ForgEd

LEARN SKILLS THAT MATTER.

Multivariable Calculus & Quantum Mechanics

FROM VECTORS TO THE SCHRÖDINGER EQUATION

FORGED MATH SERIES · VOLUME IV

FORGED OFFICIAL TEXTBOOK

An eighteen-chapter, upper-division text on vector calculus, complex analysis, linear algebra, differential equations, Fourier analysis, and nonrelativistic quantum mechanics.

First edition · Theorems, proofs, worked examples, and verified exercises

OMNISTRATA, LLC

PART I Multivariable Calculus PART II Mathematics of Quantum Theory PART III Quantum Mechanics



Multivariable Calculus & Quantum Mechanics

FROM VECTORS TO THE SCHRÖDINGER EQUATION

Omnistrata, LLC

FORGED MATH SERIES · VOLUME IV

ForgEd Official Textbook · First Edition

ForgEd • forgedlearn.com
Learn skills that matter.

Multivariable Calculus & Quantum Mechanics
ForgEd Math Series, Volume IV
First Edition

Copyright © 2026 Omnistrata, LLC. All rights reserved.

Published by Omnistrata, LLC under the ForgEd imprint, forgedlearn.com.

No part of this book may be reproduced, stored, or transmitted in any form or by any means without the prior written permission of the publisher, except for brief quotations in reviews and for classroom use by the purchaser.

The ForgEd Math Series
Volume I Algebra
Volume II Precalculus & Trigonometry
Volume III Calculus I–II
Volume IV Multivariable Calculus & Quantum Mechanics

Physical constants are the CODATA 2022 recommended values. Numerical results in the text, worked problems, and answers were checked by computer algebra and numerical computation.

PREFACE

This book is a single path from the vector algebra of three-dimensional space to the quantum mechanics of atoms and spins. It is Volume IV of the ForgEd Math Series, written for a reader who has finished single-variable calculus (Volume III) and is willing to work. The pace is that of an upper-division university course: definitions are stated precisely, theorems are proved when the proof teaches something, and the exercises are meant to take time.

HOW THE BOOK IS ORGANIZED

Part I (Chapters 1–7) develops multivariable and vector calculus, ending with the theorems of Green, Stokes, and Gauss. **Part II** (Chapters 8–11) supplies the mathematics quantum theory is written in: complex analysis, linear algebra on complex inner-product spaces, ordinary and partial differential equations, and Fourier analysis. **Part III** (Chapters 12–18) is a first course in nonrelativistic quantum mechanics, from the experiments that forced it into existence through the hydrogen atom, angular momentum and spin, identical particles, and entanglement. Three appendices collect physical constants, a mathematical reference, and five numerical methods for quantum mechanics, with working Python programs.

PREREQUISITES

Chapters	Assumed background
1	Algebra and trigonometry (Volumes I–II)
2–7	Single-variable calculus through techniques of integration (Volume III)
8–11	Part I
12–18	Parts I and II; Chapters 9 and 10 carry the most weight

How to read this book

Each chapter opens with its learning objectives. Boxed material comes in four kinds.

Definition 0.1 (Definitions, theorems, propositions). Shaded boxes contain the statements you should be able to reproduce from memory. Theorems carry a gold rule.

Example 0.2 (Examples). Worked problems. Cover the solution and attempt the problem before reading it.

COMMON ERROR. Black-ruled boxes describe the specific mistakes students make most often, and how to detect them.

LOOKING AHEAD. Dark boxes connect the current material to where it is used later, usually in quantum mechanics.

Theorems, definitions, and examples share one numbering sequence within each chapter, so Example 3.7 follows Theorem 3.6.

Worked problems, exercises, and answers

Every chapter closes with a section of *Worked Problems*, just before the exercises: longer problems, solved in full, that combine several ideas from the chapter and show the reasoning a careful solver would record. They are the bridge between the examples and the exercises. Exercises are grouped as *Computation*, *Theory and Proof*, and *Challenge*. Answers to odd-numbered computational exercises appear at the back of the book. Every numerical answer in the examples, the worked problems, and the answer section was checked by computer algebra or numerical computation before publication; if your result disagrees, find the step where it diverges.

Notation

Vectors are set in bold ($\mathbf{u}$, $\mathbf{F}$); unit vectors carry a hat ($\hat{\mathbf{n}}$). Components are written $\langle u_1, u_2, u_3 \rangle$. In spherical coordinates this book uses the physics convention

throughout: r is distance from the origin, $\theta \in [0, \pi]$ is the polar angle measured from the positive z-axis, and $\phi \in [0, 2\pi)$ is the azimuthal angle. Many calculus texts swap θ and ϕ; quantum mechanics does not.

CONTENTS

Preface iv

Part I Multivariable Calculus 1

1 Vectors and the Geometry of Space 2
- 1.1 Euclidean Space . . . 3
- 1.2 Vectors and Their Algebra . . . 4
- 1.3 The Dot Product . . . 7
- 1.4 The Cross Product . . . 10
- 1.5 Lines and Planes . . . 14
- 1.6 Cylinders and Quadric Surfaces . . . 16
- 1.7 Cylindrical and Spherical Coordinates . . . 17
- 1.8 Worked Problems . . . 19
- Exercises . . . 24

2 Vector Functions and Motion in Space 28
- 2.1 Vector-Valued Functions . . . 28
- 2.2 Derivatives and Integrals . . . 30
- 2.3 Arc Length and Reparametrization . . . 31
- 2.4 Curvature . . . 33
- 2.5 The Frenet Frame and Torsion . . . 34
- 2.6 Motion in Space . . . 36
- 2.7 Worked Problems . . . 38
- Exercises . . . 42

3 Partial Derivatives and the Gradient **46**
3.1 Functions of Several Variables . . . 47
3.2 Limits and Continuity . . . 48
3.3 Partial Derivatives . . . 49
3.4 Differentiability and Linear Approximation . . . 51
3.5 The Chain Rule . . . 53
3.6 Directional Derivatives and the Gradient . . . 54
3.7 The Implicit and Inverse Function Theorems . . . 56
3.8 Worked Problems . . . 57
Exercises . . . 62

4 Optimization in Several Variables **66**
4.1 Taylor's Theorem in Several Variables . . . 67
4.2 Quadratic Forms . . . 68
4.3 Local Extrema . . . 69
4.4 Absolute Extrema . . . 71
4.5 Lagrange Multipliers . . . 72
4.6 Applications . . . 74
4.7 Worked Problems . . . 76
Exercises . . . 80

5 Multiple Integrals **83**
5.1 Double Integrals over Rectangles . . . 83
5.2 Double Integrals over General Regions . . . 85
5.3 Polar Coordinates . . . 86
5.4 Applications of Double Integrals . . . 88
5.5 Triple Integrals . . . 89
5.6 Cylindrical and Spherical Coordinates . . . 90
5.7 Change of Variables . . . 91
5.8 The Gamma Function and Gaussian Moments . . . 93
5.9 Worked Problems . . . 95
Exercises . . . 99

6 Vector Fields and Line Integrals **102**
6.1 Vector Fields . . . 103
6.2 Line Integrals of Scalar Functions . . . 104
6.3 Line Integrals of Vector Fields . . . 105
6.4 The Fundamental Theorem for Line Integrals . . . 106
6.5 Divergence and Curl . . . 109
6.6 Green's Theorem . . . 111
6.7 Worked Problems . . . 114
Exercises . . . 117

7 Surface Integrals and the Integral Theorems **121**
7.1 Parametric Surfaces . . . 122
7.2 Surface Area . . . 123
7.3 Surface Integrals . . . 123
7.4 Stokes' Theorem . . . 125
7.5 The Divergence Theorem . . . 127
7.6 The Dirac Delta Function . . . 129
7.7 Curvilinear Coordinates . . . 131
7.8 The Unity of the Integral Theorems . . . 133
7.9 Worked Problems . . . 134
Exercises . . . 138

Part II The Mathematics of Quantum Theory **143**

8 Complex Numbers and Complex Analysis **144**
8.1 Complex Numbers . . . 145
8.2 Polar Form and Euler's Formula . . . 146
8.3 Complex Functions . . . 148
8.4 Analytic Functions and the Cauchy–Riemann Equations . . . 149
8.5 Contour Integrals . . . 150
8.6 Cauchy's Theorem and Integral Formula . . . 151
8.7 Series, Singularities, and Residues . . . 153

8.8 Evaluating Real Integrals . . . 155
8.9 Worked Problems . . . 157
Exercises . . . 162

9 Linear Algebra and Hilbert Space 166
9.1 Vector Spaces . . . 167
9.2 Linear Maps and Matrices . . . 168
9.3 Inner Product Spaces . . . 169
9.4 Eigenvalues and Eigenvectors . . . 172
9.5 Hermitian and Unitary Operators . . . 173
9.6 Dirac Notation . . . 177
9.7 Infinite Dimensions: Spaces of Functions . . . 178
9.8 Worked Problems . . . 179
Exercises . . . 184

10 Ordinary and Partial Differential Equations 188
10.1 First-Order Equations . . . 189
10.2 Second-Order Linear Equations . . . 190
10.3 Linear Systems and Normal Modes . . . 192
10.4 Series Solutions and Special Functions . . . 193
10.5 Sturm–Liouville Theory . . . 195
10.6 Partial Differential Equations: Separation of Variables . . . 196
10.7 Worked Problems . . . 198
Exercises . . . 203

11 Fourier Analysis 206
11.1 Fourier Series . . . 207
11.2 Convergence and Parseval's Identity . . . 208
11.3 The Fourier Transform . . . 210
11.4 The Delta Function and Completeness . . . 212
11.5 Convolution and the Heat Equation on the Line . . . 213
11.6 The Uncertainty Principle for Fourier Transforms . . . 213
11.7 The Discrete Fourier Transform . . . 215

11.8 Worked Problems 215
Exercises 219

Part III Quantum Mechanics 223

12 The Breakdown of Classical Physics 224
12.1 Equipartition and Its Discontents 225
12.2 Blackbody Radiation 226
12.3 The Photoelectric Effect 228
12.4 Compton Scattering 230
12.5 Atoms: Rutherford, Bohr, and the Spectrum of Hydrogen 231
12.6 Matter Waves 233
12.7 The Double-Slit Experiment 233
12.8 The Stern–Gerlach Experiment 235
12.9 Toward a New Theory 236
12.10Worked Problems 236
Exercises 241

13 The Wavefunction and the Schrödinger Equation 244
13.1 The Schrödinger Equation 245
13.2 The Statistical Interpretation 246
13.3 Conservation of Probability 247
13.4 Expectation Values and Operators 248
13.5 Ehrenfest's Theorem and the Classical Limit 250
13.6 The Uncertainty Principle 250
13.7 The Free Gaussian Wave Packet 251
13.8 Stationary States 253
13.9 Worked Problems 255
Exercises 260

14 Stationary States in One Dimension 263
14.1 The Infinite Square Well 264

14.2 The Harmonic Oscillator: Algebraic Method 265
14.3 The Harmonic Oscillator: Analytic Method 268
14.4 The Free Particle and Scattering States 269
14.5 The Delta-Function Potential 270
14.6 The Finite Square Well 271
14.7 Steps, Barriers, and Tunneling 273
14.8 Worked Problems 275
Exercises 280

15 The Formalism: Hilbert Space, Operators, and Uncertainty 283
15.1 The Postulates 284
15.2 The Generalized Uncertainty Principle 285
15.3 Time Evolution 287
15.4 Symmetry and Conservation 288
15.5 The Variational Principle 289
15.6 Time-Independent Perturbation Theory 290
15.7 Worked Problems 292
Exercises 297

16 Quantum Mechanics in Three Dimensions: The Hydrogen Atom 300
16.1 Three Dimensions and Degeneracy 301
16.2 Central Potentials and Separation of Variables 301
16.3 The Angular Equation and Spherical Harmonics 302
16.4 The Radial Equation for Hydrogen 304
16.5 Properties of the Hydrogen States 306
16.6 The Spectrum 307
16.7 Beyond the Schrödinger Atom 307
16.8 Worked Problems 308
Exercises 312

17 Angular Momentum and Spin 315
17.1 Orbital Angular Momentum 316
17.2 The Algebraic Theory of Angular Momentum 317

17.3 Spin One-Half . . . 318
17.4 Spin in a Magnetic Field . . . 319
17.5 Addition of Angular Momenta . . . 320
17.6 Worked Problems . . . 322
Exercises . . . 326

18 Two-Level Systems, Identical Particles, and Entanglement 329
18.1 Two-Level Systems . . . 330
18.2 Composite Systems and Entanglement . . . 332
18.3 Identical Particles . . . 333
18.4 Bell's Theorem . . . 335
18.5 Qubits and Quantum Information . . . 336
18.6 Worked Problems . . . 337
Exercises . . . 342

A Physical Constants and Units 345

B Mathematical Reference 348
B.1 Vector Identities . . . 348
B.2 Curvilinear Coordinates . . . 349
B.3 The Integral Theorems . . . 350
B.4 Integrals . . . 350
B.5 Series . . . 351
B.6 Fourier Transforms . . . 351
B.7 Special Functions . . . 352
B.8 Spin and Angular Momentum . . . 353

C Numerical Methods for Quantum Mechanics 355
C.1 Finite Differences: The Hamiltonian as a Matrix . . . 355
C.2 Shooting: Quantization as a Boundary Condition . . . 356
C.3 Basis Expansion: Rayleigh–Ritz with Gaussians . . . 358
C.4 Time Evolution: The Crank–Nicolson Method . . . 359
C.5 The Split-Operator Method . . . 361

C.6 Choosing a Method 362

Further Reading **363**

Answers to Selected Exercises **365**

Index **372**



PART I

Multivariable Calculus

Vectors, curves, partial derivatives, optimization, multiple integrals, and the integral theorems of vector calculus.

