^p D_j A_{kj}, \quad j = 1, \dots, p. \tag{30.24}$$

trivial conservation law and the equivalence of two conservation laws

**30.3.6. Definition.** We say that  $\mathbf{D} \cdot \mathbf{J} = 0$  is a **trivial conservation law** if there exist antisymmetric smooth functions  $A_{kj}(x, u^{(n)})$  satisfying Equation (30.24) for all solutions of the system of DEs  $\Delta(x, u^{(n)}) = 0$ . Two conservation laws are **equivalent** if they differ by a trivial conservation law.

We shall not distinguish between conservation laws that are equivalent. It turns out that to within this equivalence, some systems of DEs have current densities  $\mathbf{J}$

such that

$$\mathbf{D} \cdot \mathbf{J} = \sum_{\nu=1}^l Q_{\nu} \Delta_{\nu} \quad \text{for some } l\text{-tuple } \mathbf{Q} = (Q_1, \dots, Q_l), \quad (30.25)$$

where  $\{Q_{\nu}\}$  are smooth functions of  $x, u$ , and all derivatives of  $u$ .

characteristic form of  
a conservation law

**30.3.7. Definition.** Equation (30.25) is called the **characteristic form** of the conservation law for the current density  $\mathbf{J}$ , and the  $l$ -tuple  $\mathbf{Q}$ , the **characteristic** of the conservation law.

We are now in a position to prove the celebrated Noether's theorem. However, we first need a lemma.

**30.3.8. Lemma.** Let  $\mathbf{v} = \sum_{i=1}^p X^i \partial/\partial x^i + \sum_{\alpha=1}^q U^{\alpha} \partial/\partial u^{\alpha}$  where  $X^i$  and  $U^{\alpha}$  are functions of  $x$  and  $u$ . Let

$$Q^{\alpha}(x, u^{(1)}) \equiv U^{\alpha}(x, u) - \sum_{i=1}^p X^i(x, u) u_i^{\alpha}, \quad \alpha = 1, \dots, q.$$

Then

$$\text{pr}^{(n)} \mathbf{v} = \text{pr}^{(n)} \mathbf{v}_Q + \sum_{i=1}^p X^i D_i, \quad (30.26)$$

where

$$\mathbf{v}_Q \equiv \sum_{\alpha=1}^q Q^{\alpha}(x, u^{(1)}) \frac{\partial}{\partial u^{\alpha}}, \quad \text{pr}^{(n)} \mathbf{v}_Q \equiv \sum_{\alpha=1}^q \sum_J D_J Q^{\alpha} \frac{\partial}{\partial u_J^{\alpha}}.$$

The sum over  $J$  extends over all multi-indices with  $0 \leq |J| \leq n$ , with the  $|J| = 0$  term being simply  $\mathbf{v}_Q$ .

*Proof.* Substitute  $Q^{\alpha}$  in the definition of  $U_J^{\alpha}$  as given in Theorem 29.3.5 to obtain

$$U_J^{\alpha} = D_J Q^{\alpha} + \sum_{i=1}^p X^i u_{J,i}^{\alpha},$$

where  $U_0^{\alpha} = Q^{\alpha} + \sum_{i=1}^p X^i u_i^{\alpha} = U^{\alpha}$ . It follows that (with  $J = 0$  included in the sum)

$$\begin{aligned} \text{pr}^{(n)} \mathbf{v} &= \sum_{i=1}^p X^i \frac{\partial}{\partial x^i} + \sum_J \left[ \sum_{\alpha=1}^q D_J Q^{\alpha} + \sum_{i=1}^p X^i u_{J,i}^{\alpha} \right] \frac{\partial}{\partial u_J^{\alpha}} \\ &= \sum_{\alpha=1}^q \sum_J D_J Q^{\alpha} \frac{\partial}{\partial u_J^{\alpha}} + \underbrace{\sum_{i=1}^p X^i \left[ \frac{\partial}{\partial x^i} + \sum_{\alpha=1}^q \sum_J u_{J,i}^{\alpha} \frac{\partial}{\partial u_J^{\alpha}} \right]}_{=D_i \text{ by Proposition 29.3.4}}, \end{aligned}$$

and the lemma is proved. □

The celebrated  
Noether's theorem  
connecting  
symmetries to  
conservation laws

**30.3.9. Theorem.** (Noether's theorem) *Let*

$$\mathbf{v} = \sum_{i=1}^p X^i \partial / \partial x^i + \sum_{\alpha=1}^q U^\alpha \partial / \partial u^\alpha$$

*be the infinitesimal generator of a local 1-parameter group of symmetries  $G$  of the variational problem  $\mathbf{L}[u] = \int L(x, u^{(n)}) d^p x$ . Let*

$$Q^\alpha(x, u^{(1)}) \equiv U^\alpha(x, u) - \sum_{i=1}^p X^i(x, u) u_i^\alpha, \quad u_i^\alpha = \frac{\partial u^\alpha}{\partial x^i}.$$

*Then there exists a  $p$ -tuple  $\mathbf{J} = (J_1, \dots, J_p)$  such that*

$$\mathbf{D} \cdot \mathbf{J} = \sum_{\alpha=1}^q Q^\alpha \mathbb{E}_\alpha(L) \quad (30.27)$$

*is a conservation law in characteristic form for the Euler-Lagrange equation  $\mathbb{E}_\alpha(L) = 0$ .*

*Proof.* We use Lemma 30.3.8 in the infinitesimal criterion of the variational symmetry (30.20) to obtain

$$\begin{aligned} 0 &= \text{pr}^{(n)} \mathbf{v}(L) + L \mathbf{D} \cdot \mathbf{X} \\ &= \text{pr}^{(n)} \mathbf{v}_Q(L) + \sum_{i=1}^p X^i D_i L + L \sum_{i=1}^p D_i X^i \\ &= \text{pr}^{(n)} \mathbf{v}_Q(L) + \sum_{i=1}^p D_i(LX^i) = \text{pr}^{(n)} \mathbf{v}_Q(L) + \mathbf{D} \cdot (L\mathbf{X}). \end{aligned} \quad (30.28)$$

Using the definition of  $\text{pr}^{(n)} \mathbf{v}_Q$  and the identity

$$(D_j S)T = D_j(ST) - S D_j T,$$

we can commute  $D_J = D_{j_1} \cdots D_{j_k}$  past  $Q^\alpha$  one factor at a time, each time introducing a divergence. Therefore,

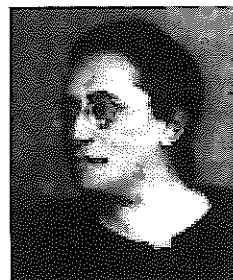
$$\begin{aligned} \text{pr}^{(n)} \mathbf{v}_Q(L) &= \sum_{\alpha, J} D_J Q^\alpha \frac{\partial L}{\partial u_J^\alpha} = \sum_{\alpha, J} Q^\alpha (-D)_J \frac{\partial L}{\partial u_J^\alpha} + \mathbf{D} \cdot \mathbf{A} \\ &= \sum_{\alpha=1}^q Q^\alpha \mathbb{E}_\alpha(L) + \mathbf{D} \cdot \mathbf{A}, \end{aligned}$$

where  $\mathbf{A} = (A_1, \dots, A_p)$  is some  $p$ -tuple of functions depending on  $L$ , the  $Q^\alpha$ 's, and their derivatives, whose precise form is not needed here. Combining this with Equation (30.28), we obtain

$$0 = \sum_{\alpha=1}^q Q^\alpha \mathbb{E}_\alpha(L) + \mathbf{D} \cdot (\mathbf{A} + L\mathbf{X}).$$

Selecting  $\mathbf{J} = -(\mathbf{A} + L\mathbf{X})$  proves the theorem.  $\square$

**Amalie Emmy Noether** (1882–1935), generally considered the greatest of all female mathematicians up to her time, was the eldest child of Max Noether, research mathematician and professor at the University of Erlangen, and Ida Amalia Kaufmann. Two of Emmy's three brothers were also scientists. Alfred, her junior by a year, earned a doctorate in chemistry at Erlangen. Fritz, two and a half years younger, became a distinguished physicist; and his son, Gottfried, became a mathematician.



At first Emmy Noether had planned to be a teacher of English and French. From 1900 to 1902 she studied mathematics and foreign languages at Erlangen. Then in 1903 she started her specialization in mathematics at the University of Göttingen. At both universities she was a nonmatriculated auditor at lectures, since at the turn of the century women could not be admitted as regular students. In 1904 she was permitted to matriculate at the University of Erlangen, which granted her the Ph.D., *summa cum laude*, in 1907. Her sponsor, the algebraist Gordan, strongly influenced her doctoral dissertation on algebraic invariants. Her divergence from Gordan's viewpoint and her progress in the direction of the "new" algebra first began when she was exposed to the ideas of Ernst Fischer, who came to Erlangen in 1911.

In 1915 Hilbert invited Emmy Noether to Göttingen. There she lectured at courses that were given under his name and applied her profound invariant-theoretic knowledge to the resolution of problems that he and Felix Klein were considering. Inspired by Hilbert and Klein's investigation into Einstein's general theory of relativity, Noether wrote her remarkable 1918 paper in which both the concept of *variational symmetry* and its connection with *conservation laws* were set down in complete generality.

Hilbert repeatedly tried to obtain her an appointment as Privatdozent, but the strong prejudice against women prevented her habilitation until 1919. In 1922 she was named a *nichtbeamteter ausserordentlicher Professor* ("unofficial associate professor"), a purely honorary position. Subsequently, a modest salary was provided through a *Lehrauftrag* ("teaching appointment") in algebra. Thus she taught at Göttingen (1922–1933), interrupted only by visiting professorships at Moscow (1928–1929) and at Frankfurt (summer of 1930).

In April 1933 she and other Jewish professors at Göttingen were summarily dismissed. In 1934 Nazi political pressures caused her brother Fritz to resign from his position at Breslau and to take up duties at the research institute in Tomsk, Siberia. Through the efforts of Hermann Weyl, Emmy Noether was offered a visiting professorship at Bryn Mawr College; she departed for the United States in October 1933. Thereafter she lectured and did research at Bryn Mawr and at the Institute for Advanced Studies, Princeton, but those activities were cut short by her sudden death from complications following surgery.