CHAPTER 1

Vectors and the Geometry of Space

LEARNING OBJECTIVES

- Work with points, distances, and spheres in $\mathbb{R}^3$, and extend these ideas to $\mathbb{R}^n$.
- State the vector-space axioms and manipulate vectors algebraically and geometrically.
- Define the dot product, prove the Cauchy–Schwarz and triangle inequalities, and compute angles and projections.
- Define the cross product, derive its geometric meaning, and use index notation with δ_{ij} and ε_{ijk}.
- Describe lines, planes, and quadric surfaces, and compute distances between them.
- Convert among Cartesian, cylindrical, and spherical coordinates using the physics convention.

Physics happens in space, and the language for space is the vector. A velocity, a force, an electric field, the gradient of a temperature, the angular momentum of an electron: each has a magnitude and a direction, and each obeys the same algebra. This chapter builds that algebra carefully, because every later chapter depends on it. In quantum mechanics the same structure reappears in a more abstract form. States are vectors, measurement amplitudes are inner products, and the cross product's index form becomes the commutation relations of angular momentum.

1.1 Euclidean Space

Fix a point O, the *origin*, and three mutually perpendicular number lines through it: the x-, y-, and z-axes. We orient them as a *right-handed system*. If the fingers of the right hand curl from the positive x-axis toward the positive y-axis, the thumb points along the positive z-axis. Each point P in space then corresponds to exactly one ordered triple (x, y, z) of real numbers, its *coordinates*. The set of all such triples is

$$\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}.$$

The three *coordinate planes* $z = 0$, $x = 0$, and $y = 0$ divide space into eight *octants*. The *first octant* is the one in which all three coordinates are positive.

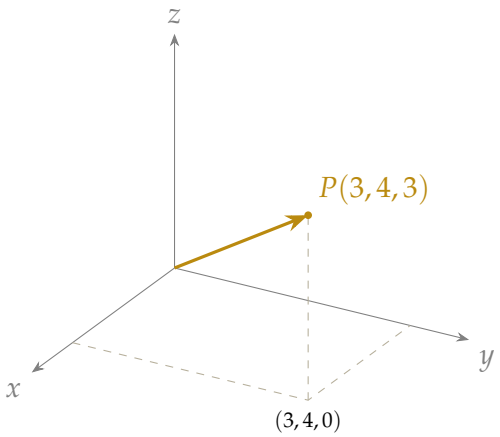


Figure 1.1: The point $P(3, 4, 3)$ and its projection onto the xy-plane.

Theorem 1.1 (Distance formula). *The distance between $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ is*

$$|P_1P_2| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Proof. Let $Q = (x_2, y_2, z_1)$. The segment P_1Q lies in the horizontal plane $z = z_1$, so by the planar distance formula $|P_1Q|^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$. The segment QP_2 is vertical, with $|QP_2| = |z_2 - z_1|$, and it is perpendicular to P_1Q. The Pythagorean theorem in triangle P_1QP_2 gives $|P_1P_2|^2 = |P_1Q|^2 + |QP_2|^2$. ∎

A *sphere* with center $C(h, k, l)$ and radius $r > 0$ is the set of points at distance r

from C. By Theorem 1.1 it has equation

$$(x-h)^2+(y-k)^2+(z-l)^2=r^2. \tag{1.1}$$

Expanding (1.1) gives $x^2+y^2+z^2+ax+by+cz+d=0$. Conversely, completing the square shows that an equation of this form describes a sphere, a single point, or the empty set, according to whether the resulting right-hand side is positive, zero, or negative.

Example 1.2 (Recognizing a sphere). Describe the set $x^2+y^2+z^2-4x+6y-2z-2=0$.

Solution. Completing the square in each variable,

$$(x-2)^2+(y+3)^2+(z-1)^2=2+4+9+1=16.$$

This is the sphere with center $(2,-3,1)$ and radius 4.

Example 1.3 (A region between spheres). Describe the solid region $\{(x,y,z):1\le x^2+y^2+z^2\le 4,\ z\ge 0\}$.

Solution. The inequalities say that the distance from the origin lies between 1 and 2. The region is the upper half of a spherical shell of inner radius 1 and outer radius 2. Its volume is $\frac{1}{2}\cdot\frac{4}{3}\pi(2^3-1^3)=\frac{14\pi}{3}$.

More generally, $\mathbb{R}^n$ is the set of ordered n-tuples $(x_1,\dots,x_n)$, with distance $\left(\sum_{i=1}^{n}(x_i-y_i)^2\right)^{1/2}$. Nothing in this chapter except the cross product depends on $n=3$. In Chapter 9 we will replace $\mathbb{R}^n$ by abstract vector spaces, including spaces of functions, and the geometry developed here will carry over almost unchanged.

1.2 Vectors and Their Algebra

Definition 1.4 (Vector). A *vector* in $\mathbb{R}^3$ is an ordered triple $\mathbf{v}=\langle v_1,v_2,v_3\rangle$ of real numbers, its *components*. Geometrically, $\mathbf{v}$ is represented by any directed line segment whose head minus tail equals (v_1,v_2,v_3). Two directed segments represent the same vector exactly when they have the same length and direction. The vector from $P(x_1,y_1,z_1)$ to $Q(x_2,y_2,z_2)$ is

$$\overrightarrow{PQ}=\langle x_2-x_1,\ y_2-y_1,\ z_2-z_1\rangle\,.$$

The vector $\overrightarrow{OP}$ from the origin to P is the *position vector* of P.

Angle brackets distinguish the vector $\langle 1, 2, 3\rangle$ from the point $(1, 2, 3)$. Real numbers, in this context, are called *scalars*.

Definition 1.5 (Vector operations). For $\mathbf{u} = \langle u_1, u_2, u_3\rangle, \mathbf{v} = \langle v_1, v_2, v_3\rangle$, and $c \in \mathbb{R}$:

$$\mathbf{u} + \mathbf{v} = \langle u_1 + v_1,\ u_2 + v_2,\ u_3 + v_3\rangle\,,$$
$$c\,\mathbf{u} = \langle cu_1,\ cu_2,\ cu_3\rangle\,, \qquad |\mathbf{u}| = \sqrt{u_1^2 + u_2^2 + u_3^2}.$$

The *zero vector* is $\mathbf{0} = \langle 0, 0, 0\rangle$, and $-\mathbf{u} = (-1)\mathbf{u}$. The number $|\mathbf{u}|$ is the *magnitude* or *norm* of $\mathbf{u}$. A *unit vector* is a vector of norm 1.

Geometrically, $\mathbf{u} + \mathbf{v}$ is the diagonal of the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$ (the *parallelogram law*). Equivalently, place the tail of $\mathbf{v}$ at the head of $\mathbf{u}$; the sum runs from the tail of $\mathbf{u}$ to the head of $\mathbf{v}$ (the *triangle law*). Scalar multiplication stretches by the factor $|c|$ and reverses direction when $c < 0$.

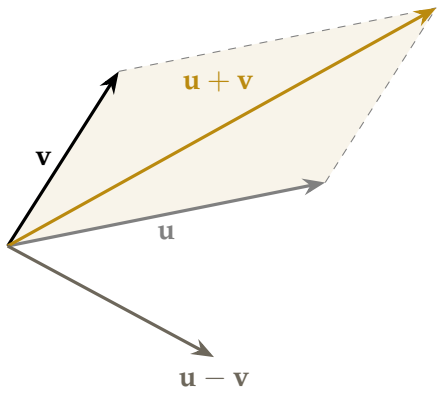


Figure 1.2: The parallelogram law. The other diagonal of the parallelogram is $\mathbf{u} - \mathbf{v}$ (drawn from the origin).

Theorem 1.6 (Vector-space properties). *For all vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *in* $\mathbb{R}^3$ *and scalars* a, b*:*

(i) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$	(v) $a(\mathbf{u}+\mathbf{v})=a\mathbf{u}+a\mathbf{v}$
(ii) $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$	(vi) $(a+b)\mathbf{u}=a\mathbf{u}+b\mathbf{u}$
(iii) $\mathbf{u}+\mathbf{0}=\mathbf{u}$	(vii) $(ab)\mathbf{u}=a(b\mathbf{u})$
(iv) $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$	$(viii)$ $1\mathbf{u}=\mathbf{u}$

Moreover $|c\,\mathbf{u}|=|c|\,|\mathbf{u}|$, *and* $|\mathbf{u}|=0$ *if and only if* $\mathbf{u}=\mathbf{0}$.

Proof. Each identity reduces, component by component, to the corresponding property of real numbers. For instance, (v) holds because $a(u_i+v_i)=au_i+av_i$ for $i=1,2,3$. For the norm, $|c\mathbf{u}|^2=\sum(cu_i)^2=c^2\sum u_i^2$; take square roots. ∎

Properties (i)–(viii) are the *axioms of a vector space*. In Chapter 9 any set with operations satisfying them will be called a vector space, whether its elements are arrows, polynomials, matrices, or wavefunctions.

1.2.1 BASES AND COMPONENTS

The *standard basis vectors* are $\hat{\mathbf{i}}=\langle 1,0,0\rangle$, $\hat{\mathbf{j}}=\langle 0,1,0\rangle$, and $\hat{\mathbf{k}}=\langle 0,0,1\rangle$, also written $\hat{\mathbf{e}}_1,\hat{\mathbf{e}}_2,\hat{\mathbf{e}}_3$. Every vector is uniquely a *linear combination* of them:

$$\langle v_1,v_2,v_3\rangle=v_1\hat{\mathbf{i}}+v_2\hat{\mathbf{j}}+v_3\hat{\mathbf{k}}=\sum_{i=1}^{3}v_i\,\hat{\mathbf{e}}_i.$$

For any nonzero $\mathbf{v}$, the vector $\hat{\mathbf{v}}=\mathbf{v}/|\mathbf{v}|$ is the unit vector in the direction of $\mathbf{v}$; the process of dividing by the norm is called *normalization*. Every nonzero vector factors as magnitude times direction: $\mathbf{v}=|\mathbf{v}|\,\hat{\mathbf{v}}$. In the plane, a vector of length r making angle θ with the positive x-axis is $r\,\langle\cos\theta,\sin\theta\rangle$.

Example 1.7 (Resultant of forces). Forces of 50 N at 30° and 30 N at 120° (measured from the positive x-axis) act on a particle. Find the resultant.

Solution. In components,

$$\mathbf{F}_1=50\,\langle\cos 30^\circ,\sin 30^\circ\rangle=\left\langle 25\sqrt{3},\,25\right\rangle,$$
$$\mathbf{F}_2=30\,\langle\cos 120^\circ,\sin 120^\circ\rangle=\left\langle -15,\,15\sqrt{3}\right\rangle.$$

So $\mathbf{F}=\mathbf{F}_1+\mathbf{F}_2=\left\langle 25\sqrt{3}-15,\;25+15\sqrt{3}\right\rangle\approx\langle 28.30,\,50.98\rangle$. Because the two forces are perpendicular, $|\mathbf{F}|=\sqrt{50^2+30^2}=10\sqrt{34}\approx 58.31$ N. Its direction is $\tan^{-1}(50.98/28.30)\approx 60.96^\circ$.

Example 1.8 (Dividing a segment). Find the point R that is two-thirds of the way from $P(1,-2,4)$ to $Q(7,4,-2)$.

Solution. The position vector of R is $\mathbf{r} = \mathbf{p} + \frac{2}{3}(\mathbf{q}-\mathbf{p}) = \frac{1}{3}\mathbf{p} + \frac{2}{3}\mathbf{q}$. Hence $R = \left(\frac{1}{3}+\frac{14}{3}, -\frac{2}{3}+\frac{8}{3}, \frac{4}{3}-\frac{4}{3}\right) = (5,2,0)$. More generally, $(1-t)\mathbf{p} + t\mathbf{q}$ traces the segment PQ as t runs from 0 to 1.

1.3 The Dot Product

Definition 1.9 (Dot product). The *dot product* (or *scalar product*) of $\mathbf{u} = \langle u_1, u_2, u_3\rangle$ and $\mathbf{v} = \langle v_1, v_2, v_3\rangle$ is the scalar

$$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 = \sum_{i=1}^{3} u_iv_i.$$

Theorem 1.10 (Algebraic properties). *For vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ and scalar c:*

(*i*) $\mathbf{u}\cdot\mathbf{u} = |\mathbf{u}|^2 \geq 0$, *with equality only for* $\mathbf{u}=\mathbf{0}$ (*positive definiteness*)

(*ii*) $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$ (*symmetry*)

(*iii*) $\mathbf{u}\cdot(\mathbf{v}+\mathbf{w}) = \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}$ *and* $(c\mathbf{u})\cdot\mathbf{v} = c(\mathbf{u}\cdot\mathbf{v})$ (*linearity*)

Proof. All three follow directly from the component formula. For (i), $\mathbf{u}\cdot\mathbf{u} = \sum u_i^2$, a sum of squares that vanishes only when every $u_i = 0$. ∎

These three properties are exactly what an *inner product* must satisfy (Chapter 9). Everything that follows in this section uses only them, which is why it will generalize.