Emmy Noether's most important contributions to mathematics were in the area of abstract algebra. One of the traditional postulates of algebra, namely the commutative law of multiplication, was relinquished in the earliest example of a generalized algebraic structure, e.g., in Hamilton's quaternion algebra and also in many of the 1844 Grassmann algebras. From 1927 to 1929 Emmy Noether contributed notably to the theory of representations, the object of which is to provide realizations of noncommutative algebras by means of matrices, or linear transformations. From 1932 to 1934 she was able to probe profoundly

into the structure of noncommutative algebras by means of her concept of the *verschränktes* (“cross”) product.

Emmy Noether wrote some forty-five research papers and was an inspiration to many future mathematicians. The so-called Noether school included such algebraists as Hasse and W. Schmeidler, with whom she exchanged ideas and whom she converted to her own special point of view. She was particularly influential in the work of B. L. van der Waerden, who continued to promote her ideas after her death and to indicate the many concepts for which he was indebted to her.

### 30.4 Application to Classical Field Theory

It is clear from the proof of Noether’s theorem that if we are interested in the conserved current, we need to find  $A$ . In general, the expression for  $A$  is very complicated. However, if the variational problem is of first order (which in most cases of physical interest it is), then we can easily find the explicit form of  $A$ , and, consequently the conserved current  $J$ . We leave it for the reader to prove the following:

**30.4.1. Corollary.** *Let  $v = \sum_{i=1}^p X^i \partial/\partial x^i + \sum_{\alpha=1}^q U^\alpha \partial/\partial u^\alpha$  be the infinitesimal generator of a local 1-parameter group of symmetries  $G$  of the first-order variational problem  $L[u] = \int L(x, u^{(1)}) d^p x$ . Then<sup>5</sup>*

$$J_i = \sum_{\alpha=1}^q \sum_{j=1}^p X^j u_j^\alpha \frac{\partial L}{\partial u_i^\alpha} - \sum_{\alpha=1}^q U^\alpha \frac{\partial L}{\partial u_i^\alpha} - X^i L, \quad i = 1, \dots, p$$

form the components of a conserved current for the Euler–Lagrange equation  $\mathbb{E}_\alpha(L) = 0$ .

This corollary can be applied to most DEs in physics derivable from a Lagrangian. We are interested in partial DEs studied in classical field theories. The case of ODEs, studied in point mechanics, is relegated to Problem (30.11).

First consider spacetime translation  $v^i = \eta^{ij} \partial_j$ , where we have introduced the Lorentz metric  $\eta^{ij}$  to include non-Euclidean cases. In order for  $v^i$  to be an infinitesimal variational symmetry, it has to satisfy Equation (30.20), which in the case at hand, reduces to  $v^i(L) = 0$ , or  $\partial_i L = 0$ .

**30.4.2. Box.** *In order for a variational problem to be invariant under spacetime translations, its Lagrangian must not depend explicitly on the coordinates.*

<sup>5</sup>We have multiplied  $J_i$  by a negative sign to conform to physicists’ convention.

If spacetime translation happens to be a symmetry, then  $X^i \rightarrow \eta^{ij}$ , and the (double-indexed) conserved current, derived from Corollary 30.4.1, takes the form

$$T^{ij} = \sum_{\alpha=1}^q \frac{\partial u^\alpha}{\partial x_j} \frac{\partial L}{\partial u_i^\alpha} - \eta^{ij} L.$$

energy momentum  
current density

Using Greek indices to describe space-time coordinates, and Latin indices to label the components of  $\mathbb{R}^q$ , we write

$$T^{\mu\nu} = \sum_{j=1}^q \frac{\partial \phi^j}{\partial x_\mu} \frac{\partial L}{\partial \phi_\nu^j} - \eta^{\mu\nu} L \equiv \sum_{j=1}^q \eta^{\mu\sigma} \frac{\partial \phi^j}{\partial x^\sigma} \frac{\partial L}{\partial \phi_\nu^j} - \eta^{\mu\nu} L, \quad (30.29)$$

where we changed the dependent variable  $u$  to  $\phi$  to adhere to the notation used in the physics literature. Recall that  $\phi_\nu^j \equiv \partial \phi^j / \partial x^\nu$ .  $T^{\mu\nu}$  is called the **energy momentum current density**.

The quantity  $T^{\mu\nu}$ , having a vanishing divergence, is really a density, just as the continuity equation (vanishing of the divergence) for the electric charge involves the electric charge and current densities. In the electric case, we find the charge by integrating the charge density, the zeroth component of the electric 4-current density. Similarly, we find the “charge” associated with  $T^{\mu\nu}$  by integrating its zeroth component. This yields the energy momentum 4 vector:

$$P^\nu = \int_V T^{0\nu} d^3x.$$

We note that

$$\begin{aligned} \frac{dP^\nu}{dt} &= \frac{d}{dt} \int_V T^{0\nu} d^3x = \int_V \frac{\partial T^{0\nu}}{\partial t} d^3x \\ &= - \int_V \sum_{i=1}^3 \frac{\partial T^{i\nu}}{\partial x^i} d^3x = - \int_S \sum_{i=1}^3 T^{i\nu} da_i, \end{aligned}$$

where we have used the three-dimensional divergence theorem. By taking  $S$  to be infinite, and assuming that  $T^{i\nu} \rightarrow 0$  at infinity (faster than the element of area  $da_i$  diverges), we obtain  $dP^\nu/dt = 0$ , the conservation of the 4-momentum.

**30.4.3. Example.** A relativistic scalar field of mass  $m$  is a 1-component field satisfying the Klein–Gordon equation, which is, as the reader may check, the Euler–Lagrange equation of

$$\mathbf{L}[\phi] = \int L(\phi, \phi_\mu) d^4x \equiv \int \frac{1}{2} [\eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - m^2 \phi^2] d^4x.$$

The energy momentum current for the scalar field is found to be

$$T^{\mu\nu} = \partial^\mu \phi \partial^\nu \phi - \eta^{\mu\nu} L, \quad \partial^\mu \phi \equiv \eta^{\mu\nu} \frac{\partial \phi}{\partial x^\nu}.$$



Note that  $T^{\mu\nu}$  is symmetric under interchange of its indices. This is a desired feature of the energy momentum current that holds for the scalar field but is not satisfied in general, as Equation (30.29) indicates. The reader is urged to show directly that  $\partial_\mu T^{\mu\nu} = 0 = \partial_\nu T^{\mu\nu}$ , i.e., that energy momentum is conserved. ■

To go beyond translation, we consider classical (nonquantized) fields<sup>6</sup>  $\{\phi^j\}_{j=1}^{n_\alpha}$ , which, as is the case in most physical situations, transform among themselves as the rows of the  $\alpha$ th irreducible representation of a Lie group  $G$  that acts on the independent variables. Under these circumstances, the generators of the symmetry are given by Equation (27.59):

$$\mathcal{D}_{ij}(\xi) = \mathfrak{T}_{ij}^{(\alpha)}(\xi)\phi^k(\mathbf{x})\frac{\partial}{\partial\phi^k} + \delta_{ij}X^\nu(\mathbf{x};\xi)\frac{\partial}{\partial x^\nu}, \tag{30.30}$$

where  $\nu$  labels the independent variables. Corollary 30.4.1 now gives the conserved current as

$$J_{ij}^\mu = \left\{ X^\mu(\mathbf{x};\xi)\phi_\nu^k(\mathbf{x})\frac{\partial L}{\partial\phi_\nu^k} - X^\mu(\mathbf{x};\xi)L \right\} \delta_{ij} - \phi^k(\mathbf{x})\frac{\partial L}{\partial\phi_\mu^k}\mathfrak{T}_{ij}^{(\alpha)}(\xi),$$

where summation over repeated indices is understood with  $1 \leq k \leq n_\alpha$  and  $1 \leq \nu \leq p$ . We can rewrite this equation in the form

$$\mathbf{J}^\mu = \left\{ X^\mu(\mathbf{x};\xi)\phi_\nu^k(\mathbf{x})\frac{\partial L}{\partial\phi_\nu^k} - X^\mu(\mathbf{x};\xi)L \right\} \mathbf{1} - \phi^k(\mathbf{x})\frac{\partial L}{\partial\phi_\mu^k}\mathfrak{T}^{(\alpha)}(\xi), \tag{30.31}$$

where  $\mathbf{J}^\mu$  and  $\mathfrak{T}^{(\alpha)}(\xi)$  are  $n_\alpha \times n_\alpha$  matrices whose elements are  $J_{ij}^\mu$  and  $\mathfrak{T}_{ij}^{(\alpha)}(\xi)$ , respectively, and  $\mathbf{1}$  is the unit matrix of the same dimension.

We note that the conserved current has a coordinate part (the term that includes  $X^\mu$  and multiplies the unit matrix), and an “intrinsic” part (the term with no  $X^\mu$ ) represented by the term involving  $\mathfrak{T}^{(\alpha)}(\xi)$ . If the field has only one component (a scalar field), then  $\mathfrak{T}^{(\alpha)}(\xi) = 0$ , and only the coordinate part contributes to the current.

The current  $\mathbf{J}^\mu$  acquires an extra index when a *component* of  $\xi$  is chosen. As a concrete example, consider the case where  $G$  is the rotation group in  $\mathbb{R}^p$ . Then a typical component of  $\xi$  will be  $\xi^{\rho\sigma}$ , corresponding to a rotation in the  $\rho\sigma$ -plane, and the current will be written as  $\mathbf{J}^{\mu;\rho\sigma}$ . These extra indices are also reflected in  $X^\mu$ , as that too is a function of  $\xi$ :

$$X^\mu(\mathbf{x};\xi^{\rho\sigma})\frac{\partial}{\partial x^\mu} = x^\rho\partial_\sigma - x^\sigma\partial_\rho \Rightarrow X^\mu(\mathbf{x};\xi^{\rho\sigma}) = x^\rho\delta^{\mu\sigma} - x^\sigma\delta^{\mu\rho}.$$

orbital angular momentum and intrinsic spin

The volume integral of  $\mathbf{J}^{0;\rho\sigma}$  will give the components of angular momentum. When integrated, the term multiplying  $\mathbf{1}$  becomes the **orbital angular momentum**, and the remaining term gives the **intrinsic spin**. The conservation of  $\mathbf{J}^{\mu;\rho\sigma}$  is the

<sup>6</sup>The reader notes that the superscript  $\alpha$ , which labeled components of the independent variable  $u$ , is now the label of the irreducible representation. The components of the dependent variable (now denoted by  $\phi$ ) are labeled by  $j$ .

statement of the conservation of *total* angular momentum. The label  $\alpha$  denotes various representations of the rotation group. If  $p = 3$ , then  $\alpha$  is simply the value of the spin. For example, the spin- $\frac{1}{2}$  representation corresponds to  $\alpha = \frac{1}{2}$ , and

$$\mathfrak{T}^{(1/2)}(\xi) = \frac{1}{2}(\sigma^1, \sigma^2, \sigma^3), \quad \text{or} \quad \mathfrak{T}^{(1/2)}(\xi^a) = \frac{1}{2}\sigma^a, \quad a = 1, 2, 3,$$

with  $a$  labeling the three different “directions” of rotation.<sup>7</sup> If the field is a scalar,  $\mathfrak{T}^{(\alpha)}(\xi) = 0$ , and the field has only an orbital angular momentum.