Theorem 1.11 (Geometric form of the dot product). *If $\theta \in [0,\pi]$ is the angle between nonzero vectors $\mathbf{u}$ and $\mathbf{v}$, then*

$$\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}|\,|\mathbf{v}|\cos\theta.$$

Proof. The vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u}-\mathbf{v}$ form a triangle in which θ is the angle opposite the side $\mathbf{u}-\mathbf{v}$. By the Law of Cosines,

$$|\mathbf{u}-\mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2|\mathbf{u}||\mathbf{v}|\cos\theta.$$

By Theorem 1.10, the left side equals $(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) = |\mathbf{u}|^2 - 2\,\mathbf{u}\cdot\mathbf{v} + |\mathbf{v}|^2$. Comparing the two expressions gives the result. ∎

Corollary 1.12. *Nonzero vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *are perpendicular* (orthogonal) *if and only if* $\mathbf{u}\cdot\mathbf{v}=0$. *The angle between them is acute when* $\mathbf{u}\cdot\mathbf{v}>0$ *and obtuse when* $\mathbf{u}\cdot\mathbf{v}<0$.

By convention $\mathbf{0}$ is orthogonal to every vector. The geometric form suggests that $|\mathbf{u}\cdot\mathbf{v}|$ can never exceed $|\mathbf{u}||\mathbf{v}|$. The point of the next theorem is that this inequality can be proved from the algebraic properties alone, with no appeal to pictures. That is what lets it survive in spaces where pictures are unavailable.

Theorem 1.13 (Cauchy–Schwarz inequality). *For all vectors* $\mathbf{u},\mathbf{v}$,

$$|\mathbf{u}\cdot\mathbf{v}| \le |\mathbf{u}|\,|\mathbf{v}|,$$

with equality if and only if one of the vectors is a scalar multiple of the other.

Proof. If $\mathbf{v}=\mathbf{0}$ both sides vanish. Otherwise, for every real t, positive definiteness gives

$$0 \le |\mathbf{u}-t\mathbf{v}|^2 = |\mathbf{v}|^2 t^2 - 2(\mathbf{u}\cdot\mathbf{v})\,t + |\mathbf{u}|^2.$$

A quadratic in t with positive leading coefficient that is never negative has discriminant at most zero: $4(\mathbf{u}\cdot\mathbf{v})^2 - 4|\mathbf{u}|^2|\mathbf{v}|^2 \le 0$, which is the inequality. Equality holds exactly when the quadratic has a real root t_0, that is, when $|\mathbf{u}-t_0\mathbf{v}|=0$, or $\mathbf{u}=t_0\mathbf{v}$. ∎

Corollary 1.14 (Triangle inequality). $|\mathbf{u}+\mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}|$.

Proof. $|\mathbf{u}+\mathbf{v}|^2 = |\mathbf{u}|^2 + 2\,\mathbf{u}\cdot\mathbf{v} + |\mathbf{v}|^2 \le |\mathbf{u}|^2 + 2|\mathbf{u}||\mathbf{v}| + |\mathbf{v}|^2 = (|\mathbf{u}|+|\mathbf{v}|)^2$, using Theorem 1.13. Take square roots. ∎

Example 1.15 (Angle between vectors). Find the angle between $\mathbf{u} = \langle 1,2,2\rangle$ and $\mathbf{v}=\langle 3,0,4\rangle$.

Solution. $\mathbf{u}\cdot\mathbf{v} = 3+0+8 = 11$, $|\mathbf{u}|=3$, $|\mathbf{v}|=5$, so $\cos\theta = 11/15$ and $\theta = \cos^{-1}(11/15) \approx 42.83^\circ$.

Example 1.16 (Angle in a cube). Find the angle between a main diagonal of a cube and one of its edges, and between two main diagonals.

Solution. Take the unit cube with a corner at the origin. A diagonal is $\langle 1,1,1\rangle$ and an edge is $\langle 1,0,0\rangle$, so $\cos\theta = 1/\sqrt{3}$ and $\theta \approx 54.74^\circ$. Two diagonals are $\langle 1,1,1\rangle$ and $\langle 1,1,-1\rangle$, giving $\cos\theta = 1/3$ and $\theta \approx 70.53^\circ$. The second angle is the tetrahedral bond angle's supplement: $180^\circ - 70.53^\circ = 109.47^\circ$ is the H–C–H angle in methane.

1.3.1 DIRECTION COSINES

The angles α, β, γ that a nonzero vector $\mathbf{v}$ makes with $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ are its *direction angles*. By Theorem 1.11, $\cos\alpha = v_1/|\mathbf{v}|$, $\cos\beta = v_2/|\mathbf{v}|$, and $\cos\gamma = v_3/|\mathbf{v}|$. These are the components of $\hat{\mathbf{v}}$, so

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1.$$

1.3.2 PROJECTIONS AND ORTHOGONAL DECOMPOSITION

Definition 1.17 (Projection). Let $\mathbf{v} \neq \mathbf{0}$. The *scalar component* of $\mathbf{u}$ along $\mathbf{v}$ and the *vector projection* of $\mathbf{u}$ onto $\mathbf{v}$ are

$$\operatorname{comp}_{\mathbf{v}}\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|} = \mathbf{u}\cdot\hat{\mathbf{v}}, \qquad \operatorname{proj}_{\mathbf{v}}\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|^2}\,\mathbf{v} = (\mathbf{u}\cdot\hat{\mathbf{v}})\,\hat{\mathbf{v}}.$$

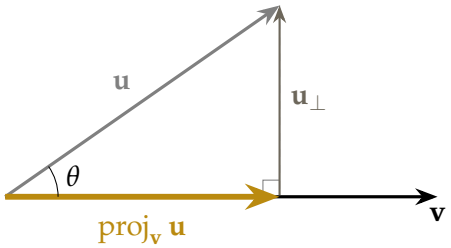


Figure 1.3: Orthogonal decomposition $\mathbf{u} = \operatorname{proj}_{\mathbf{v}}\mathbf{u} + \mathbf{u}_\perp$.

Theorem 1.18 (Orthogonal decomposition). *For $\mathbf{v} \neq \mathbf{0}$, every vector $\mathbf{u}$ can be written uniquely as $\mathbf{u} = \mathbf{u}_\parallel + \mathbf{u}_\perp$ with $\mathbf{u}_\parallel$ a multiple of $\mathbf{v}$ and $\mathbf{u}_\perp \cdot \mathbf{v} = 0$. Namely $\mathbf{u}_\parallel = \operatorname{proj}_{\mathbf{v}}\mathbf{u}$ and $\mathbf{u}_\perp = \mathbf{u} - \operatorname{proj}_{\mathbf{v}}\mathbf{u}$. Moreover $|\mathbf{u}|^2 = |\mathbf{u}_\parallel|^2 + |\mathbf{u}_\perp|^2$.*

Proof. Suppose $\mathbf{u} = c\mathbf{v} + \mathbf{w}$ with $\mathbf{w}\cdot\mathbf{v} = 0$. Dotting with $\mathbf{v}$ gives $\mathbf{u}\cdot\mathbf{v} = c|\mathbf{v}|^2$,

which forces $c = (\mathbf{u}\cdot\mathbf{v})/|\mathbf{v}|^2$. This proves uniqueness, and the same c shows existence, since then $(\mathbf{u}-c\mathbf{v})\cdot\mathbf{v}=0$. The norm identity follows by expanding $(\mathbf{u}_\parallel+\mathbf{u}_\perp)\cdot(\mathbf{u}_\parallel+\mathbf{u}_\perp)$; the cross terms vanish. ■

Example 1.19 (Projection). Decompose $\mathbf{u}=\langle 4,1,3\rangle$ into components parallel and perpendicular to $\mathbf{v}=\langle 1,2,2\rangle$.

Solution. $\mathbf{u}\cdot\mathbf{v}=12$ and $|\mathbf{v}|^2=9$, so $\mathbf{u}_\parallel=\frac{4}{3}\langle 1,2,2\rangle=\langle\frac{4}{3},\frac{8}{3},\frac{8}{3}\rangle$ and $\mathbf{u}_\perp=\langle\frac{8}{3},-\frac{5}{3},\frac{1}{3}\rangle$. Check: $\mathbf{u}_\perp\cdot\mathbf{v}=\frac{8}{3}-\frac{10}{3}+\frac{2}{3}=0$, and $|\mathbf{u}_\parallel|^2+|\mathbf{u}_\perp|^2=16+10=26=|\mathbf{u}|^2$.

COMMON ERROR. The projection divides by $|\mathbf{v}|^2$, not $|\mathbf{v}|$. A quick test: replacing $\mathbf{v}$ by $2\mathbf{v}$ describes the same line, so $\text{proj}_{\mathbf{v}}\,\mathbf{u}$ must not change. Only the $|\mathbf{v}|^2$ version passes. Dividing by $|\mathbf{v}|$ once gives the scalar component, a number rather than a vector.

1.3.3 WORK

A constant force $\mathbf{F}$ that moves an object along a straight displacement $\mathbf{d}$ does work $W=\mathbf{F}\cdot\mathbf{d}=|\mathbf{F}||\mathbf{d}|\cos\theta$. Only the component of force along the motion does work. In Chapter 6 this becomes the line integral $W=\int_C \mathbf{F}\cdot d\mathbf{r}$ for variable forces along curved paths.

Example 1.20 (Work along a ramp). A 40 N force with direction $\langle 2,1,2\rangle$ pushes a crate from $(0,0,0)$ to $(3,4,0)$, in meters. Find the work.

Solution. $\mathbf{F}=40\cdot\frac{1}{3}\langle 2,1,2\rangle$ and $\mathbf{d}=\langle 3,4,0\rangle$, so $W=\frac{40}{3}(6+4+0)=\frac{400}{3}\approx 133.3\,\text{J}$.

1.4 THE CROSS PRODUCT

The dot product turns two vectors into a scalar. In three dimensions, and only there, there is also a product that turns two vectors into a third vector perpendicular to both.

Definition 1.21 (Cross product). For $\mathbf{u} = \langle u_1, u_2, u_3 \rangle$ and $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$,

$$\mathbf{u} \times \mathbf{v} = \langle u_2 v_3 - u_3 v_2,\ u_3 v_1 - u_1 v_3,\ u_1 v_2 - u_2 v_1 \rangle = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix},$$

where the determinant is expanded formally along its first row.

Theorem 1.22 (Properties of the cross product). *For vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *and scalar* c*:*

(*i*) $\mathbf{v} \times \mathbf{u} = -\,\mathbf{u} \times \mathbf{v}$; *in particular* $\mathbf{u} \times \mathbf{u} = \mathbf{0}$.

(*ii*) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ *and* $(c\mathbf{u}) \times \mathbf{v} = c(\mathbf{u} \times \mathbf{v})$.

(*iii*) $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{u} = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{v} = 0$.

(*iv*) $|\mathbf{u} \times \mathbf{v}|^2 = |\mathbf{u}|^2 |\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2$ (*Lagrange's identity*)

(*v*) $|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}||\mathbf{v}| \sin\theta$, *where* $\theta \in [0, \pi]$ *is the angle between them.*

(*vi*) $\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}$, $\hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}}$, $\hat{\mathbf{k}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}}$.

(*vii*) *The cross product is* not *associative: in general* $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} \neq \mathbf{u} \times (\mathbf{v} \times \mathbf{w})$.

Proof. (i), (ii), and (vi) are immediate from the definition. For (iii),

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{u} = (u_2 v_3 - u_3 v_2) u_1 + (u_3 v_1 - u_1 v_3) u_2 + (u_1 v_2 - u_2 v_1) u_3 = 0,$$

since the six terms cancel in pairs; the computation with $\mathbf{v}$ is identical. Lagrange's identity (iv) is a direct expansion of both sides, which we leave as Item 1.40. Then (v) follows from (iv) and Theorem 1.11: $|\mathbf{u} \times \mathbf{v}|^2 = |\mathbf{u}|^2 |\mathbf{v}|^2 (1 - \cos^2\theta) = |\mathbf{u}|^2 |\mathbf{v}|^2 \sin^2\theta$, and $\sin\theta \geq 0$ on $[0, \pi]$. For (vii), $(\hat{\mathbf{i}} \times \hat{\mathbf{i}}) \times \hat{\mathbf{j}} = \mathbf{0}$ while $\hat{\mathbf{i}} \times (\hat{\mathbf{i}} \times \hat{\mathbf{j}}) = \hat{\mathbf{i}} \times \hat{\mathbf{k}} = -\hat{\mathbf{j}}$. ∎

Properties (iii) and (v) pin $\mathbf{u} \times \mathbf{v}$ down geometrically. It is perpendicular to both factors, and its length is the area of the parallelogram they span. Of the two perpendicular directions, it points along the one given by the right-hand rule: curl the fingers from $\mathbf{u}$ toward $\mathbf{v}$, and the thumb gives $\mathbf{u} \times \mathbf{v}$. Two nonzero vectors are parallel if and only if their cross product is $\mathbf{0}$.