## 30.5 Problems

**30.1.** Show that the derivative of a linear map from one Hilbert space to another is the map itself.

**30.2.** Show that a complex function  $f : \mathbb{C} \supset \Omega \rightarrow \mathbb{C}$  considered as a map  $f : \mathbb{R}^2 \supset \Omega \rightarrow \mathbb{R}^2$  is differentiable iff it satisfies the Cauchy–Riemann conditions. Hint: Consider the Jacobian matrix of  $f$ , and note that a linear complex map  $\mathbf{T} : \mathbb{C} \rightarrow \mathbb{C}$  is necessarily of the form  $\mathbf{T}(z) = \lambda z$  for some constant  $\lambda \in \mathbb{C}$ .

**30.3.** Show that

$$\frac{\delta \mathbf{E}_{y,i}[u]}{\delta u}(\mathbf{x}) = -\partial_i \delta(\mathbf{x} - \mathbf{y}).$$

**30.4.** Show that the first functional derivative of  $\mathbf{L}[u] \equiv \int_{x_1}^{x_2} \sqrt{1 + u_x^2} dx$ , obtained using Equation (30.9), is  $\mathbf{E}(L)$ .

**30.5.** Show that for the proper time of special relativity

$$\frac{\delta \mathbf{L}[x]}{\delta x(s)} = \frac{x_{ss}}{(1 - x_s^2)^{3/2}}.$$

Use this to show that the contribution of the second variational derivative to the Taylor expansion of the functional is always negative.

**30.6.** Show that the first prolongation of the Lorentz generator  $\mathbf{v} = u \partial_x + x \partial_u$  is

$$\text{pr}^{(1)} \mathbf{v} = \mathbf{v} + (1 - u_x^2) \frac{\partial}{\partial u_x}.$$

**30.7.** Verify that rotation in the  $xu$ -plane is a symmetry of the arc-length variational problem (see Example 30.1.13).

**30.8.** Show that  $\mathbf{v}_4$ ,  $\mathbf{v}_6$ , and  $\mathbf{v}_7$  of Table 29.3 are variational symmetries of Equation (30.21), but  $\mathbf{v}_5$ ,  $\mathbf{v}_8$ ,  $\mathbf{v}_9$ , and  $\mathbf{v}_{10}$  are not. Find the constant  $c$  (if it exists) such that  $\mathbf{v}_5 + cu \partial_u$  is a variational symmetry. Show that no linear combination of inversions produces a symmetry.

Kepler problem **30.9.** The two-dimensional **Kepler problem** (for a unit point mass) starts with the functional

$$L = \int \left[ \frac{1}{2}(\dot{x}_i^2 + \dot{y}_i^2) - V(r) \right] dt, \quad r = \sqrt{x^2 + y^2}.$$

- Show that  $L$  is invariant under  $t$  translation and rotation in the  $xy$ -plane.
- Find the generators of  $t$  translation and rotation in polar coordinates and conclude that  $r$  is the best choice for the *independent* variable.
- Rewrite  $L$  in polar coordinates and show that it is independent of  $t$  and  $\theta$ .
- Write the Euler–Lagrange equations and integrate them to get  $\theta$  as an integral over  $r$ .

**30.10.** Prove Corollary 30.4.1.

**30.11.** Consider a system of  $N$  particles whose total kinetic energy  $K$  and potential energy  $U$  are given by

$$K(\dot{\mathbf{x}}) = \frac{1}{2} \sum_{\alpha=1}^N m_{\alpha} |\dot{\mathbf{x}}^{\alpha}|^2, \quad U(t, \mathbf{x}) = \sum_{\alpha \neq \beta} k_{\alpha\beta} |\mathbf{x}^{\alpha} - \mathbf{x}^{\beta}|^{-1},$$

where  $\mathbf{x}^{\alpha} = (x^{\alpha}, y^{\alpha}, z^{\alpha})$  is the position of the  $\alpha$ th particle. The variational problem is of the form

$$L[\mathbf{x}] = \int_{-\infty}^{\infty} L(t, \mathbf{x}, \dot{\mathbf{x}}) dt = \int_{-\infty}^{\infty} [K(\dot{\mathbf{x}}) - U(t, \mathbf{x})] dt.$$

- Show that the Euler–Lagrange equations are identical to Newton’s second law of motion.
- Write the infinitesimal criterion for the vector field

$$\mathbf{v} = \tau(t, \mathbf{x}) \frac{\partial}{\partial t} + \sum_{\alpha} \left[ \xi^{\alpha}(t, \mathbf{x}) \frac{\partial}{\partial x^{\alpha}} + \eta^{\alpha}(t, \mathbf{x}) \frac{\partial}{\partial y^{\alpha}} + \zeta^{\alpha}(t, \mathbf{x}) \frac{\partial}{\partial z^{\alpha}} \right]$$

to be the generator of a 1-parameter group of variational symmetries of  $L$ .

- Show that the conserved “current” derived from Corollary 30.4.1 is

$$T = \sum_{\alpha=1}^N m_{\alpha} (\xi^{\alpha} \dot{x}^{\alpha} + \eta^{\alpha} \dot{y}^{\alpha} + \zeta^{\alpha} \dot{z}^{\alpha}) - \tau E,$$

where  $E = K + U$  is the total energy of the system.

- Find the conditions on  $U$  such that (i) time translation, (ii) space translations, and (iii) rotations become symmetries of  $L$ . In each case, compute the corresponding conserved quantity.

<sup>7</sup>Only in three dimensions can one label rotations with a single index. This is because each coordinate plane has a unique direction (by the use of the right-hand rule) perpendicular to it that can be identified as the direction of rotation.

30.12. Show that the Euler–Lagrange equation of

$$\mathcal{L}[\phi] = \int L(\phi, \phi_{,\mu}) d^4x \equiv \int \frac{1}{2}[\eta^{\mu\nu} \partial_{\mu}\phi \partial_{\nu}\phi - m^2\phi^2] d^4x$$

is the Klein–Gordon equation. Verify that  $T^{\mu\nu} = \partial^{\mu}\phi \partial^{\nu}\phi - \eta^{\mu\nu}L$  are the currents associated with the invariance under translations. Show directly that  $T^{\mu\nu}$  is conserved.

## Additional Reading

The first three books cited below discuss what we have not covered in our study of Lie groups and DEs, namely “generalized symmetries,” whereby the generators of symmetries are allowed to be not only functions of the independent and dependent variables, but also of the derivatives of the dependent variables. This leads to a more general version of Noether’s theorem than presented in this chapter.

1. Bluman, G. and Kumei, S. *Symmetries and Differential Equations*, Springer-Verlag, 1989.
2. Olver, P. *Application of Lie Groups to Differential Equations*, Springer-Verlag, 1986.
3. Stephani, H. *Differential Equations: Their solutions using symmetries*, Cambridge University Press, 1989.
4. All books on relativistic quantum field theories have a discussion of symmetries and conservation laws. See, for example, Weinberg, S. *The Quantum Theory of Fields* (2 volumes), Cambridge University Press, 1995.

---

# Bibliography

- [Abra 85] Abraham, R. and Marsden, J. *Foundations of Mechanics*, 2nd ed., Addison-Wesley, 1985.
- [Abra 88] Abraham, R., Marsden, J., and Ratiu, T. *Manifolds, Tensor Analysis, and Applications*, 2nd ed., Springer-Verlag, 1988.
- [Axle 96] Axler, S. *Linear Algebra Done Right*, Springer-Verlag, 1996.
- [Baru 86] Barut, A. and Raczka, R. *Theory of Group Representations and Applications*, World Scientific, 1986.
- [Birk 77] Birkhoff, G. and MacLane, S. *Modern Algebra*, 4th ed., Macmillan, 1977.
- [Birk 78] Birkhoff, G. and Rota, G.-C. *Ordinary Differential Equations*, 3rd ed., Wiley, 1978.
- [Bish 80] Bishop, R. and Goldberg, S. *Tensor Analysis on Manifolds*, Dover, 1980.
- [Blum 89] Bluman, G. and Kumei, S. *Symmetries and Differential Equations*, Springer-Verlag, 1989.
- [Bocc 90] Boccara, N. *Functional Analysis*, Academic Press, 1990.
- [Boer 63] Boerner, H. *Representation of Groups*, North-Holland, 1963.
- [Bott 82] Bott, R. and Tu, L. *Differential Forms in Algebraic Topology*, Springer-Verlag, 1982.
- [Chev 46] Chevalley, C. *Theory of Lie Groups*, Princeton University Press, 1946.
- [Choq 82] Choquet-Bruhat, Y., DeWitt-Morette, C., and Dillard-Bleick, M. *Analysis, Manifolds, and Physics*, 2nd ed., North-Holland, 1982.
- [Chur 74] Churchill, R. and Verhey, R. *Complex Variables and Applications*, 3rd ed., McGraw-Hill, 1974.
- [Cour 62] Courant, R. and Hilbert, D. *Methods of Mathematical Physics*, vol. 1, Interscience, 1962.