Example 1.23 (Computing a cross product). $\langle 1, 2, 3 \rangle \times \langle 4, 5, 6 \rangle = \langle 2 \cdot 6 - 3 \cdot 5,\ 3 \cdot 4 - 1 \cdot 6,\ 1 \cdot 5 - 2 \cdot 4 \rangle = \langle -3, 6, -3 \rangle$. Check: $\langle -3, 6, -3 \rangle \cdot \langle 1, 2, 3 \rangle = -3 + 12 - 9 = 0$ and $\langle -3, 6, -3 \rangle \cdot \langle 4, 5, 6 \rangle = -12 + 30 - 18 = 0$.

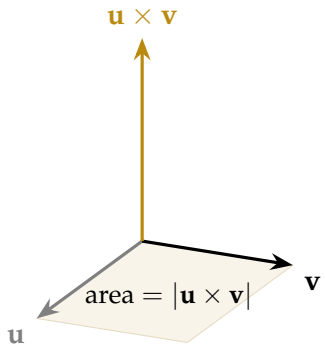


Figure 1.4: The cross product is normal to the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$, with length equal to its area.

COMMON ERROR (the middle component). In the determinant expansion, the $\hat{\mathbf{j}}$ term carries a minus sign: $-(u_1v_3 - u_3v_1)$. Dropping it is the most common cross-product error. Dotting the result with both factors, as in Example 1.23, catches it every time.

Example 1.24 (Area of a triangle). Find the area of the triangle with vertices $P(1,0,0)$, $Q(0,2,0)$, $R(0,0,3)$.

Solution. $\overrightarrow{PQ} \times \overrightarrow{PR} = \langle -1,2,0\rangle \times \langle -1,0,3\rangle = \langle 6,3,2\rangle$, with norm 7. The triangle is half the parallelogram, so its area is $7/2$.

1.4.1 TRIPLE PRODUCTS

Theorem 1.25 (Scalar triple product).

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix},$$

and its absolute value is the volume of the parallelepiped spanned by $\mathbf{u}, \mathbf{v}, \mathbf{w}$. *The product is unchanged by cyclic permutation,* $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})$, *and changes sign under any swap of two factors. Three vectors are coplanar if and only if their triple product is zero.*

Proof. The determinant formula is the cofactor expansion along the first row. For the volume, the base of the parallelepiped is the parallelogram spanned by $\mathbf{v}, \mathbf{w}$,

with area $|\mathbf{v}\times\mathbf{w}|$. Its height is the length of the projection of $\mathbf{u}$ onto the normal $\mathbf{v}\times\mathbf{w}$, namely $|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|/|\mathbf{v}\times\mathbf{w}|$. Multiply. The symmetry properties are those of determinants under row permutations. ∎

Theorem 1.26 (Vector triple product, "BAC–CAB"). $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}\,(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}\,(\mathbf{a}\cdot\mathbf{b})$.

We prove this below using index notation, where it takes two lines.

1.4.2 INDEX NOTATION

Writing every component formula out in full quickly becomes unmanageable. Physicists use *index notation*: a vector is named by its components u_i with $i\in\{1,2,3\}$, and two symbols do most of the work.

Definition 1.27 (Kronecker delta and Levi-Civita symbol).

$$\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}\qquad \varepsilon_{ijk}=\begin{cases}+1, & (i,j,k)\text{ an even permutation of }(1,2,3)\\ -1, & (i,j,k)\text{ an odd permutation of }(1,2,3)\\ 0, & \text{if any index repeats.}\end{cases}$$

So $\varepsilon_{123}=\varepsilon_{231}=\varepsilon_{312}=1$ and $\varepsilon_{132}=\varepsilon_{213}=\varepsilon_{321}=-1$. We adopt the *Einstein summation convention*: an index that appears twice in a term is summed from 1 to 3. With it,

$$\mathbf{u}\cdot\mathbf{v}=u_iv_i=\delta_{ij}u_iv_j,\qquad (\mathbf{u}\times\mathbf{v})_i=\varepsilon_{ijk}u_jv_k,\qquad \mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})=\varepsilon_{ijk}u_iv_jw_k.$$

For example, $(\mathbf{u}\times\mathbf{v})_1=\varepsilon_{123}u_2v_3+\varepsilon_{132}u_3v_2=u_2v_3-u_3v_2$, as it should.

Lemma 1.28 (The ε–δ identity). $\varepsilon_{ijk}\varepsilon_{imn}=\delta_{jm}\delta_{kn}-\delta_{jn}\delta_{km}$.

Proof. Both sides vanish unless $j\neq k$ and $\{j,k\}=\{m,n\}$. If $(m,n)=(j,k)$, the left side is $\varepsilon_{ijk}^2=1$, with i the remaining index, and the right side is $1-0=1$. If $(m,n)=(k,j)$, both sides equal -1. ∎

Proof of Theorem 1.26. Using cyclicity $\varepsilon_{kij}=\varepsilon_{ijk}$ and Lemma 1.28,

$$[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_i=\varepsilon_{ijk}a_j\varepsilon_{kmn}b_mc_n=(\delta_{im}\delta_{jn}-\delta_{in}\delta_{jm})a_jb_mc_n=b_i(a_jc_j)-c_i(a_jb_j).$$ ∎

LOOKING AHEAD. In Chapter 17, the components of orbital angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ become operators that fail to commute. Their commutation relations are written $[L_i, L_j] = i\hbar\,\varepsilon_{ijk} L_k$, which is exactly the index structure of the cross product. Fluency with ε_{ijk} now pays off directly there.

1.4.3 TORQUE AND ANGULAR MOMENTUM

If a force $\mathbf{F}$ acts at a point with position $\mathbf{r}$ relative to a pivot, the *torque* is $\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$. A particle of momentum $\mathbf{p} = m\mathbf{v}$ at position $\mathbf{r}$ has *angular momentum* $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Newton's second law gives $\mathrm{d}\mathbf{L}/\mathrm{d}t = \boldsymbol{\tau}$ (Chapter 2), so angular momentum is conserved under any central force, one parallel to $\mathbf{r}$.

Example 1.29 (Volume of a parallelepiped). The vectors $\mathbf{u} = \langle 1,1,0\rangle$, $\mathbf{v} = \langle 0,1,1\rangle$, $\mathbf{w} = \langle 1,0,1\rangle$ span a parallelepiped of volume $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| = |\,\langle 1,1,0\rangle \cdot \langle 1,1,-1\rangle\,| = 2$.

1.5 LINES AND PLANES

1.5.1 LINES

A line is determined by one point P_0 (position vector $\mathbf{r}_0$) and a nonzero *direction vector* $\mathbf{v} = \langle a,b,c\rangle$. A point lies on it exactly when its position vector has the form

$$\mathbf{r}(t) = \mathbf{r}_0 + t\mathbf{v}, \qquad t \in \mathbb{R}. \tag{1.2}$$

In components these are the *parametric equations* $x = x_0 + at$, $y = y_0 + bt$, $z = z_0 + ct$. If a, b, c are all nonzero, eliminating t gives the *symmetric equations*

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Two lines in space are *parallel* if their direction vectors are parallel. Otherwise they either *intersect* or are *skew*: not parallel, and yet never meeting.

Example 1.30 (Line through two points). Find the line through $P(1,-1,2)$ and $Q(3,2,-1)$ and the point where it crosses the xy-plane.

Solution. With $\mathbf{v} = \overrightarrow{PQ} = \langle 2,3,-3\rangle$: $x = 1 + 2t$, $y = -1 + 3t$, $z = 2 - 3t$. Setting $z = 0$ gives $t = 2/3$, so the crossing point is $(7/3,\ 1,\ 0)$.

1.5.2 PLANES

A plane is determined by a point P_0 in it and a nonzero *normal vector* $\mathbf{n} = \langle a, b, c \rangle$ perpendicular to it. The point $\mathbf{r}$ lies in the plane if and only if $\mathbf{r} - \mathbf{r}_0$ is orthogonal to $\mathbf{n}$:

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0 \iff a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \iff ax + by + cz = d, \tag{1.3}$$

where $d = \mathbf{n} \cdot \mathbf{r}_0$. The coefficients of x, y, z in a linear equation are always a normal vector. The angle between two planes is defined to be the angle between their normals, taken in $[0, \pi/2]$.

Example 1.31 (Plane through three points). Find the plane through $P(1,0,0)$, $Q(0,2,0)$, $R(0,0,3)$.

Solution. A normal is $\overrightarrow{PQ} \times \overrightarrow{PR} = \langle 6, 3, 2 \rangle$. Using P: $6(x-1) + 3y + 2z = 0$, so the plane is $6x + 3y + 2z = 6$. (Dividing by 6 gives the *intercept form* $x/1 + y/2 + z/3 = 1$.)

Example 1.32 (Line of intersection). Find the line of intersection of $x + y + z = 1$ and $x - 2y + 3z = 1$, and the angle between the planes.

Solution. The line lies in both planes, so it is perpendicular to both normals. Its direction is $\langle 1,1,1 \rangle \times \langle 1,-2,3 \rangle = \langle 5,-2,-3 \rangle$. To find a point, set $z = 0$: $x + y = 1$ and $x - 2y = 1$ give $(1,0,0)$. The line is $\mathbf{r}(t) = \langle 1,0,0 \rangle + t\langle 5,-2,-3 \rangle$. The normals satisfy $\cos\theta = |1 - 2 + 3|/(\sqrt{3}\sqrt{14}) = 2/\sqrt{42}$, so $\theta \approx 72.02^\circ$.

1.5.3 DISTANCES

Theorem 1.33 (Distance formulas). *(a) The distance from a point P_1 to the plane $ax + by + cz = d$ is*

$$D = \frac{|ax_1 + by_1 + cz_1 - d|}{\sqrt{a^2 + b^2 + c^2}}.$$

(b) The distance from P_1 to the line through P_0 with direction $\mathbf{v}$ is

$$D = \frac{|\overrightarrow{P_0P_1} \times \mathbf{v}|}{|\mathbf{v}|}.$$

(c) If the lines $\mathbf{r}_0 + t\mathbf{v}_1$ and $\mathbf{s}_0 + s\mathbf{v}_2$ are not parallel, the distance between them

is

$$D = \frac{|(\mathbf{s}_0 - \mathbf{r}_0) \cdot (\mathbf{v}_1 \times \mathbf{v}_2)|}{|\mathbf{v}_1 \times \mathbf{v}_2|}.$$

Proof. (a) Let P_0 be any point of the plane. The distance is the length of the projection of $\overrightarrow{P_0P_1}$ onto the normal: $|\overrightarrow{P_0P_1} \cdot \mathbf{n}|/|\mathbf{n}| = |\mathbf{n} \cdot \mathbf{r}_1 - \mathbf{n} \cdot \mathbf{r}_0|/|\mathbf{n}|$, and $\mathbf{n} \cdot \mathbf{r}_0 = d$. (b) The parallelogram spanned by $\overrightarrow{P_0P_1}$ and $\mathbf{v}$ has area $|\overrightarrow{P_0P_1} \times \mathbf{v}|$ and base $|\mathbf{v}|$; its height is the distance. (c) The vector $\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2$ is perpendicular to both lines. They therefore lie in parallel planes with normal $\mathbf{n}$, and the distance between those planes is the length of the projection of $\mathbf{s}_0 - \mathbf{r}_0$ onto $\mathbf{n}$. ∎

Example 1.34 (Distance to a plane). The distance from the origin to $6x + 3y + 2z = 6$ is $|-6|/\sqrt{49} = 6/7$.

1.6 Cylinders and Quadric Surfaces

The graph of an equation in x, y, z is generally a surface. Two families recur throughout this book.

A *cylinder* is a surface consisting of all lines (*rulings*) parallel to a fixed line and passing through a fixed plane curve. When one variable is missing from an equation, the surface is a cylinder whose rulings are parallel to that variable's axis. For example, $x^2 + y^2 = 1$ in $\mathbb{R}^3$ is a circular cylinder about the z-axis, and $z = y^2$ is a parabolic cylinder with rulings parallel to the x-axis.

A *quadric surface* is the graph of a second-degree equation in x, y, z. By rotating and translating axes (a process we justify with eigenvalues in Chapter 9), every nondegenerate quadric can be brought to one of the standard forms in Table 1.1. The tool for identifying them is the *trace*: the curve in which the surface meets a plane parallel to a coordinate plane.

Example 1.35 (Identifying a quadric). Identify $4x^2 - y^2 + 2z^2 + 4 = 0$.

Solution. Dividing by -4: $-x^2 + \frac{y^2}{4} - \frac{z^2}{2} = 1$. One positive squared term on the right-hand side of the standard form, with the other two negative, means a hyperboloid of two sheets. Here the axis is the y-axis. The traces $y = k$ are the ellipses $x^2 + \frac{z^2}{2} = \frac{k^2}{4} - 1$, which exist only for $|k| \geq 2$. The two sheets lie in $y \geq 2$ and $y \leq -2$.

Surface	Standard equation	Traces
Ellipsoid	$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$	all ellipses
Elliptic paraboloid	$\frac{z}{c}=\frac{x^2}{a^2}+\frac{y^2}{b^2}$	horizontal: ellipses; vertical: parabolas
Hyperbolic paraboloid	$\frac{z}{c}=\frac{x^2}{a^2}-\frac{y^2}{b^2}$	horizontal: hyperbolas; vertical: parabolas
Elliptic cone	$\frac{z^2}{c^2}=\frac{x^2}{a^2}+\frac{y^2}{b^2}$	horizontal: ellipses; vertical through axis: line pairs
Hyperboloid of one sheet	$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$	horizontal: ellipses; vertical: hyperbolas
Hyperboloid of two sheets	$-\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$	horizontal ($\lvert z\rvert > c$): ellipses; vertical: hyperbolas

Table 1.1: The nondegenerate quadric surfaces. The axis of symmetry is the z-axis.

Example 1.36 (Completing the square). Identify $x^2+2z^2-6x-y+10=0$.

Solution. $(x-3)^2+2z^2=y-1$. This is an elliptic paraboloid with vertex $(3,1,0)$ opening in the positive y-direction.

The hyperbolic paraboloid $z=x^2-y^2$, the "saddle," reappears in Chapter 4 as the model of a critical point that is neither a maximum nor a minimum.

1.7 Cylindrical and Spherical Coordinates

Problems with an axis of symmetry, or with spherical symmetry, are almost always simpler in coordinates adapted to that symmetry. The hydrogen atom of Chapter 16 is solvable precisely because spherical coordinates separate its Schrödinger equation.

1.7.1 Cylindrical Coordinates

A point P has *cylindrical coordinates* (ρ,ϕ,z), where (ρ,ϕ) are polar coordinates of the projection of P onto the xy-plane and z is unchanged:

$$x=\rho\cos\phi,\quad y=\rho\sin\phi,\quad z=z;\qquad \rho=\sqrt{x^2+y^2},\quad \tan\phi=\frac{y}{x}.$$

The quadrant of (x,y) determines ϕ; the formula $\tan\phi=y/x$ alone does not.

1.7.2 SPHERICAL COORDINATES

Definition 1.37 (Spherical coordinates, physics convention). The *spherical coordinates* (r, θ, ϕ) of a point P are: $r = |\overrightarrow{OP}| \geq 0$; the *polar angle* $\theta \in [0, \pi]$ between $\overrightarrow{OP}$ and the positive z-axis; and the *azimuthal angle* $\phi \in [0, 2\pi)$, the same angle as in cylindrical coordinates. Then

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta,$$

and conversely $r = \sqrt{x^2+y^2+z^2}$, $\cos\theta = z/r$.

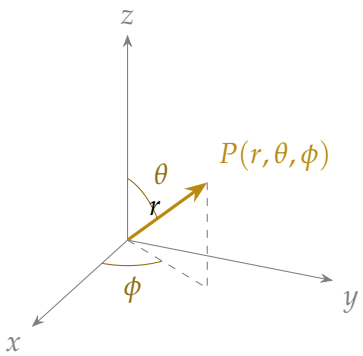


Figure 1.5: Spherical coordinates in the physics convention: θ from the z-axis, ϕ around it.

COMMON ERROR (conventions). Most calculus texts call the polar angle ϕ and the azimuth θ, the reverse of this book. Physics, and all of quantum mechanics, uses θ for the polar angle. When reading other sources, check which convention is in force before using any formula.

Example 1.38 (Conversions). Convert $(x, y, z) = (1, 1, \sqrt{6})$ to cylindrical and spherical coordinates.

Solution. $\rho = \sqrt{2}$ and $\phi = \pi/4$, so the cylindrical coordinates are $(\sqrt{2}, \pi/4, \sqrt{6})$. For the spherical coordinates, $r = \sqrt{1+1+6} = 2\sqrt{2}$, $\cos\theta = \sqrt{6}/(2\sqrt{2}) = \sqrt{3}/2$ so $\theta = \pi/6$, and $\phi = \pi/4$. The result is $(2\sqrt{2}, \pi/6, \pi/4)$.

Example 1.39 (Describing surfaces). In spherical coordinates, $r = a$ is a sphere, $\theta = \pi/4$ is the upper half of the cone $z = \sqrt{x^2 + y^2}$, and $\phi = \pi/3$ is a vertical half-plane. The equation $r = 2\cos\theta$ becomes $r^2 = 2r\cos\theta$, or $x^2 + y^2 + z^2 = 2z$: the sphere of radius 1 centered at $(0, 0, 1)$.

At each point there are mutually orthogonal unit vectors pointing in the directions of increasing coordinates. In spherical coordinates these are

$$\begin{aligned}\hat{\mathbf{r}} &= \langle \sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta \rangle\,,\\ \hat{\boldsymbol{\theta}} &= \langle \cos\theta\cos\phi,\ \cos\theta\sin\phi,\ -\sin\theta \rangle\,,\\ \hat{\boldsymbol{\phi}} &= \langle -\sin\phi,\ \cos\phi,\ 0 \rangle\,.\end{aligned}$$

Unlike $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$, they change from point to point. That fact is the source of the extra terms in the gradient, divergence, and Laplacian in curvilinear coordinates (Chapter 7).

LOOKING AHEAD. A quantum state of a two-level system, such as an electron's spin, can be represented by a unit vector, and the Born rule says measurement probabilities are squared components along orthonormal directions. If $\boldsymbol{\psi} = \langle 3/5, 4/5 \rangle$, the two outcomes occur with probabilities $9/25$ and $16/25$, which sum to 1 because $|\boldsymbol{\psi}| = 1$. The real theory uses complex components (Chapter 8) and the inner product of Chapter 9. In Chapter 17 we will see that spin states are parameterized by exactly the two spherical angles (θ, ϕ) introduced here: the *Bloch sphere*.

1.8 Worked Problems

The problems below are solved in full. Each combines several ideas from the chapter, and the solutions show the reasoning a careful solver would record, not only the result. Attempt each one before reading its solution.

Problem 1.1 Distance between skew lines

Let L_1 be the line $\mathbf{r}_1(t) = \langle 1, 0, 2 \rangle + t\,\langle 1, 2, -1 \rangle$ and L_2 the line $\mathbf{r}_2(s) = \langle 0, 1, -1 \rangle + s\,\langle 2, 1, 1 \rangle$.

(a) Show that L_1 and L_2 are skew.

(b) Find the distance between them.

(c) Find the points $Q_1 \in L_1$ and $Q_2 \in L_2$ that realize this distance.

Solution. (a) The directions $\mathbf{d}_1 = \langle 1, 2, -1 \rangle$ and $\mathbf{d}_2 = \langle 2, 1, 1 \rangle$ are not parallel, so the lines are either intersecting or skew. They intersect only if $(1+t,\ 2t,\ 2-t) = (2s,\ 1+s,\ -1+s)$ for some t, s. The first two components give $t = 2s - 1$ and $2t = 1 + s$, so $4s - 2 = 1 + s$, $s = 1$, $t = 1$. The third component then requires $2 - 1 = -1 + 1$, which is false. The lines do not meet, so they are skew.

(b) The common perpendicular has direction

$$\mathbf{n} = \mathbf{d}_1 \times \mathbf{d}_2 = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{\mathbf{k}} \\ 1 & 2 & -1 \\ 2 & 1 & 1 \end{vmatrix} = \langle 3, -3, -3 \rangle .$$

Every vector joining a point of L_1 to a point of L_2 has the same component along $\mathbf{n}$, because moving along either line changes the joining vector by a multiple of $\mathbf{d}_1$ or $\mathbf{d}_2$, both orthogonal to $\mathbf{n}$. With $\mathbf{w} = \overrightarrow{P_1P_2} = \langle -1, 1, -3 \rangle$,

$$d = \frac{|\mathbf{w} \cdot \mathbf{n}|}{|\mathbf{n}|} = \frac{|-3 - 3 + 9|}{3\sqrt{3}} = \frac{1}{\sqrt{3}}.$$

(c) The segment Q_1Q_2 is the common perpendicular, so $\mathbf{r}_1(t) - \mathbf{r}_2(s)$ is orthogonal to both $\mathbf{d}_1$ and $\mathbf{d}_2$. With $\mathbf{r}_1(t) - \mathbf{r}_2(s) = (1 + t - 2s,\ 2t - 1 - s,\ 3 - t - s)$, the two conditions are

$$(1 + t - 2s) + 2(2t - 1 - s) - (3 - t - s) = 6t - 3s - 4 = 0,$$
$$2(1 + t - 2s) + (2t - 1 - s) + (3 - t - s) = 3t - 6s + 4 = 0.$$

Hence $t = s = \frac{4}{3}$, and $Q_1 = \left(\frac{7}{3}, \frac{8}{3}, \frac{2}{3}\right)$, $Q_2 = \left(\frac{8}{3}, \frac{7}{3}, \frac{1}{3}\right)$. Check: $\overrightarrow{Q_2Q_1} = \frac{1}{3}\langle -1, 1, 1 \rangle = -\frac{1}{9}\mathbf{n}$, which is parallel to $\mathbf{n}$, with length $\frac{1}{\sqrt{3}}$, in agreement with (b).

Problem 1.2 A tetrahedron three ways

Let $A = (1, 1, 0)$, $B = (0, 2, 1)$, $C = (2, 0, 3)$, $D = (1, 3, 2)$. Find the volume of the tetrahedron $ABCD$, the area of face ABC, and the height of D above the plane ABC, and check that the three results are consistent.

Solution. Put the edge vectors from A into the columns of a determinant: $\mathbf{u} = \overrightarrow{AB} = \langle -1, 1, 1 \rangle$, $\mathbf{v} = \overrightarrow{AC} = \langle 1, -1, 3 \rangle$, $\mathbf{w} = \overrightarrow{AD} = \langle 0, 2, 2 \rangle$. The parallelepiped

they span has volume $|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|$, and a tetrahedron is one sixth of it:

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = \begin{vmatrix} -1 & 1 & 1 \\ 1 & -1 & 3 \\ 0 & 2 & 2 \end{vmatrix} = -1(-2-6) - 1(2-0) + 1(2-0) = 8,$$

$$V = \frac{8}{6} = \frac{4}{3}.$$

For the face, $\mathbf{u}\times\mathbf{v} = \langle 1\cdot 3 - 1\cdot(-1),\ 1\cdot 1 - (-1)\cdot 3,\ (-1)(-1) - 1\cdot 1\rangle = \langle 4,4,0\rangle$, so the area is $\frac{1}{2}|\mathbf{u}\times\mathbf{v}| = \frac{1}{2}\cdot 4\sqrt{2} = 2\sqrt{2}$. The height of D is the scalar component of $\mathbf{w}$ along the unit normal:

$$h = \frac{|\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})|}{|\mathbf{u}\times\mathbf{v}|} = \frac{|0+8+0|}{4\sqrt{2}} = \sqrt{2}.$$

Consistency: $V = \frac{1}{3}(\text{base area})(\text{height}) = \frac{1}{3}\cdot 2\sqrt{2}\cdot\sqrt{2} = \frac{4}{3}$. This is no coincidence. Expanding the triple product as $\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})$ makes the formula $\frac{1}{6}|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})| = \frac{1}{3}\cdot\frac{1}{2}|\mathbf{u}\times\mathbf{v}|\cdot h$ an identity, by the cyclic symmetry of the triple product.

Problem 1.3 The BAC–CAB rule and Lagrange's identity

(a) Using $\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$, prove $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})$.

(b) Deduce Lagrange's identity $|\mathbf{a}\times\mathbf{b}|^2 = |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a}\cdot\mathbf{b})^2$, and use it to rederive $|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta$.

(c) Prove the Jacobi identity $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) + \mathbf{b}\times(\mathbf{c}\times\mathbf{a}) + \mathbf{c}\times(\mathbf{a}\times\mathbf{b}) = \mathbf{0}$.

Solution. (a) In components, with summation over repeated indices,

$$\begin{aligned}[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_i &= \varepsilon_{ijk}a_j(\mathbf{b}\times\mathbf{c})_k = \varepsilon_{ijk}\varepsilon_{klm}a_j b_l c_m \\ &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})a_j b_l c_m = b_i(a_j c_j) - c_i(a_j b_j).\end{aligned}$$

This is the ith component of $\mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})$.

(b) By the cyclic property of the triple product, $|\mathbf{a}\times\mathbf{b}|^2 = (\mathbf{a}\times\mathbf{b})\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{a}\cdot\big(\mathbf{b}\times(\mathbf{a}\times\mathbf{b})\big)$. By (a), $\mathbf{b}\times(\mathbf{a}\times\mathbf{b}) = \mathbf{a}(\mathbf{b}\cdot\mathbf{b}) - \mathbf{b}(\mathbf{b}\cdot\mathbf{a})$. Dotting with $\mathbf{a}$ gives $|\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a}\cdot\mathbf{b})^2$. Since $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta$, the right side is $|\mathbf{a}|^2|\mathbf{b}|^2(1-\cos^2\theta)$, and $|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta$ because $\sin\theta \geq 0$ for $\theta\in[0,\pi]$. Note the logic: the magnitude formula for the cross product, which might seem to be part of the definition, is a consequence of the component formula.