- [Denn 67] Dennerly, P., and Krzywicki, A. *Mathematics for Physicists*, Harper and Row, 1967.
- [DeVi 90] DeVito, C. *Functional Analysis and Linear Operator Theory*, Addison-Wesley, 1990.
- [DeVr 94] DeVries, P. *A First Course in Computational Physics*, Wiley, 1994.
- [Econ 83] Economou, E. *Green's Functions in Quantum Physics*, Springer-Verlag, 1983.
- [Elli 79] Elliott, J., and Dawber, P. *Symmetry in Physics* (2 volumes), Oxford University Press, 1979.
- [Flan 89] Flanders, H. *Differential Forms with Applications to Physical Sciences*, Dover, 1989.
- [Foll 95] Folland, G. *Introduction to Partial Differential Equations*, 2nd ed., Princeton University Press, 1995.
- [Frie 82] Friedman, A. *Foundations of Modern Analysis*, Dover, 1982.
- [Fult 91] Fulton, W. and Harris, J. *Representation Theory*, Springer-Verlag, 1991.
- [Gill 72] Gillispie, C., ed. *Dictionary of Scientific Biography*, Charles Scribner's, New York, 1970.
- [Gilm 74] Gilmore, R. *Lie Groups, Lie Algebras, and Some of Their Applications*, Wiley, 1974.
- [Glim 87] Glimm, J. and Jaffe, A. *Quantum Physics*, 2nd ed., Springer-Verlag, 1987.
- [Grad 65] Gradshteyn, I. and Ryzhik, I. *Table of Integrals, Series, and Products*, Academic Press, 1965.
- [Greu 75] Greub, W. *Linear Algebra*, 4th ed., Springer-Verlag, 1975.
- [Halm 58] Halmos, P. *Finite Dimensional Vector Spaces*, 2nd ed., Van Nostrand, 1958.
- [Halm 74] Halmos, P. *Naive Set Theory*, Springer-Verlag, 1974.
- [Hame 89] Hamermesh, M. *Group Theory and Its Application to Physical Problems*, Dover, 1989.
- [Hass 00] Hassani, S. *Mathematical Methods for Students of Physics and Related Fields*, Springer-Verlag, 2000.
- [Hell 67] Hellwig, G. *Differential Operators of Mathematical Physics*, Addison-Wesley, 1967.
- [Hild 87] Hildebrand, F. *Introduction to Numerical Analysis*, 2nd ed., Dover, 1987.
- [Hill 87] Hildebrand, F. *Statistical Mechanics*, Dover, 1987.
- [Jack 75] Jackson, J. *Classical Electrodynamics*, 2nd ed., Wiley, 1975.
- [Jorg 82] Jörgen, K. *Linear Integral Operators*, Pitman, 1982.
- [Kell 85] Kelley, J. *General Topology*, Springer-Verlag, 1985.
- [Koba 63] Kobayashi, S. and Nomizu, K. *Foundations of Differential Geometry*, vol. I, Wiley, 1963.
- [Lang 85] Lang, S. *Complex Analysis*, 2nd ed., Springer-Verlag, 1985.

- [Lorr 88] Lorrain, P., Corson D., and Lorrain, F. *Electromagnetic Fields and Waves*, 3rd ed., W. H. Freeman, 1988.
- [Mack 68] Mackey, G. *Induced Representations*, Benjamin, 1968.
- [Mari 80] Marion, J. and Heald, M. *Classical Electromagnetic Radiation*, 2nd ed., Academic Press, 1980.
- [Math 70] Mathews, J. and Walker, R. *Mathematical Methods of Physics*, 2nd ed., Benjamin, 1970.
- [Mess 66] Messiah, A. *Quantum Mechanics* (2 volumes), Wiley, 1966.
- [Mill 68] Miller, W. *Lie Theory and Special Functions*, Academic Press, 1968.
- [Misn 73] Misner, C., Thorne, K., and Wheeler, J. *Gravitation*, Freeman, 1973.
- [Mors 53] Morse, P. and Feshbach, M. *Methods of Theoretical Physics*, McGraw-Hill, 1953.
- [Naka 90] Nakahara, M. *Geometry, Topology, and Physics*, Adam Hilger, 1990.
- [Olve 86] Olver, P. *Application of Lie Groups to Differential Equations*, Springer-Verlag, 1986.
- [Reed 80] Reed, M. and Simon, B. *Functional Analysis* (4 volumes), Academic Press, 1980.
- [Rich 78] Richtmyer, R. *Principles of Advanced Mathematical Physics*, Springer-Verlag, 1978.
- [Roac 70] Roach, G. *Green's Functions*, Van Nostrand, 1970.
- [Rotm 84] Rotman, J. *An Introduction to the Theory of Groups*, 3rd ed., Allyn and Bacon, 1984.
- [Rudi 74] Rudin, W. *Functional Analysis*, McGraw-Hill, 1991.
- [Saun 89] Saunders, D. *The Geometry of Jet Bundles*, Cambridge University Press, 1989.
- [Simm 83] Simmons, G. *Introduction to Topology and Modern Analysis*, Krieger, 1983.
- [Simm 92] Simmons, G. *Calculus Gems*, McGraw-Hill, 1992.
- [Stak 79] Stakgold, I. *Green's Functions and Boundary Value Problems*, Wiley, 1979.
- [Step 89] Stephani, H. *Differential Equations: Their solutions using symmetries*, Cambridge University Press, 1989.
- [Tric 55] Tricomi, F. *Vorlesungen über Orthogonalreihen*, Springer, 1955.
- [Vara 84] Varadarajan, V. *Lie Groups, Lie Algebras and Their Representations*, Springer-Verlag, 1984.
- [Wald 84] Wald, R. *General Relativity*, University of Chicago Press, 1984.
- [Warn 83] Warner, F. *Foundations of Differentiable Manifolds and Lie Groups*, Springer-Verlag, 1983.
- [Wats 52] Watson, G. *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, 1952.
- [Wein 95] Weinberg, S. *The Quantum Theory of Fields* (2 volumes), Cambridge University Press, 1995.
- [Wign 59] Wigner, E. *Group Theory*, Academic Press, 1959.
- [Zeid 95] Zeidler, E. *Applied Functional Analysis*, Springer-Verlag, 1995.





---

# Index

- Abel, 63, 177, 183, 233, 416, 462, 653
  - biography, 662
- abelian group, 654
- abelian Lie algebra, 834
- absolute convergence, 252
- additive identity, 20
- adjoint
  - formal, 560
- adjoint action, 827
- adjoint algebra, 840
- adjoint BC, 563
- adjoint DO, 364–367
- adjoint Green's function, 566
- adjoint map, 825
- adjoint of a matrix, 87
- adjoint of an operator, 61
- adjoint representation, 679
  - character, 684
- affine group, 876
- algebra, 41–45
  - definition, 41
  - derivation of an, 42
  - dimension of an, 41
  - homomorphism, 43
  - ideal, 44
  - isomorphism, 43
  - of operators, 49–76
  - quaternions, 44
  - structure constants, 43
  - symmetric, 737
  - tensor, 731
- algebraic equation, 937
  - symmetry group of, 937
- analytic continuation, 302–309
- analytic function, 228–235
  - definition, 233
  - derivative, 228
  - derivatives as integrals, 249
  - entire, 233
  - poles of, 273
  - roots (zeros) of, 262
- angular momentum, 831, 869
  - addition theorem, 866
  - eigenvalues, 337
  - operator, 331
    - construction of eigenvalues, 334
- annihilator, 39
- anti-hermitian operator, 63
- anticommutator, 59
- antisymmetric representation, 679
- antisymmetrizer, 739
- arc length, 890
- associated Legendre functions, 341
- associative algebra, 41
- atlas, 765
- automorphism, 36
- averaging operator, 71
- azimuthal symmetry, 343

- backward difference operator, 70, 384–387
- Baker–Campbell–Hausdorff, 59
- Banach space, 148
- basis, 22
  - dual of a, 38, 729
  - oriented, 745
  - orthonormal, 26
  - transformation matrix, 91
- Becquerel, 799
- Bernoulli, 423, 981
- Bessel, 177, 620, 982
  - biography, 422
- Bessel equation
  - Liouville substitution, 514
- Bessel function, 423–426
  - asymptotic behavior, 443
    - large argument, 444
    - large order, 443
  - confluent hypergeometric, 424
  - generating function, 440
  - integral representation of, 438
  - modified, 425
    - second kind, 425
  - oscillation of, 363
  - second kind, 424
  - third kind, 425
- Bessel functions
  - spherical, 542
- Bessel imaginary (bei), 539
- Bessel inequality, 149
- Bessel real (ber), 539
- beta function, 309–312
  - definition, 310
- Bianchi identity, 889
- bijjective map, 6
- binary operation, 7
- binomial theorem, 13
- Birkhoff's theorem, 925
- Bolzano, 11
- Bolzano-Weierstrass theorem, 461
- Boole, 704
- boundary conditions
  - Dirichlet, 591
  - homogeneous, 559
  - mixed, 564
  - Neumann, 591
  - periodic, 515
  - separated, 510
  - unmixed, 564
- boundary functionals, 559
- boundary point, 459
- boundary value problem, 560, 583
  - Dirichlet, 591
  - Neumann, 591
  - numerical solution, 383
- bounded operator, 453–457
  - continuity, 454
- branch cut, 296
- Brouwer, 850
- bundle of tensors, 786
- BWHB theorem, 853
- canonical basis, 747, 802
  - second order linear DE, 404
- canonical coordinates, 802
- canonical transformation, 802, 806
- Cantor, 463, 739, 799
  - biography, 10
- Cantor set, 12
- cardinality, 12
- Cartan, 798, 841, 942
  - biography, 743
- Cartan metric tensor, 843, 862
- Cartan subalgebra, 845
- Cartan's lemma, 743
- Cartesian product, 2
- Casimir operator, 861–863, 868–870, 872, 873
- Cauchy, 96, 271, 296, 416, 478, 528, 588, 653, 891
  - biography, 233
- Cauchy data, 584
- Cauchy integral formula, 245
  - for operators, 482
- Cauchy problem, 584
  - ill-posed, 590
- Cauchy sequence, 10, 146
- Cauchy–Goursat theorem, 242
- Cauchy–Riemann conditions, 230
  - differentiability, 232
- Cayley, 704, 744
- center of a group, 658
- central difference operator, 70