(c) Apply (a) to each term:

$$\big[\mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})\big] + \big[\mathbf{c}(\mathbf{b}\cdot\mathbf{a}) - \mathbf{a}(\mathbf{b}\cdot\mathbf{c})\big] + \big[\mathbf{a}(\mathbf{c}\cdot\mathbf{b}) - \mathbf{b}(\mathbf{c}\cdot\mathbf{a})\big] = \mathbf{0},$$

since the six terms cancel in pairs. The Jacobi identity says the cross product is not associative, but its failure to be associative is controlled. The same identity, for commutators $[\hat{A},[\hat{B},\hat{C}]]+\text{cyclic}=0$, governs the angular momentum operators of Chapter 17, whose commutators reproduce the cross product.

Problem 1.4 Two planes, their line, and a reflection

Let $\Pi_1: x+2y-z=4$ and $\Pi_2: 2x-y+z=1$.

(a) Find parametric equations for the line $\Pi_1\cap\Pi_2$.

(b) Find the angle between the planes.

(c) Find the reflection P' of $P=(3,1,2)$ across Π_1.

Solution. (a) The line lies in both planes, so it is orthogonal to both normals $\mathbf{n}_1=\langle 1,2,-1\rangle$ and $\mathbf{n}_2=\langle 2,-1,1\rangle$. Its direction is

$$\mathbf{n}_1\times\mathbf{n}_2=\langle 2\cdot 1-(-1)(-1),\ (-1)\cdot 2-1\cdot 1,\ 1\cdot(-1)-2\cdot 2\rangle=\langle 1,-3,-5\rangle\,.$$

For a point on the line, the direction has nonzero z-component, so the line meets the plane $z=0$. There $x+2y=4$ and $2x-y=1$, so $x=\frac{6}{5}$, $y=\frac{7}{5}$. The line is $\mathbf{r}(t)=\left(\frac{6}{5}+t,\ \frac{7}{5}-3t,\ -5t\right)$. Check with $t=1$: $(\frac{11}{5},-\frac{8}{5},-5)$ gives $\frac{11}{5}-\frac{16}{5}+5=4$ and $\frac{22}{5}+\frac{8}{5}-5=1$.

(b) The angle between planes is the acute angle between their normals: $\cos\theta=\frac{|\mathbf{n}_1\cdot\mathbf{n}_2|}{|\mathbf{n}_1||\mathbf{n}_2|}=\frac{|2-2-1|}{\sqrt{6}\sqrt{6}}=\frac{1}{6}$, so $\theta=\arccos\frac{1}{6}\approx 80.4^\circ$.

(c) Write Π_1 as $\mathbf{n}_1\cdot\mathbf{x}=4$. The signed quantity $f(P)=\mathbf{n}_1\cdot P-4=3+2-2-4=-1$ measures how far P is from the plane in units of $|\mathbf{n}_1|$. To reflect, move P along $\mathbf{n}_1$ twice the distance to the plane:

$$P'=P-\frac{2f(P)}{|\mathbf{n}_1|^2}\mathbf{n}_1=(3,1,2)+\frac{2}{6}\langle 1,2,-1\rangle=\left(\frac{10}{3},\frac{5}{3},\frac{5}{3}\right).$$

Two checks: $\overrightarrow{PP'}=\frac{1}{3}\langle 1,2,-1\rangle$ is parallel to $\mathbf{n}_1$, and the midpoint $\left(\frac{19}{6},\frac{4}{3},\frac{11}{6}\right)$ satisfies $\frac{19}{6}+\frac{16}{6}-\frac{11}{6}=4$, so it lies on Π_1.

Problem 1.5 Identifying a quadric

Identify the surface $x^2-2y^2+z^2+4x+8y-2z=0$. Find its center, axis, and any vertices, and describe its cross-sections perpendicular to the axis.

Solution. Complete the square in each variable:

$$(x+2)^2 - 4 - 2\big[(y-2)^2 - 4\big] + (z-1)^2 - 1 = 0,$$
$$2(y-2)^2 - (x+2)^2 - (z-1)^2 = 3.$$

With $X = x+2$, $Y = y-2$, $Z = z-1$, this is $\dfrac{Y^2}{3/2} - \dfrac{X^2}{3} - \dfrac{Z^2}{3} = 1$: one positive square, two negative, and a positive constant. That is a *hyperboloid of two sheets*, centered at $(-2, 2, 1)$, with axis the line through the center parallel to the y-axis.

The cross-section $Y = c$ is $X^2 + Z^2 = 2c^2 - 3$. It is empty for $|c| < \sqrt{3/2}$, a single point for $|c| = \sqrt{3/2}$, and a circle of radius $\sqrt{2c^2 - 3}$ for $|c| > \sqrt{3/2}$. So the vertices are $(-2,\, 2 \pm \sqrt{3/2},\, 1)$, and the sheets open away from the center along the axis. The circular sections reflect the equal coefficients of X^2 and Z^2. The surface is a surface of revolution about its axis. For large $|Y|$ it approaches the cone $2Y^2 = X^2 + Z^2$ with half-angle $\arctan\sqrt{2}$.

Problem 1.6 When are three vectors coplanar?

For which real k are $\mathbf{a} = \langle 1, 2, k\rangle$, $\mathbf{b} = \langle 2, 1, 1\rangle$, $\mathbf{c} = \langle k, 1, 1\rangle$ linearly dependent? For the other values, is the ordered triple $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ right-handed or left-handed?

Solution. Three vectors are coplanar exactly when the parallelepiped they span has zero volume, that is, when their triple product vanishes. Expanding along the first row,

$$\begin{vmatrix} 1 & 2 & k \\ 2 & 1 & 1 \\ k & 1 & 1 \end{vmatrix} = 1(1-1) - 2(2-k) + k(2-k) = -k^2 + 4k - 4 = -(k-2)^2.$$

So the vectors are dependent only for $k = 2$. Check: then $\mathbf{a} = \langle 1, 2, 2\rangle$, $\mathbf{c} = \langle 2, 1, 1\rangle = \mathbf{b}$, so $\mathbf{b}$ and $\mathbf{c}$ coincide.

For $k \neq 2$ the triple product is strictly negative, so the triple is always left-handed. This is worth noticing: the determinant touches zero at $k = 2$ but does not change sign. As k passes through 2 the three vectors become coplanar for an instant and then separate again on the same side. A sign change would require the vectors to pass *through* coplanarity, which happens at a simple root, not a double one.

Problem 1.7 Spherical coordinates and a curve on a sphere

(a) Find spherical coordinates (r, θ, ϕ) for the point $(1, -1, \sqrt{2})$.

(b) Show that $r = 4\sin\theta\cos\phi$ is a sphere, and find its center and radius.

(c) Describe the curve in which this sphere meets the cone $\theta = \pi/4$, and find its projection onto the xy-plane.

Solution. (a) $r = \sqrt{1+1+2} = 2$. Then $\cos\theta = z/r = \sqrt{2}/2$, so $\theta = \pi/4$. The point $(x, y) = (1, -1)$ is in the fourth quadrant, so $\phi = 7\pi/4$ (with the convention $\phi \in [0, 2\pi)$). Note that $\tan\phi = y/x = -1$ alone cannot distinguish $(1, -1)$ from $(-1, 1)$. The quadrant must be read off separately.

(b) Multiply by r: $r^2 = 4r\sin\theta\cos\phi$, that is, $x^2 + y^2 + z^2 = 4x$, or $(x-2)^2 + y^2 + z^2 = 4$. This is the sphere of radius 2 centered at $(2, 0, 0)$, which passes through the origin. Multiplying by r could only add the origin, which is already on the surface, so nothing spurious was introduced. Note that $r \geq 0$ forces $\cos\phi \geq 0$: the sphere lies in the half-space $x \geq 0$.

(c) On the cone $\theta = \pi/4$ we have $z^2 = x^2 + y^2$ with $z \geq 0$. Substituting into $x^2 + y^2 + z^2 = 4x$ gives $2(x^2 + y^2) = 4x$, that is,

$$(x-1)^2 + y^2 = 1.$$

So the curve lies on the vertical cylinder over the circle of radius 1 centered at $(1, 0)$, and its projection onto the xy-plane is that circle. In parametric form, with $\phi \in [-\pi/2, \pi/2]$, $r = 2\sqrt{2}\cos\phi$ and

$$\mathbf{r}(\phi) = \langle 2\cos^2\phi,\ 2\sin\phi\cos\phi,\ 2\cos\phi\rangle = \langle 1 + \cos 2\phi,\ \sin 2\phi,\ 2\cos\phi\rangle\,.$$

The curve passes through the origin (at $\phi = \pm\pi/2$) and reaches its highest point $(2, 0, 2)$ at $\phi = 0$. It is a closed curve on the sphere, the intersection of two quadrics: a sphere and a cone, or equivalently a sphere and a cylinder.

EXERCISES

A. COMPUTATION

1.1 Find the distance between $A(-2, 3, 1)$ and $B(4, -1, 13)$.

1.2 Show that the triangle with vertices $(1, 2, 3)$, $(3, 4, 4)$, $(3, 0, 4)$ is isosceles, and find its perimeter.

1.3 Find the center and radius of the sphere $x^2 + y^2 + z^2 + 8x - 2y + 6z + 1 = 0$.

1.4 Find an equation of the sphere that has the segment from $(2, 1, 4)$ to $(4, 3, 10)$ as a diameter.

1.5 Find $\overrightarrow{PQ}$, its norm, and the unit vector in its direction, for $P(3, -2, 5)$ and $Q(-1, 2, 7)$.

1.6 For $\mathbf{u} = \langle 2, -1, 3\rangle$ and $\mathbf{v} = \langle -1, 4, 2\rangle$, compute $3\mathbf{u} - 2\mathbf{v}$ and $|\mathbf{u} + \mathbf{v}|$.

1.7 A plane flies at 200 mph on a bearing of 30° (clockwise from north). A wind blows due east at 40 mph. Find the ground speed and the true bearing.

1.8 Find the point one quarter of the way from $(-3, 5, 1)$ to $(5, 1, 9)$.

1.9 Find the angle between $\mathbf{u} = \langle 1, 1, 0\rangle$ and $\mathbf{v} = \langle 1, 0, 1\rangle$.

1.10 Find the angle between $\mathbf{u} = \langle 2, 2, 1\rangle$ and $\mathbf{v} = \langle 1, -2, 2\rangle$.

1.11 Find all k for which $\langle k, 3, -2\rangle$ is orthogonal to $\langle 4, k, 5\rangle$.

1.12 Find the direction cosines of $\langle 2, -3, 6\rangle$ and verify that their squares sum to 1.

1.13 Decompose $\mathbf{u} = \langle 2, 3, -1\rangle$ into parts parallel and perpendicular to $\mathbf{v} = \langle 1, 1, 1\rangle$.

1.14 Find $\operatorname{comp}_{\mathbf{v}} \mathbf{u}$ and $\operatorname{proj}_{\mathbf{v}} \mathbf{u}$ for $\mathbf{u} = \langle 3, -1, 4\rangle$ and $\mathbf{v} = \langle 2, 2, 1\rangle$.

1.15 A force $\mathbf{F} = \langle 3, -2, 5\rangle$ newtons moves an object from $(1, 0, 2)$ to $(4, 4, 3)$, in meters. Find the work done.

1.16 A sled is pulled 25 m by a rope at 40° above horizontal with tension 60 N. Find the work done, to the nearest joule.

1.17 Compute $\langle 2, -1, 3\rangle \times \langle 1, 4, -2\rangle$ and verify that the result is orthogonal to both factors.

1.18 Compute $(\hat{\mathbf{i}} + \hat{\mathbf{j}}) \times (\hat{\mathbf{i}} - \hat{\mathbf{j}})$ using only Theorem 1.22.

1.19 Find the area of the triangle with vertices $A(1, 1, 1)$, $B(3, 2, 1)$, $C(1, 3, 4)$.

1.20 Find a unit vector orthogonal to both $\langle 1, 2, -1\rangle$ and $\langle 3, 0, 1\rangle$.

1.21 Find the volume of the parallelepiped spanned by $\langle 1, 2, 0\rangle$, $\langle 0, 1, 3\rangle$, and $\langle 2, 0, 1\rangle$.

1.22 Determine whether $\langle 1, 2, 3\rangle$, $\langle 4, 5, 6\rangle$, and $\langle 7, 8, 9\rangle$ are coplanar.

1.23 A wrench 0.3 m long lies along the positive x-axis. A force $\mathbf{F} = \langle 0, 40, -30\rangle$ newtons is applied at its end. Find the torque about the origin and its magnitude.

1.24 Find parametric equations of the line through $(2, -1, 4)$ parallel to $\langle 3, 0, -2\rangle$, and the point where it meets the xy-plane.

1.25 Find symmetric equations of the line through $(1,2,3)$ and $(4,0,7)$.

1.26 Determine whether the lines $\langle 1,1,0\rangle + t\,\langle 1,-1,2\rangle$ and $\langle 2,0,2\rangle + s\,\langle 1,1,0\rangle$ are parallel, intersecting, or skew. If they intersect, find the point.