- centralizer, 658  
 Champollion, 198  
 character  
   adjoint representation, 684  
   compound, 683, 685  
   conjugacy class, 684  
   group and its subgroup, 692  
   of a representation, 683  
   simple, 683, 685  
 character table, 691  
 character table  
   for  $S_2$ , 694  
 character table  
   for  $S_3$ , 694  
 characteristic hypersurface, 585–589  
   definition, 586  
 characteristic polynomial, 115  
   linear DEs, 376  
 chart, 764  
 Chebyshev, 588  
 Chebyshev polynomials, 185  
 Chevalley's theorem, 862  
 Chevalley, biography, 864  
 Christoffel symbol, 890  
 Christoffel, biography, 890  
 circuit matrix of a DE, 403  
 circular heat-conducting plate, 537  
 Clebsch, 178, 704  
   biography, 702  
 Clebsch–Gordan  
   coefficients, 705, 706  
   decomposition, 701–705, 721, 866  
   series, 703, 706  
 Clifford, 744  
 closed form, 797  
 closed subset, 459  
 closure, 459  
 codomain, 5  
 cofactor, 95  
 commutative algebra, 41  
 commutative group, 654  
 commutative Lie algebra, 834  
 commutator  
   diagonalizability, 124  
   commutator subgroup, 658  
   commutators, 55  
   compact Lie algebra, 843  
   compact Lie group  
     representation, 845–855  
 compact operator, 464–466  
   definition, 464  
   spectrum, 467–473  
 compact resolvent, 508  
 compact set, 458–463  
 compact subset, 461  
 compact support, 165n, 801  
 comparison theorem, 361–363  
 complement of a set, 2  
 complete o.n. sequence, 149  
 completeness relation, 68, 150  
 complex coordinate space, 21  
 complex exponential function, 234  
 complex FOLDEs, 401–402  
 complex function, 227  
   integration, 241  
 complex  $GL(V)$ , 822  
 complex plane  
   contour in the, 242  
   curve in the, 242  
   multiply connected region, 245  
   path in the, 242  
   simply connected region, 245  
 complex potential, 236  
 complex series, 252  
 complex SOLDE, 404–410  
 complex vector space, 20  
 composition of maps, 6  
 compound character, 685  
 conducting cylindrical can, 535  
 confluent hypergeometric function  
   definition, 420  
   integral representation of, 437  
 conformal group, 912  
   in 2 dimensions, 912  
 conformal Killing vector, 911  
 conformal map, 236–241  
   definition, 237  
   translation, 238  
 conformal transformation, 911  
   special, 912  
 conic sections, 132  
 conjugacy class, 661, 684  
 conjugate, 661  
 conjugate subgroup, 657

- conjugation of operators, 61–63
- conjunct, 560, 563, 576
- connection coefficients, 886, 900
- conservation law, 855, 992
  - characteristic, 994
  - trivial, 993
  - trivial of the first kind, 992
  - trivial of the second kind, 993
- conserved current density, 992
- constant of the motion, 992
- constrained systems, 804n
- continuous index, 159
- contraction, 735
- contravariant degree, 731
- contravariant tensor, 731
- convex subset, 474
- convolution theorem, 223
- coordinate curve, 774
- coordinate frame, 774
- coordinate functions, 764
- coordinate representation of  $L_{g^*}$ , 820
- coordinate transformation
  - orientation preserving, 800
  - orientation reversing, 800
- coset, 658
- cosmological constant, 921
- Coulomb, 985
- countably infinite set, 12
- covariant degree, 731
- covariant derivative, 897–908
  - and curvature, 903
  - Lie derivative, 902
- covariant tensor, 731
- criterion for irreducibility, 685
- crystallography, 675
- current density, 992
  - energy momentum, 998
- curvature
  - and gravity, 914
  - as relative acceleration, 913
  - matrix, 888
  - two-form, 885
- curve
  - coordinate, 774
  - differentiable, 770
- curvilinear coordinates, 892
- cycle, 666
- cyclic permutation, 666
- cyclic subgroup, 657
- d'Alembert, 984
  - biography, 330
- Darboux, 744, 942
- Darboux inequality, 245
- Darboux theorem, 802
- de Broglie, 807
- Dedekind, 11, 711, 739, 891
- degeneracy, 697
  - energy, 605
- degenerate kernels, 501–504
- delta function, 160, 592
  - and step function, 163
  - expansion
    - Fourier, 204
    - general, 189
  - limit of sequence, 161, 163
  - potential, 601
- dense subset, 460
- density function, 833
- density of states, 532
- derivation, 840
- derivation algebra, 840
- derivation of an algebra, 42
- derivation property
  - tangent vector, 772
- derivative
  - covariant, 897
- derivative operator
  - unboundedness of, 455
- Descartes, 738
- determinant, 93–101, 744
  - analytic definition of, 137
  - connection with trace, 102
  - definition, 93
  - derivative of, 102
  - expansion of, 95
  - exponential of trace, 103
  - of an operator, 101
  - products of matrices, 97
  - relation to trace, 101
- diffeomorphism, 768
- differentiable curve, 770
  - tangent vector, 772
- differentiable manifold, 763–769

- dimension of a, 764
- differentiable map, 768
  - coordinate expression of, 768
- differential
  - of a constant map, 777
  - of a map, 776
  - one-form, 786
  - real-valued maps, 778
- differential equation
  - analytic, 401
  - analytic properties, 401
  - associated Legendre, 343
  - Bessel, 407
  - confluent hypergeometric, 420
  - constant coefficients, 376–379
  - Euler, 411
  - first-order
    - existence, 350
    - linear, 351
    - normal form, 351
    - Peano existence theorem, 351
    - uniqueness, 350
    - uniqueness theorem, 352
  - first-order linear
    - irregular singular point, 402
    - regular singular point, 402
  - Fuchsian, 411
  - higher-order
    - numerical solution, 393
  - homogeneous, 349
  - hypergeometric, 407
    - definition, 413
    - Jacobi functions, 418
    - Kummer's solutions, 418
  - inhomogeneous, 349
  - linear, 349
    - superposition principle, 354
  - numerical solution, 384
    - Adam's method, 386
    - Euler's method, 386
  - Riemann, 412
  - second order linear
    - behavior at infinity, 410
    - canonical basis, 404
    - characteristic exponents, 406
  - second-order linear
    - adjoint, 365
    - Frobenius method, 371
    - general properties, 352
    - integrating factor, 364
    - regular, 353
    - uniqueness theorem, 354
    - Wronskian, 356
  - symmetry group of, 941–950
- differential form, 791
  - Maxwell's equations, 793
- differential forms
  - Lorentz force law, 794
- differential operator, 349, 863
- differentiation operator, 73
- diffusion equation, 528, 591, 621
  - time-dependent, 530
- diffusion operator
  - Green's function for, 632
- dilation, 238
- dilatation, 912
- dimension theorem, 35
- Dirac, 850
  - biography, 167
- Dirac delta function, 160
- direct products, 661–663
- direct sum, 109–111
  - definition, 110
- directional derivative, 976
- Dirichlet, 177, 738, 888, 891
  - biography, 614
- Dirichlet boundary condition, 591
- Dirichlet BVP, 591, 613–619
  - in two dimensions, 638
- discrete Fourier transform, 217–219
- dispersion relation
  - with one subtraction, 308
- dispersion relations, 306–309
- distribution, 166
  - density, 166
  - derivative of, 169
  - Fourier transform of a, 219
  - limit of functions, 168
- divergence, 987
  - of tensors, 903
- domain, 5
- dual
  - basis, 38, 729
  - of an operator, 39

- space, 34
- eigenfunction expansion technique, 636–641
  - 2D Laplacian, 638
- eigenspace, 115
- eigenvalue, 114–116
  - characteristic polynomial, 115
  - definition, 114
  - of angular momentum, 337
  - simple, 115
- eigenvector, 114–116
  - angular momentum, 338–346
  - definition, 114
  - generalized, 468
- Einstein, 799, 849, 889, 891, 899, 920, 921, 996
- Einstein tensor, 919
- Einstein's equation, 918–921
  - Schwarzschild solution, 924
- Einstein's summation convention, 728
- electromagnetic field tensor, 796, 797
- elementary column operation, 99
- elementary row operation, 99
- elliptic PDE, 589, 613–621
- elsewhere, 838
- empty set, 2
- endomorphism, 33
- energy function, 805
- entire function, 233
- equivalence class, 3
- equivalence relation, 3
- equivalent representations, 674
- essential singularity, 273
- essentially idempotent, 689
- $\eta$ -orthogonal matrices, 837
- Euclid, 151, 807
- Euclidean metric, 892
- Euclidean space, 754
- Euler, 233, 390, 415, 423, 514, 888, 984
  - biography, 981
- Euler angles, 89, 865
- Euler kernel, 434
- Euler operator, 980
- Euler theorem, 865
- Euler transform, 433
- Euler–Lagrange equations, 981
- Euler–Mascheroni constant, 310
- evaluation function, 977
- event, 837
- evolution operator, 58, 626
- exact form, 797
- expectation value, 64
- exponential function
  - complex, 234
- exponential map, 824
- exterior algebra, 739–749
- exterior calculus, 791–801
- exterior derivative, 791
- exterior product, 740
- $F$ -related vector fields, 780
- faithful representation, 674
- Fermi energy, 532
- Feynman diagram, 602
- Feynman propagator, 636
- field, 20
- finite-difference operators, 70–73
- finite-rank operator, 464
- first integral, 992
- first variation, 983
- flat manifold, 895, 896
- flat map, 747, 802
- flow, 784
- form factor, 216
- formal adjoint, 560
- formally self-adjoint, 560
- forward difference operator, 70
- Fourier, 614, 653, 982
  - biography, 198
- Fourier series, 196–208
  - angular variable, 197
  - fundamental cell, 198
  - general variable, 200
  - group theory, 853
  - higher dimensions, 207
  - main theorem, 202
  - sawtooth, 201
  - square wave, 200
  - to Fourier transform, 208–210
- Fourier transform, 208–220, 433
  - and derivatives, 216–217
  - and GF, 628–636
  - and quark model, 216

- Coulomb potential
  - charge distribution, 215
  - point charge, 214
- definition, 210
- distribution, 219
- Gaussian, 212
- higher dimensions, 213
- Fourier–Bessel series, 536
- Fredholm, 151
- Fredholm alternative, 496
- Fredholm equation, 488, 600
- Fredholm, biography, 496
- Friedmann metric, 892
- Friedmann, biography, 921
- Frobenius, 681, 850, 874
  - biography, 711
- Frobenius method, 370–376
- Fuchsian DE, 410–413
  - definition, 411
- function, 5
- function algebra, 42
- function of operator, 53, 125–128
  - derivative, 56
- functional, 980
  - linear, 34
- functional derivative, 977
- functions of trig functions
  - integrals of, 281
- fundamental solution, 597–600
- fundamental theorem of algebra, 251
- future light cone, 837
  