1.27 Find the plane through $(1,2,3)$ with normal $\langle 2,-1,4\rangle$.

1.28 Find the plane through $A(1,0,1)$, $B(2,1,0)$, $C(0,2,3)$.

1.29 Find the plane through $(1,-1,2)$ parallel to $3x - y + 2z = 7$.

1.30 Find the distance from $(3,1,-2)$ to the plane $2x - y + 2z = 4$.

1.31 Find the distance between the parallel planes $x + 2y - 2z = 3$ and $x + 2y - 2z = 12$.

1.32 Find the distance from $(1,2,3)$ to the line $\langle 0,0,0\rangle + t\,\langle 1,1,1\rangle$.

1.33 Find the line of intersection of $x - y + z = 2$ and $2x + y - z = 1$.

1.34 Find the angle between the planes $x + y = 1$ and $y + z = 2$.

1.35 Identify and sketch $x^2 + 4y^2 - z^2 = 4$. Describe its traces in $z = 0$, $z = 2$, and $x = 0$.

1.36 Identify $z = 4 - x^2 - y^2$ and $9x^2 - y^2 + z^2 = 0$.

1.37 Convert $(x,y,z) = (-1,\sqrt{3},2)$ to cylindrical and to spherical coordinates (physics convention).

1.38 Convert the spherical point $(r,\theta,\phi) = (4,\pi/3,3\pi/4)$ to Cartesian coordinates.

1.39 Write $x^2 + y^2 + z^2 = 4z$ in spherical coordinates and $z = x^2 + y^2$ in cylindrical coordinates.

B. THEORY AND PROOF

1.40 Prove Lagrange's identity $|\mathbf{u}\times\mathbf{v}|^2 + (\mathbf{u}\cdot\mathbf{v})^2 = |\mathbf{u}|^2|\mathbf{v}|^2$ (a) by expanding components and (b) using Lemma 1.28.

1.41 Prove the *parallelogram law* $|\mathbf{u}+\mathbf{v}|^2 + |\mathbf{u}-\mathbf{v}|^2 = 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2$ and the *polarization identity* $\mathbf{u}\cdot\mathbf{v} = \tfrac{1}{4}\big(|\mathbf{u}+\mathbf{v}|^2 - |\mathbf{u}-\mathbf{v}|^2\big)$. Interpret both geometrically.

1.42 Prove that the diagonals of a parallelogram are perpendicular if and only if the parallelogram is a rhombus.

1.43 Prove the reverse triangle inequality $\big||\mathbf{u}| - |\mathbf{v}|\big| \le |\mathbf{u}-\mathbf{v}|$.

1.44 Prove the Jacobi identity $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})+\mathbf{b}\times(\mathbf{c}\times\mathbf{a})+\mathbf{c}\times(\mathbf{a}\times\mathbf{b})=\mathbf{0}$.

1.45 Use index notation to prove $(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d})=(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})-(\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c})$.

1.46 Show that $\varepsilon_{ijk}\varepsilon_{ijk}=6$ and $\varepsilon_{ijk}\varepsilon_{ijn}=2\delta_{kn}$.

1.47 Prove that the spherical unit vectors $\hat{\mathbf{r}},\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\phi}}$ are mutually orthogonal, and that $\hat{\mathbf{r}}\times\hat{\boldsymbol{\theta}}=\hat{\boldsymbol{\phi}}$.

1.48 Prove that the distance formula of Theorem 1.33(c) gives zero exactly when the lines intersect.

C. CHALLENGE

1.49 Show that the lines $\langle 1,0,0\rangle+t\langle 1,1,0\rangle$ and $\langle 0,0,2\rangle+s\langle 0,1,1\rangle$ are skew, and find the distance between them. Then find the points on each line that achieve this distance.

1.50 A regular tetrahedron has vertices at $(1,1,1)$, $(1,-1,-1)$, $(-1,1,-1)$, $(-1,-1,1)$. Find its volume and the angle between two of its faces (the dihedral angle).

1.51 Let $\mathbf{n}$ be a unit vector. Show that reflection across the plane through the origin with normal $\mathbf{n}$ is $R(\mathbf{u})=\mathbf{u}-2(\mathbf{u}\cdot\mathbf{n})\mathbf{n}$, and that $|R(\mathbf{u})|=|\mathbf{u}|$ and $R(R(\mathbf{u}))=\mathbf{u}$.

1.52 (Rodrigues' formula.) Show that rotating $\mathbf{v}$ by angle α about a unit axis $\hat{\mathbf{k}}$ (right-hand sense) gives $\mathbf{v}_{\text{rot}}=\mathbf{v}\cos\alpha+(\hat{\mathbf{k}}\times\mathbf{v})\sin\alpha+\hat{\mathbf{k}}\,(\hat{\mathbf{k}}\cdot\mathbf{v})(1-\cos\alpha)$.

1.53 A two-state system has state $\boldsymbol{\psi}=\left\langle 1/\sqrt{3},\sqrt{2/3}\right\rangle$. (a) Verify it is normalized. (b) Find the probability of the first outcome. (c) Find a unit vector orthogonal to $\boldsymbol{\psi}$, and explain why its two outcome probabilities are the reverse of $\boldsymbol{\psi}$'s.

CHAPTER 2

VECTOR FUNCTIONS AND MOTION IN SPACE

LEARNING OBJECTIVES

- Define limits, continuity, derivatives, and integrals of vector-valued functions, and prove the product rules.
- Compute arc length and reparametrize a curve by arc length.
- Define curvature and torsion, construct the Frenet frame $\{\hat{\mathbf{T}}, \hat{\mathbf{N}}, \hat{\mathbf{B}}\}$, and prove the Frenet–Serret formulas.
- Decompose acceleration into tangential and normal components.
- Use vector calculus to analyze projectile motion and central forces, and derive Kepler's laws.

A particle moving through space traces a curve, and at each instant its position is a vector. Calculus applied to that vector gives velocity and acceleration, and Newton's second law $\mathbf{F} = m\mathbf{a}$ becomes a differential equation for a vector-valued function. This chapter develops the calculus of such functions and the geometry of the curves they trace. It ends with one of the great achievements of classical physics: the derivation of Kepler's laws of planetary motion from the inverse-square law of gravity. The breakdown of that classical picture at atomic scales is where quantum mechanics begins (Chapter 12).

2.1 VECTOR-VALUED FUNCTIONS

Definition 2.1 (Vector function). A *vector-valued function* on an interval $I \subseteq \mathbb{R}$ is a rule $\mathbf{r} : I \to \mathbb{R}^3$,

$$\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle = f(t)\,\hat{\mathbf{i}} + g(t)\,\hat{\mathbf{j}} + h(t)\,\hat{\mathbf{k}},$$

whose *component functions* f, g, h are real-valued. As t varies, the head of $\mathbf{r}(t)$ traces a *space curve* C; the equations $x = f(t)$, $y = g(t)$, $z = h(t)$ are *parametric equations* of C with *parameter* t.

The same curve has many parametrizations. The line through P_0 in the direction $\mathbf{v}$ is traced by $\mathbf{r}_0 + t\mathbf{v}$, but equally by $\mathbf{r}_0 + t^3\mathbf{v}$, which moves along it at a different rate. Properties of the curve itself, such as length and curvature, must turn out to be independent of the parametrization. Properties of the motion, such as speed, are not.

Definition 2.2 (Limit and continuity). $\lim_{t\to a} \mathbf{r}(t) = \mathbf{L}$ means $\lim_{t\to a} |\mathbf{r}(t) - \mathbf{L}| = 0$. The function $\mathbf{r}$ is *continuous* at a if $\lim_{t\to a} \mathbf{r}(t) = \mathbf{r}(a)$.

Theorem 2.3 (Limits are computed componentwise). *If* $\mathbf{r} = \langle f, g, h\rangle$ *and* $\mathbf{L} = \langle L_1, L_2, L_3\rangle$, *then* $\lim_{t\to a} \mathbf{r}(t) = \mathbf{L}$ *if and only if* $\lim_{t\to a} f(t) = L_1$, $\lim_{t\to a} g(t) = L_2$, *and* $\lim_{t\to a} h(t) = L_3$.

Proof. Each component satisfies $|f(t) - L_1| \le |\mathbf{r}(t) - \mathbf{L}|$, which gives one direction. Conversely, $|\mathbf{r}(t) - \mathbf{L}| \le |f(t) - L_1| + |g(t) - L_2| + |h(t) - L_3|$, by the triangle inequality applied to $\sum(\text{component})\,\hat{\mathbf{e}}_i$, which gives the other. ∎

Example 2.4 (The circular helix). The curve $\mathbf{r}(t) = \langle a\cos t,\ a\sin t,\ bt\rangle$ with $a > 0$, $b \neq 0$ lies on the cylinder $x^2 + y^2 = a^2$. Its height increases by $2\pi b$ in each full turn, a quantity called the *pitch*. It is the shape of a spring, a screw thread, and the DNA double helix.

Example 2.5 (Intersection of surfaces). Parametrize the curve in which the cylinder $x^2 + y^2 = 1$ meets the plane $y + z = 2$.

Solution. On the cylinder set $x = \cos t$, $y = \sin t$. The plane then forces $z = 2 - \sin t$. So $\mathbf{r}(t) = \langle \cos t, \sin t, 2 - \sin t\rangle$, $0 \le t \le 2\pi$. The curve is an ellipse: it is the intersection of a plane with a cylinder, tilted at 45°.

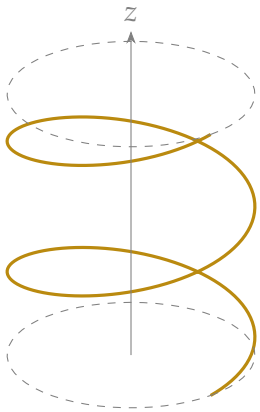


Figure 2.1: Two turns of the helix $\mathbf{r}(t) = \langle \cos t, \sin t, t/4 \rangle$ on the cylinder $x^2 + y^2 = 1$.

2.2 Derivatives and Integrals

Definition 2.6 (Derivative). The *derivative* of $\mathbf{r}$ at t is

$$\mathbf{r}'(t) = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \lim_{h \to 0} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h},$$

when the limit exists. If $\mathbf{r}'(t) \neq \mathbf{0}$, it is a *tangent vector* to the curve at $\mathbf{r}(t)$, and $\hat{\mathbf{T}}(t) = \mathbf{r}'(t)/|\mathbf{r}'(t)|$ is the *unit tangent vector*. A parametrization is *smooth* if $\mathbf{r}'$ is continuous and never zero.

By Theorem 2.3, $\mathbf{r}'(t) = \langle f'(t), g'(t), h'(t) \rangle$: vector functions are differentiated componentwise. The difference quotient is a secant vector scaled by $1/h$; as $h \to 0$ it turns into the tangent direction. A curve can have a corner where the parametrization is not smooth. For example, $\mathbf{r}(t) = \langle t^3, t^2 \rangle$ traces the cusp $y = x^{2/3}$, and $\mathbf{r}'(0) = \mathbf{0}$ there.

Theorem 2.7 (Differentiation rules). *Let $\mathbf{u}, \mathbf{v}$ be differentiable vector functions, f a differentiable real function, and c a scalar. Then*

(*i*) $(\mathbf{u} + \mathbf{v})' = \mathbf{u}' + \mathbf{v}'$ *and* $(c\mathbf{u})' = c\mathbf{u}'$;

(*ii*) $(f\mathbf{u})' = f'\mathbf{u} + f\mathbf{u}'$;

(*iii*) $(\mathbf{u} \cdot \mathbf{v})' = \mathbf{u}' \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{v}'$;

(*iv*) $(\mathbf{u} \times \mathbf{v})' = \mathbf{u}' \times \mathbf{v} + \mathbf{u} \times \mathbf{v}'$ (*the order of the factors matters*);

(*v*) $\big(\mathbf{u}(f(t))\big)' = f'(t)\,\mathbf{u}'(f(t))$ (*chain rule*).

Proof. We prove (iii); the others are similar. In index notation $\mathbf{u}\cdot\mathbf{v} = u_i v_i$, a sum of products of real functions, so the ordinary product rule gives $(u_i v_i)' = u_i' v_i + u_i v_i'$. For (iv), $(\mathbf{u}\times\mathbf{v})_i = \varepsilon_{ijk} u_j v_k$ and the same argument applies; because ε_{ijk} is antisymmetric, the order of factors in each term must be preserved. ∎

Corollary 2.8 (Constant length). *If $|\mathbf{r}(t)|$ is constant, then $\mathbf{r}(t)\cdot\mathbf{r}'(t) = 0$ for all t.*

Proof. Differentiate $\mathbf{r}\cdot\mathbf{r} = c^2$ using (iii): $2\,\mathbf{r}\cdot\mathbf{r}' = 0$. ∎

Geometrically: a curve on a sphere centered at the origin always moves perpendicular to its radius. Applied to $\hat{\mathbf{T}}$, which has constant length 1, the corollary says $\hat{\mathbf{T}}' \perp \hat{\mathbf{T}}$. This fact drives the entire theory of curvature.

Definition 2.9 (Integral). For continuous $\mathbf{r} = \langle f, g, h\rangle$ on $[a,b]$, $\displaystyle\int_a^b \mathbf{r}(t)\,dt = \left\langle \int_a^b f\,dt,\ \int_a^b g\,dt,\ \int_a^b h\,dt \right\rangle$.

The Fundamental Theorem of Calculus carries over componentwise. If $\mathbf{R}' = \mathbf{r}$, then $\int_a^b \mathbf{r}\,dt = \mathbf{R}(b) - \mathbf{R}(a)$, and every antiderivative of $\mathbf{r}$ has the form $\mathbf{R}(t) + \mathbf{C}$ for a constant *vector* $\mathbf{C}$.