- $G$ -invariance, 937
- $g$ -orthogonal, 751
- $g$ -orthonormal, 752
- Galois, 96, 711, 842, 942
  - biography, 653
- gamma function, 181, 309–312
  - definition, 309
- gauge invariance, 797
- Gauss, 96, 183, 233, 422, 463, 478, 614, 738, 797, 888, 891, 981
  - biography, 415
- Gay-Lussac, 528
- Gegenbauer functions, 419
- Gegenbauer polynomials, 185
- general linear group, 655
  - representation, 856–859
- general relativity, 918–932
- generalized eigenvector, 468
- generalized Fourier coefficients, 150
- generalized functions, 165–169
- generalized Green's identity, 561, 574, 597
- generating function, 189
- geodesic, 905–908
  - relative acceleration, 913
- geodesic deviation, 913–918
  - equation of, 913
- geodesic equation
  - massive particles, 926, 927
  - massless particles, 926, 928
- geometric multiplicity, 115
- Gibbs, 463, 807
- Gibbs phenomenon, 205–207
- $GL(n, \mathbb{R})$  as a Lie group, 816
- $GL(\mathcal{V})$ 
  - as a Lie group, 816
  - representation of, 856
- Gödel, 799
- Gordan, 996
  - biography, 704
- gradient
  - for Hilbert spaces, 976
- gradient operator, 901
- Gram, 28
- Gram–Schmidt process, 27–29
- Gram–Schmidt process, 172
- graph, 5
- Grassmann, 743, 996
- Grassmann product, 740
- gravitational red-shift, 930
- gravity
  - and curvature, 914
- Green's function, 289
  - adjoint, 566
  - advanced, 634
  - and Fourier transform, 628
  - construction, 569
  - diffusion operator, 632
  - Dirichlet BC
    - circle, 618
  - Dirichlet BVP, 613
  - eigenfunction expansion, 577–579

- for SOLDO, 565–577
- for the Laplacian, 595–596
- formal considerations, 557
- Helmholtz operator, 630
  - in 2D, 643
- in one dimension, 553
- indefinite, 554
- inhomogeneous BCs, 574
- integral equations, 600
- Laplacian, 629
- Neumann BVP, 620
  - exterior, 621
  - interior, 620
- physical interpretation, 577
- properties, 567
- retarded, 634
- second order DO, 562
- symmetry, 567
- wave equation, 633
- Green's function
  - for  $d/dx$ , 554
  - for  $d^2/dx^2$ , 555
- Green's identity, 567, 596, 623, 627
- Green, biography, 561
- group, 652–656
  - abelian, 654
  - affine, 876
  - algebra, 687
    - symmetric group, 718
  - automorphism, 655
  - center of  $a$ , 658
  - commutative, 654
  - commutator, 657
  - homomorphism, 655
  - isomorphism, 655
  - left action, 663
  - multiplication, 652
  - multiplication table, 656
  - of affine motions, 817
  - order of, 652
  - orthogonal, 656
  - realization, 664
  - representation
    - definition, 674
    - irreducible, projection operator, 698
    - matrix, 675
    - particles and fields, 699
    - tensor product, 699
  - right action, 663
  - rigid rotations, 657
  - special orthogonal, 657
  - special unitary, 657
  - symplectic, 657, 749
  - unitary, 656
- group action, 663–664, 817–820
  - effective, 663
  - infinitesimal generator, 827
  - transitive, 663, 818
- guided waves
  - TE, 533
  - TEM, 533
  - TM, 533
- Haar measure, 832
- Halley, 422
- Hamilton, 178, 490, 996
  - biography, 806
- Hamiltonian system, 805
- Hamiltonian vector field, 805
- Hankel function, 425
  - asymptotic expansion of, 316
- Hankel transform, 434
- harmonic functions, 236
- heat equation, 327, 591, 621
  - symmetry group, 957–961
- heat transfer, 528
  - steady-state, 528
  - time-dependent, 530
- heat-conducting plate, 528
- Hegel, 738
- Heisenberg, 64, 167
- helicity, 876
- Helmholtz, 178, 588, 849
- Helmholtz equation, 542
- Helmholtz operator
  - Green's function for, 630
- Hermite, 183, 798
  - biography, 63
- Hermite polynomials, 176, 180, 517
- hermitian conjugate, 61
- hermitian kernel, 497–501
- hermitian operator, 63–67
- Hilbert, 11, 28, 199, 463, 704, 799, 848, 920, 996



- biography, 151
- Hilbert space, *145–157*
  - basis, 150
  - definition, 148
  - differential of functions, 976
  - functions on
    - derivative of, 974
- Hilbert transforms, 308
- Hilbert–Schmidt operator, 465
- Hilbert–Schmidt theorem, 498
- Hilbert–Schmidt kernel, 466, 494
- Hilbert–Schmidt operator, 847
- Hodge star operator, *756–758*, 795
- Hölder, 463
- homogeneous SOLDE
  - exact, 364
  - second solution, 358
- homographic transformations, 239
- homomorphism
  - algebra, 43
  - kernel, 658
  - Lie group, 816
  - symmetric, 655
  - trivial, 655
- hyperbolic PDE, 589, *626–628*
- hypergeometric DE, 407
- hypergeometric function, *413–419*
  - contiguous, 417
  - Euler formula, 436
  - integral representation of, 434
- hypergeometric series, 414
- hypersurface, 583
  
- ideal, 44
- idempotent, 688
  - essentially, 689
  - primitive, 689
- identity
  - additive, 20
  - multiplicative, 20
- identity map, 5
- identity operator, 50
- identity representation, 674
- ignorable coordinate, 593
- image of a subset, 5
- implicit function theorem, 350
- indicial polynomial, 406
  
- induced representations, 871
- induction principle, 12
- inductive definition, 14
- inequality
  - Bessel, 149
  - Darboux, 245
  - Parseval, 149
  - Schwarz, 29
  - triangle, 31
- infinitesimal action, *826–831*
  - adjoint, 827
- infinitesimal generator, 827, 830
- inhomogeneous BCs, *574–577*
- inhomogeneous SOLDE
  - general solution, 360
- initial value problem, 559, 583
  - numerical solution, 383
- injective map, 6
- inner automorphism, 825
- inner product, 23–32
  - bra and ket notation, 25
  - definition of, 24
  - norm and, 31
  - positive definite, 25
  - pseudo-Riemannian, 25
  - Riemannian, 25
  - sesquilinear, 25
- inner product space, 25
- integrability condition, 888
- integral curve, 783
- integral equation
  - characteristic value, 489
  - Fredholm, *494–504*
  - Green’s functions, 600
  - kernel of, 488
  - of the first kind, 488
  - of the second kind, 488
  - Volterra, 488
  - Volterra, of second kind
    - solution, 490
- integral operator, 452
- integral transforms, 433
  - Bessel function, 434
- integration on manifolds, *800–801*
- integration operator, 73
- interior product, 793
- intersection, 2

- intrinsic spin, 999
- invariant, 937
- invariant map, 937
- invariant operator
  - matrix representation, 112
- invariant subspace, 112–114, 676, 677
  - definition, 112
- inverse image, 5
- inverse mapping theorem, 777
- inverse of a map, 6
- inverse of a matrix, 97–100
- inversion, 238, 912
- irreducibility criterion, 685
- irreducible representation, 677
  - i*-th row
    - functions, 696
    - norm of functions, 696
- irreducible tensor operators, 705–707
- isometric map, 908
- isometry, 908–912
  - time translation, 923
- isomorphism, 36
  - algebra, 43
  - Lie algebra, 822
  - Lie group, 816
- Jacobi, 183, 416, 490, 615, 663, 702, 704, 738, 807, 888
  - biography, 177
- Jacobi functions, 418
- Jacobi identity, 782, 790, 826
- Jacobi polynomials, 176, 184, 418
  - special cases, 176
- Jacobian matrix, 777
- Jordan arc, 242
- Jordan canonical form, 485
- Jordan's lemma, 275
- Kant, 738
- Kelvin, 561
- Kelvin equation, 538
- Kepler problem, 1001
- kernel, 35
  - degenerate, 501
  - integral operator, 452
  - integral transforms, 433
  - separable, 501
- Killing, 743, 942
  - biography, 842
- Killing equation, 909
- Killing form, 841, 845
- Killing form
  - of  $\mathfrak{gl}(n, \mathbb{R})$ , 844
- Killing vector field, 908–912, 923, 926, 930
  - conformal, 911
- Kirchhoff, 588
- Klein, 388, 743, 798, 849, 891, 920, 942, 996
- Klein–Gordon equation, 328
- Korteweg–de Vries equation, 971
- Kovalevskaya, 463
  - biography, 587
- Kramers–Kronig relation, 309
- Kronecker, 11, 96, 388
  - biography, 738
- Kronecker delta, 26
- Kummer, 30, 704, 738, 842, 891
- Kutta, biography, 390
- Lagrange, 96, 178, 182, 198, 199, 415, 423, 528, 704
  - biography, 984
- Lagrange identity, 366, 434, 525, 560
- Lagrange multiplier, 991
- Lagrange theorem, 670
- Lagrangian, 804, 980
  - null, 987
- Laguerre polynomials, 176, 180–181
- Laplace, 28, 199, 614, 806, 920, 982, 985
  - biography, 527
- Laplace transform, 433
- Laplace's equation, 327
  - Cartesian coordinates, 526
  - cylindrical coordinates, 535
  - spherical coordinates, 541
- Laplacian
  - Green's function for, 595, 629
  - separation of angular part, 331
- Laurent series, 252–262, 606
  - construction, 254
  - uniqueness, 258
- Laurent, biography, 271
- Lavoisier, 95, 985
- left coset, 658

- left ideal, 44, 689
  - minimal, 689
- left translation, 820
  - as action, 827
- left-invariant 1-form, 820
- left-invariant vectors, 820
- Legendre, 177, 199, 490, 614, 888
  - biography, 182
- Legendre equation, 517
- Legendre functions, 419
- Legendre polynomials, 155, 182–184, 340, 359
  - and delta function, 188
  - and Laplacian, 188
- Legendre transformation, 804
- Leibniz, 95, 738
- Leibniz rule, 15
- length of a vector, 30–32
- Levi-Civita, 891
  - biography, 898
- Levi-Civita tensor, 745, 869
- Lie, 711, 743, 798, 842
  - biography, 941
- Lie algebra, 833–845
  - abelian, 834
  - adjoint map, 825
  - Cartan metric tensor, 843
  - Cartan theorem, 844
  - center, 834
  - commutative, 834
  - compact, 843
  - decomposition, 844
  - derivation, 840
  - ideal, 834
  - Killing form of  $\mathfrak{a}$ , 841
  - of a Lie group, 820–826
  - of unitary group, 824
  - of vector fields, 782
  - representation, 859–876
    - definition, 845
  - semisimple, 844
  - simple, 844
  - structure constants, 834
- Lie algebra
  - of  $SL(V)$ , 823
- Lie bracket, 782
- Lie derivative, 788
  - covariant derivative, 902
  - of a 1-form, 790
  - of vectors, 789
- Lie derivative
  - of  $p$ -forms, 793
- Lie group, 815
  - compact
    - characters, 853
    - matrix representation, 852
    - representation, 852
    - unitary representation, 846
    - Weyl operator, 847
  - homomorphism, 816
  - integration, 832–833
    - density function, 833
  - local, 817
    - representation, 845
- Lie multiplication, 833
- Lie subalgebra, 834
- Lie's first theorem, 829
- Lie's second theorem, 826
- Lie's third theorem, 826
- light cone, 837
- linear combination, 21
- linear functional, 34, 455
- linear independence, 21
- linear operator, 33
  - null space of  $\mathfrak{a}$ , 35
- linear PDE, 584
- linear transformation, 32–40
  - bounded, 454
  - definition, 33
  - pullback of  $\mathfrak{a}$ , 39
- Liouville, 512, 654
  - biography, 514
- Liouville substitution, 513, 520
- Liouville's theorem, 807
- Lipschitz condition, 351
- little algebra, 871
- little group, 871, 873
- local diffeomorphism, 768
- local group of transformations, 817
- local Lie group, 817
- local operator, 452
- logarithmic function, 295
- Lorentz, 799
- Lorentz algebra, 864

- Lorentz force law, 794
- Lorentz group, 657, 837
- Lorentz metric, 892
- Lorentz transformations, 837
- lowering indices, 750
- lowering operator, 335
  
- Maclaurin series, 254
- manifold, 764
  - coordinate functions, 764
  - flat, 895
  - orientable, 800
  - product, 767
  - pseudo-Riemannian, 887
  - symplectic, 801
- map, 4–7
  - bijjective, 6
  - codomain, 5
  - differentiable, 768
  - domain, 5
  - equality of, 5
  - functions and, 5
  - graph of a, 5
  - identity, 5
  - image of a subset, 5
  - injective, 6
  - inverse of a, 6
  - one-to-one, 6
  - onto, 6
  - range of a, 5
  - surjective, 6
  - target space, 5
- mathematical induction, 12
- matrix, 82–86
  - antisymmetric, 87
  - basis transformation, 91
  - complex conjugate of, 87
  - diagonal, 88
  - hermitian, 87
  - hermitian conjugate of, 87
  - inverse of a, 98
  - irreducible, 112
  - operations on a, 87
  - orthogonal, 87
  - rank of a, 86
  - reducible, 112, 113
  - representation
    - orthonormal basis, 89
    - row-echelon, 99
    - symmetric, 87
    - transpose of a, 87
    - triangular, 99
    - unitary, 87
- matrix algebra, 42
- maximally symmetric spaces, 910
- Maxwell's equations, 796
- Mellin transform, 434
- Mendelssohn, 615, 738
- meromorphic functions, 293–294
- method of images, 616
  - sphere, 617
- method of steepest descent, 521
- metric space, 7–10
  - convergence, 9
  - definition, 8
- minimal ideal, 44, 856
- Minkowski, 920
- Minkowski metric, 892
- Minkowski space, 754
- Mittag-Leffler, 388, 463, 588
- Mittag-Leffler expansion, 294
- modified Bessel functions, 425
- moment of inertia matrix, 88
- Monge, 95, 198
- Morera's theorem, 252
- multilinear mapping, 729–735
  - tensor-valued, 735
- multiplicative identity, 20
- multivalued functions, 295–302
  
- $n$ -equivalent functions, 945
- $n$ -sphere, 764
- $n$ -th jet space, 945
- Napoleon, 198, 527
- natural isomorphism, 733, 756
- natural pairing, 731
- neighborhood, 459
- Neumann, 177, 702
  - biography, 620
- Neumann BC, 591
- Neumann BVP, 591, 619–621
- Neumann function, 424
- Neumann series, 493, 601, 602
- Newton, 330, 415, 527, 798, 806, 982

- Newtonian gravity, 916–918  
 Noether, 704  
     biography, 996  
 Noether's theorem, 995  
 non-local potential, 631  
 norm  
     of a vector, 30  
     operator, 454  
     product of operators, 456  
 normal operator, 117  
 normal subgroup, 659  
 normal vectors, 26  
 normed linear space, 31  
 normed vector space  
     compact subset of, 461  
 null divergence, 993  
 null Lagrangian, 987  
 null space, 35  
 null vector, 751, 837  
 numerical integration, 74–76  
     Simpson's 1/3 rule, 75  
     Simpson's 3/8 rule, 75  
     trapezoidal rule, 75  
  
 Ohm, 614  
 Olbers, 422  
 one-form, 786  
 one-parameter group, 784  
 open ball, 459  
 open subset, 459  
 operations on matrices, 87–89  
 operator, 33  
     adjoint, 61  
         existence of, 457  
     angular momentum, 331  
     annihilation, 374  
     bounded, 454  
     closed, 508  
         bounded, 508  
     compact Hermitian  
         spectral theorem, 476  
     compact normal  
         spectral theorem, 477  
     compact resolvent, 508  
     creation, 374  
     diagonalizable, 116  
     differentiation, 73  
     domain of, 507  
     equality, 49  
     evolution, 58  
     expectation value of, 64  
     extension of, 508  
     finite rank, 464  
     formally self-adjoint, 597  
     functions of an, 53  
     hermitian, 63, 508  
         eigenvalue, 117  
     hermitian conjugate of, 61  
     Hilbert–Schmidt, 465, 511  
     integration, 73  
     inverse of an, 50  
     kernel of an, 35  
     local, 452  
     negative powers of, 52  
     norm of an, 454  
     normal, 117  
         diagonalizable, 119  
         eigenspace of, 118  
     null space of an, 35  
     numerical analysis, 70  
     polarization identity, 50  
     positive, 65  
     projection, 67  
     projection  
         orthogonal, 68  
     pullback of an, 39  
     raising and lowering, 335  
     regular point, 457  
     representation of an, 83  
     right-shift, 453  
     scalar, 706  
     self-adjoint, 508  
     spectrum, 457, 457–458  
     spectrum of an, 115  
     square root of, 127  
     Sturm–Liouville, 510  
     trace of an, 102  
     unbounded, 507  
     unitary, 66, 127  
         eigenvalue, 117  
     operator algebra, 49–76  
     operator polynomials, 51–53  
     operator spectrum, 457  
      $\mathfrak{o}(p, n - p)$ , 837

- optical theorem, 308
- orbit, 663, 676, 818
- orbital angular momentum, 999
- ordered pairs, 2
- orientable manifolds, 800
- orientation, 745, 746
  - positive, 746
- oriented basis, 745
- orthochronous, 837
- orthogonal group, 656, 824
  - Lie algebra of, 824
- orthogonal polynomials, 153
  - classical, 173
    - classification, 176
    - differential equation, 174
    - generating functions, 189
    - recurrence relations, 176
  - expansion, 186
  - least square fit, 155
- orthogonal transformation, 97
- orthogonal vectors, 26
- orthogonality, 26–27
- orthonormal basis, 26
- orthonormal frames, 887
  
- p*-forms, 741
- pairing
  - natural, 731
- parabolic PDE, 589, 621–626
- parallel transport, 906
- parallelogram law, 31
- parity, 667
  - Hermite polynomials, 194
  - Legendre polynomials, 194
- Parseval inequality, 149, 851
- Parseval's relation, 223
- partial DE, 584–592
  - characteristic system of, 939
  - order of, 584
  - principal part, 584
- particle in a box, 531
- particle in a cylindrical can, 539
- particle in a hard sphere, 543
- partition, 4, 669
- past light cone, 838
- Pauli spin matrices, 88, 835, 840
- Peano, 799
  
- periodic BC, 515
- permutation
  - even, 668
  - odd, 668
  - parity of, 667
- permutation group, 664
- permutation symbol, 93
- permutation tensor, 754
- perturbation theory, 603–610, 697
  - degenerate, 609
  - first-order, 608
  - nondegenerate, 607
  - second-order, 608
- Peter–Weyl theorem, 853
- Peter–Weyl theorem
  - Fourier series, 853
- photon capture
  - cross section, 929
- Planck, 388, 463, 920
- Poincaré, 63, 478, 497, 620, 744, 920
  - biography, 797
- Poincaré algebra, 840, 844
  - representation, 868–876
- Poincaré group, 657, 817, 839, 872
- Poincaré lemma, 797
  - converse of, 797
- Poisson, 178, 512, 528, 614, 653
- Poisson bracket, 808
- Poisson's equation, 327, 917
- polar decomposition, 129–130
- pole, 273
- polynomials
  - orthogonal, 153
- positive operator, 65
- positive orientation, 746
- potential
  - non-local, 631
  - separable, 631
- power series, 252
  - differentiation of, 253
  - integration of, 253
  - uniform convergence, 253
- $p(p, n - p)$ , 837
- preimage, 5
- primitive idempotent, 689
- principal value, 285
- product manifold, 767

- projectable symmetry, 944
- projection operator, 67–70, 474, 604
  - completeness relation, 68
  - orthogonal, 68
- projective group
  - density function, 833
  - one-dimensional, 820
- projective space, 4
- prolongation, 945–950
  - of a function, 946
  - of groups, 948
  - of vector fields, 949
- propagator, 626
  - Feynman, 636
- proper orthochronous, 837
- proper subset, 2
- Prüfer substitution, 518
- Puiseux, biography, 296
- pullback, 735, 786, 791, 801
  - linear transformation, 39
- pullback map, 791
- pullback of  $p$ -forms, 741
- quantization
  - harmonic oscillator
    - algebraic, 375
    - analytic, 374
  - hydrogen atom, 422
- quarks, 701, 703, 873
- quaternions, 44
- quotient group, 660
- quotient set, 4
- raising indices, 750
- raising operator, 335
- range of a map, 5
- rank of a matrix, 86
- rational function, 274
  - and trig functions
    - integrals of, 279
  - integration of, 276
- rational numbers
  - dense subset of reals, 460
- real coordinate space, 21
- real vector space, 20, 130–137
- realization, 664
- reciprocal lattice vectors, 208
- recurrence relations, 153
- redshift, 930
- reduced matrix elements, 707
- reducible representation, 677
- regular graphs, 713
- regular point, 233
- regular representation, 686
- regular singular point
  - second-order linear DE, 405
- relation, 3
- relative acceleration, 913
- relativistic electromagnetism, 792
- representation
  - abelian group, 681
  - adjoint, 679, 704
  - carrier space, 674
  - character, 683
  - compact Lie group, 852
  - complex conjugate, 679
  - dimension of, 674
  - direct sum, 679
  - equivalent, 674
  - faithful, 674
  - general linear group, 856–859
  - group, 674
  - identity, 674
  - irreducible, 677
  - Kronecker product, 700
  - Lie group, 845
  - operators, 83
  - reducible, 677
  - regular, 686
  - symmetric group, 707–723
  - tensor product, 700
    - antisymmetrized, 701
    - character, 700
    - symmetrized, 701
  - trivial, 679
  - unitary, 678
    - compact Lie group, 846
- representation of
  - $gl(n, \mathbb{R})$ , 860
  - $sl(n, \mathbb{C})$ , 861
  - $so(3)$ , 864
  - $so(3, 1)$ , 867
  - $su(n)$ , 861
  - $u(n)$ , 861
- residue, 270–273

- definite integrals, 275
- definition, 271
- residue theorem, 271
- resolution of identity, 482, 689, 721
- resolvent, 480–485
  - compact, 508
  - Laurent expansion, 481
- resolvent set, 457
  - openness of, 460
- resonant cavities, 533
- Riccati equation, 968
- Ricci, 891, 898
- Ricci tensor, 917, 919, 920
- Riemann, 30, 199, 296, 704, 798, 849, 891, 981
  - biography, 888
- Riemann curvature tensor, 889, 904
- Riemann normal coordinates, 915
- Riemann surfaces, 296–302
- Riemannian manifold, 887–896
- right coset, 658
- right ideal, 44
- right translation, 820
- right-invariant 1-form, 820
- right-invariant vectors, 820
- right-shift operator, 453
  - eigenvalues of, 458
- rigid rotations, 657
- Rodriguez formula, 174
- Rosetta stone, 198
- rotation algebra, 864
- rotation group, 675, 862
  - character, 865
- rotation matrix, 865
- Runge, biography, 388
- Runge-Kutta method, 387–391
- Russell, 11, 799
  
- saddle point approximation, 313
- scalar operator, 706, 707
- scale transformations, 819
- Schelling, 738
- Schmidt, biography, 28
- Schopenhauer, 738
- Schrödinger, 64, 807
- Schrödinger equation, 675
- Schrödinger equation, 58, 328, 410, 420, 531, 543, 631
  - classical limit, 383
  - one dimensional, 381
- Schur, 711, 850, 874
  - biography, 681
- Schur's lemma, 680, 681, 707, 845, 861
- Schwarz, 463, 739
  - biography, 30
- Schwarz inequality, 29, 149, 152
- Schwarz reflection principle, 305–306
- Schwarzschild geodesic, 925–932
- Schwarzschild metric, 892
- Schwarzschild radius, 925
- Schwarzschild, biography, 919
- second order linear DE
  - Sturm-Liouville systems, 513
- second order PDE, 589–592
  - hyperbolic, 589
- second-order PDE
  - classification, 589
  - elliptic, 589
  - parabolic, 589
  - ultrahyperbolic, 589
- selection rules, 701
- self-adjoint operator, 63
- self-adjoint SLDs, 564–565
- semisimple Lie algebra, 844
- separable kernel, 501
- separable potential, 631
- separated boundary conditions, 510
- separation of variables, 328
  - Cartesian, 526–534
  - cylindrical, 535–540
  - spherical, 540–544
- separation theorems, 361–363
- sequence, 9
  - Cauchy, 10
  - complete orthonormal, 149
- series
  - Fourier–Bessel, 536
  - Neumann, 601, 602
- sesquilinear inner product, 25
- set, 1–3
  - Cantor, 12
  - complement of, 2
  - countably infinite, 12
  - element of, 1
  - empty, 2



- intersection, 2
  - partition of  $a$ , 4
  - uncountable, 12
  - union, 2
  - universal, 2
- sharp map, 747, 802
- shifting operator, 71
- signature of  $\mathfrak{g}$ , 752
- similarity transformation, 91–92
  - orthonormal basis, 91
- simple arc, 242
- simple character, 685
- simple Lie algebra, 844
- simultaneous diagonalizability, 123
- singleton, 2
- singular point, 233
- singularity, 233
  - isolated, 270
    - classification, 273
    - rational function, 274
- skew-symmetric, 738
- $SL(V)$  as a Lie group, 816
- smooth arc, 242
- space
  - Banach, 148
  - complex coordinate, 21
  - dual, 34
  - inner product, 25
  - metric, 8
  - real coordinate, 21
  - square-integrable functions, 151
- spacelike vector, 837
- spacetime
  - spherically symmetric, 923
  - static, 923
  - stationary, 923
- span, 22
- special linear group, 656
- special orthogonal group, 657, 824
  - Lie algebra of, 824
- special relativity, 752, 837, 868, 872, 986
- special unitary group, 657
  - Lie algebra of, 824
- spectral decomposition, 117–125
  - orthogonal operator, 135
  - real case, 131
- spectral theorem
  - compact hermitian, 476
  - compact normal, 477
- spectrum
  - closure of, 461
- spherical Bessel functions, 428, 542
  - expansion of plane wave, 543
- spherical coordinates
  - in  $m$  dimensions, 593–595
- spherical harmonics, 338–346
  - addition theorem, 345, 867
  - definition, 341
  - expansion in terms of, 344
  - expansion of plane wave, 544, 646
  - first few, 343
- Spinoza, 738
- square-integrable functions, 151
- stabilizer, 663, 818
- static spacetime, 923
- stationary spacetime, 923
- steepest descent method, 312–315
- step function, 163, 288
- stereographic projection
  - $n$ -sphere, 769
  - two-sphere, 766
- Stirling approximation, 316
- Stone-Weierstrass theorem, 152
  - generalized, 196
- stress energy tensor, 920
- structure constants, 826, 900, 905
- structure equations, 888
- Sturm, biography, 512
- Sturm–Liouville
  - operator, 510
  - problem, 174, 622
  - system, 510, 637
    - completeness, 524
    - eigenfunctions, 511
    - eigenvalues, 513
    - large argument, 521
    - large eigenvalues, 518
    - regular, 510
    - singular, 516
- subalgebra, 44
- subgroup, 656–663
  - conjugate, 657
  - generated by a subset, 657
  - normal, 659

- trivial, 656
- submanifold, 767
  - open, 767
- subset, 2
  - bounded, 459
  - proper, 2
- subspace, 22
  - invariant, 112
- surjective map, 6
- Sylvester's theorem, 753
- symmetric algebra, 737
- symmetric bilinear form, 749
  - classification, 751
  - inner product, 750
  - nondegenerate, 750
- symmetric group, 655, 664–669
  - characters
    - graphical construction, 714
  - cycle, 666
  - representation, 707–723
    - analytic construction, 707
    - antisymmetric, 679
    - graphical construction, 710
    - products, 721
    - Young operators, 717
- symmetric homomorphism, 655
- symmetric product, 737
- symmetrizer, 736
- symmetry group, 944
  - defining equations, 957
  - of a DE, 941
  - of a subset, 937
  - projectable, 944
  - transform of a function, 943
  - variational, 988
- symplectic algebra, 836
- symplectic charts, 802
- symplectic form, 746, 801
  - rank of, 747
- symplectic geometry, 801–808
  - conservation of energy, 805
- symplectic group, 657, 749, 835
- symplectic manifold, 801
- symplectic map, 746, 802
- symplectic matrix, 749
- symplectic structure, 801
- symplectic transformation, 746
- symplectic vector space, 746–749
  - canonical basis, 747
  - Hamiltonian dynamics, 748
- tangent bundle, 780
- tangent space, 773
- tangent vector, 772
- tangents to a curve
  - components, 778
- target space, 5
- Taylor series, 252–262
  - construction, 254
- tensor, 731
  - components of, 732
  - contravariant, 731
  - contravariant-symmetric, 736
  - covariant, 731
  - covariant-symmetric, 736
  - Levi-Civita, 745
  - skew-symmetric, 738
  - symmetric, 736
  - transformation law, 733
  - types of, 731
- tensor algebra, 731
- tensor bundle, 786
- tensor field, 785–790
  - crucial property of, 786, 787
- tensor product, 730, 731
  - of vector spaces, 41
- tensors
  - symmetric product, 737
- theta function, 288
- timelike vector, 837
- total derivative, 954
- total divergence, 987
- trace, 101–103
  - and determinant, 102
  - definition, 101
  - log of determinant, 103
  - relation to determinant, 101
- transformation group, 655
- translation, 819
- translation operator, 138
- transpose of a matrix, 87
- transposition, 667
- traveling waves, 533
- triangle inequality, 31

- trivial homomorphism, 655
- trivial representation, 679
- trivial subgroup, 656
- twin paradox
  - as a variational problem, 986
- uncertainty principle, 78
- uncountable set, 12
- union, 2
- unitary group, 656
  - Lie algebra of, 824
- unitary operator, 63–67
- unitary representation, 678
- universal set, 2
- Vandermonde, biography, 95
- variation of constants, 360
- variational derivative, 977
- variational problem, 980
  - twin paradox, 986
- variational symmetry group, 988
- vector field, 780
  - flow of  $a$ , 784
  - integral curve of, 783
  - left-invariant, 820
  - Lie algebra of, 782
- vector space, 19–45
  - automorphism, 36
  - basis
    - components in  $a$ , 22
  - basis of  $a$ , 22
  - complete, 146
  - complex, 20
  - definition, 19
  - dual, 34
  - endomorphism of  $a$ , 33
  - finite-dimension
    - criterion for, 462
  - finite-dimensional, 22
  - isomorphism, 36
  - linear operator on  $a$ , 33
  - normed, 31
  - operator on  $a$ , 33
  - oriented, 746
  - real, 20
  - symplectic, 746
- vectors, 19
  - norm of, 30
  - normal, 26
  - orthogonal, 26
- Volterra equations, 488
- Volterra, biography, 490
- volume element, 746
- von Dyck, 390
- von Humboldt, 177, 614, 738
- von Neumann, 874
  - biography, 478
- wave equation, 328, 532
  - symmetry group, 961–963
- wave guide, 533
  - cylindrical, 537
  - rectangular, 534
- wedge product, 740
- Weierstrass, 11, 30, 296, 388, 588, 739, 841
  - biography, 462
- Weyl, 744, 842, 942, 996
  - biography, 849
- Weyl basis, 835, 844
- Weyl operator, 847
- Wigner, 168, 943
  - biography, 874
- Wigner formula, 865
- Wigner–Eckart theorem, 707
- Wigner–Seitz cell, 208
- WKB method, 380–383
  - connection formulas, 381
- Wordsworth, 806
- Wronski, biography, 357
- Wronskian, 355–363, 511
  - comparison theorem, 362
  - separation theorem, 361
- Young, 850
- Young frame, 712, 718
  - negative application, 715
  - positive application, 715
  - regular application, 714
- Young operator, 856
- Young pattern, 712
- Young tableaux, 713, 718, 857
- Yukawa potential, 214