Example 2.10 (Tangent line). Find the tangent line to $\mathbf{r}(t) = \langle t^2,\ t^3,\ e^{t-1}\rangle$ at $t = 1$.

Solution. $\mathbf{r}(1) = \langle 1,1,1\rangle$ and $\mathbf{r}'(t) = \langle 2t, 3t^2, e^{t-1}\rangle$, so $\mathbf{r}'(1) = \langle 2,3,1\rangle$. The tangent line is $\langle 1,1,1\rangle + s\,\langle 2,3,1\rangle$.

2.3 Arc Length and Reparametrization

Approximate a curve by an inscribed polygon with vertices $\mathbf{r}(t_0),\dots,\mathbf{r}(t_n)$. The length of the kth side is $|\mathbf{r}(t_k) - \mathbf{r}(t_{k-1})| \approx |\mathbf{r}'(t_k^*)|\,\Delta t$ by the mean value theorem applied to each component. As the partition is refined, the polygon's length converges to an integral. We take its limit as the definition.

Definition 2.11 (Arc length). The *length* of a smooth curve $\mathbf{r}(t)$, $a \le t \le b$,

traversed once is

$$L = \int_a^b |\mathbf{r}'(t)|\,dt = \int_a^b \sqrt{f'(t)^2 + g'(t)^2 + h'(t)^2}\,dt.$$

The *arc length function* measured from $t = a$ is $s(t) = \int_a^t |\mathbf{r}'(u)|\,du$, so that $ds/dt = |\mathbf{r}'(t)|$, the *speed*.

Theorem 2.12 (Invariance of length). *If $t = \varphi(u)$ with φ continuously differentiable and $\varphi'(u) > 0$ on $[\alpha, \beta]$, $\varphi(\alpha) = a$, $\varphi(\beta) = b$, then $\mathbf{q}(u) = \mathbf{r}(\varphi(u))$ traces the same curve and has the same length.*

Proof. By the chain rule, $|\mathbf{q}'(u)| = |\mathbf{r}'(\varphi(u))|\,\varphi'(u)$, so by substitution

$$\int_\alpha^\beta |\mathbf{q}'(u)|\,du = \int_\alpha^\beta |\mathbf{r}'(\varphi(u))|\,\varphi'(u)\,du = \int_a^b |\mathbf{r}'(t)|\,dt.$$ ∎

Because $s'(t) = |\mathbf{r}'(t)| > 0$ for a smooth curve, s is strictly increasing and can be inverted. The curve can then be parametrized by arc length itself. In that parametrization $|\,d\mathbf{r}/ds\,| = 1$: the curve is traversed at unit speed, and $d\mathbf{r}/ds = \hat{\mathbf{T}}$.

Example 2.13 (Helix by arc length). For the helix $\mathbf{r}(t) = \langle a\cos t, a\sin t, bt\rangle$, $|\mathbf{r}'(t)| = \sqrt{a^2\sin^2 t + a^2\cos^2 t + b^2} = \sqrt{a^2+b^2}$ is constant. One turn has length $2\pi\sqrt{a^2+b^2}$. This is the hypotenuse of a right triangle with legs $2\pi a$ and $2\pi b$, as you can see by unrolling the cylinder. With $s = t\sqrt{a^2+b^2}$, the arc-length parametrization is

$$\mathbf{r}(s) = \left\langle a\cos\frac{s}{c},\ a\sin\frac{s}{c},\ \frac{bs}{c}\right\rangle, \qquad c = \sqrt{a^2+b^2}.$$

Example 2.14 (A perfect-square integrand). Find the length of $\mathbf{r}(t) = \left\langle \sqrt{2}\,t,\ \ln t,\ t^2/2\right\rangle$ for $1 \le t \le e$.

Solution. $\mathbf{r}'(t) = \left\langle \sqrt{2},\ 1/t,\ t\right\rangle$, so $|\mathbf{r}'(t)| = \sqrt{t^2 + 2 + t^{-2}} = t + \frac{1}{t}$, a perfect square. Hence $L = \int_1^e (t + t^{-1})\,dt = \frac{e^2-1}{2} + 1 = \frac{e^2+1}{2}$. Most arc-length integrals have no elementary antiderivative; textbook examples are chosen so that the integrand simplifies like this one.

2.4 CURVATURE

How sharply does a curve bend? A natural measure is how fast its direction changes per unit distance travelled. The measure must use arc length rather than t, so that it depends only on the curve and not on how fast it is traced.

Definition 2.15 (Curvature). The *curvature* of a smooth curve is $\kappa = \left|\dfrac{d\hat{\mathbf{T}}}{ds}\right|$. Where $\kappa \neq 0$, $\rho = 1/\kappa$ is the *radius of curvature*.

Theorem 2.16 (Curvature formulas). *For a smooth parametrization* $\mathbf{r}(t)$,

$$\kappa(t) = \frac{|\hat{\mathbf{T}}'(t)|}{|\mathbf{r}'(t)|} = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3}.$$

For a plane curve $y = f(x)$, *this reduces to* $\kappa = \dfrac{|f''(x)|}{\left(1 + f'(x)^2\right)^{3/2}}$.

Proof. By the chain rule, $d\hat{\mathbf{T}}/dt = (d\hat{\mathbf{T}}/ds)(ds/dt)$, which gives the first formula. For the second, write $\mathbf{r}' = v\hat{\mathbf{T}}$ with $v = |\mathbf{r}'|$. Differentiating, $\mathbf{r}'' = v'\hat{\mathbf{T}} + v\hat{\mathbf{T}}'$, so

$$\mathbf{r}' \times \mathbf{r}'' = v\hat{\mathbf{T}} \times (v'\hat{\mathbf{T}} + v\hat{\mathbf{T}}') = v^2\,\hat{\mathbf{T}} \times \hat{\mathbf{T}}',$$

since $\hat{\mathbf{T}} \times \hat{\mathbf{T}} = \mathbf{0}$. By Corollary 2.8, $\hat{\mathbf{T}}' \perp \hat{\mathbf{T}}$, so $|\hat{\mathbf{T}} \times \hat{\mathbf{T}}'| = |\hat{\mathbf{T}}'|$. Thus $|\mathbf{r}' \times \mathbf{r}''| = v^2|\hat{\mathbf{T}}'| = v^3\kappa$. For $y = f(x)$, use $\mathbf{r}(x) = \langle x, f(x), 0\rangle$; then $\mathbf{r}' \times \mathbf{r}'' = \langle 0, 0, f''\rangle$ and $|\mathbf{r}'| = \sqrt{1 + f'^2}$. ∎

Example 2.17 (Circles and lines). A circle of radius a, $\mathbf{r}(t) = \langle a\cos t, a\sin t, 0\rangle$, has $\mathbf{r}' \times \mathbf{r}'' = \langle 0, 0, a^2\rangle$ and $|\mathbf{r}'| = a$, so $\kappa = a^2/a^3 = 1/a$. Small circles bend sharply. A line has $\mathbf{r}'' = \mathbf{0}$ and hence $\kappa = 0$.

Example 2.18 (Curvature of the helix). For $\mathbf{r} = \langle a\cos t, a\sin t, bt\rangle$, $\mathbf{r}' \times \mathbf{r}'' = \langle -a\sin t, a\cos t, b\rangle \times \langle -a\cos t, -a\sin t, 0\rangle = \langle ab\sin t,\ -ab\cos t,\ a^2\rangle$, with norm $a\sqrt{a^2 + b^2}$. Hence

$$\kappa = \frac{a\sqrt{a^2 + b^2}}{(a^2 + b^2)^{3/2}} = \frac{a}{a^2 + b^2}.$$

As $b \to 0$ the helix flattens to a circle and $\kappa \to 1/a$; as $b \to \infty$ it straightens and $\kappa \to 0$.

Example 2.19 (Where is a parabola most curved?). For $y = x^2$, $\kappa(x) = 2/(1+4x^2)^{3/2}$, which is maximized at the vertex $x = 0$, where $\kappa = 2$. The radius of curvature there is $1/2$: the circle $x^2 + (y - \frac{1}{2})^2 = \frac{1}{4}$ hugs the parabola at its vertex more closely than any other circle.

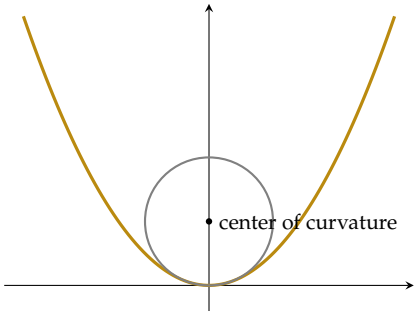


Figure 2.2: The osculating circle of $y = x^2$ at the vertex has radius $1/\kappa = 1/2$.

2.5 The Frenet Frame and Torsion

Definition 2.20 (Normal and binormal). Where $\kappa \neq 0$, the *principal unit normal* and the *binormal* are

$$\hat{\mathbf{N}} = \frac{d\hat{\mathbf{T}}/ds}{\kappa} = \frac{\hat{\mathbf{T}}'(t)}{|\hat{\mathbf{T}}'(t)|}, \qquad \hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}}.$$

The triple $\{\hat{\mathbf{T}}, \hat{\mathbf{N}}, \hat{\mathbf{B}}\}$ is the *Frenet frame*. The plane through the point spanned by $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$ is the *osculating plane*, and the circle in it of radius $1/\kappa$ centered at $\mathbf{r} + \hat{\mathbf{N}}/\kappa$ is the *osculating circle*.

$\hat{\mathbf{N}}$ is a unit vector perpendicular to $\hat{\mathbf{T}}$ (Corollary 2.8), pointing toward the side to which the curve turns. Since $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$ are orthonormal, so is the frame $\{\hat{\mathbf{T}}, \hat{\mathbf{N}}, \hat{\mathbf{B}}\}$, and it is right-handed. As the point moves, this frame rotates. Its rotation is governed by just two numbers.

Theorem 2.21 (Frenet–Serret formulas). *There is a function $\tau(s)$, the* torsion, *such that*

$$\frac{d\hat{\mathbf{T}}}{ds} = \kappa\hat{\mathbf{N}}, \qquad \frac{d\hat{\mathbf{N}}}{ds} = -\kappa\hat{\mathbf{T}} + \tau\hat{\mathbf{B}}, \qquad \frac{d\hat{\mathbf{B}}}{ds} = -\tau\hat{\mathbf{N}}.$$

In matrix form, $\dfrac{\mathrm{d}}{\mathrm{d}s}\begin{pmatrix}\hat{\mathbf{T}}\\ \hat{\mathbf{N}}\\ \hat{\mathbf{B}}\end{pmatrix} = \begin{pmatrix}0 & \kappa & 0\\ -\kappa & 0 & \tau\\ 0 & -\tau & 0\end{pmatrix}\begin{pmatrix}\hat{\mathbf{T}}\\ \hat{\mathbf{N}}\\ \hat{\mathbf{B}}\end{pmatrix}.$

Proof. The first formula is the definition of $\hat{\mathbf{N}}$. For the third, differentiate $\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}}$:

$$\hat{\mathbf{B}}' = \hat{\mathbf{T}}' \times \hat{\mathbf{N}} + \hat{\mathbf{T}} \times \hat{\mathbf{N}}' = \kappa\hat{\mathbf{N}} \times \hat{\mathbf{N}} + \hat{\mathbf{T}} \times \hat{\mathbf{N}}' = \hat{\mathbf{T}} \times \hat{\mathbf{N}}'$$

(primes denote $\mathrm{d}/\mathrm{d}s$). So $\hat{\mathbf{B}}' \perp \hat{\mathbf{T}}$. Also $\hat{\mathbf{B}}' \perp \hat{\mathbf{B}}$ by Corollary 2.8. A vector orthogonal to both $\hat{\mathbf{T}}$ and $\hat{\mathbf{B}}$ is a multiple of $\hat{\mathbf{N}}$; we name the multiple $-\tau$. For the middle formula, expand $\hat{\mathbf{N}}'$ in the orthonormal frame: $\hat{\mathbf{N}}' = (\hat{\mathbf{N}}' \cdot \hat{\mathbf{T}})\hat{\mathbf{T}} + (\hat{\mathbf{N}}' \cdot \hat{\mathbf{N}})\hat{\mathbf{N}} + (\hat{\mathbf{N}}' \cdot \hat{\mathbf{B}})\hat{\mathbf{B}}$. Differentiating $\hat{\mathbf{N}} \cdot \hat{\mathbf{T}} = 0$ gives $\hat{\mathbf{N}}' \cdot \hat{\mathbf{T}} = -\hat{\mathbf{N}} \cdot \hat{\mathbf{T}}' = -\kappa$. Next, $\hat{\mathbf{N}}' \cdot \hat{\mathbf{N}} = 0$ because $|\hat{\mathbf{N}}| = 1$. Finally, differentiating $\hat{\mathbf{N}} \cdot \hat{\mathbf{B}} = 0$ gives $\hat{\mathbf{N}}' \cdot \hat{\mathbf{B}} = -\hat{\mathbf{N}} \cdot \hat{\mathbf{B}}' = \tau$. ∎

The coefficient matrix is antisymmetric. That is no accident: an orthonormal frame can only rotate, and the generator of a rotation is always an antisymmetric matrix. We return to this in Chapter 17, where antisymmetric matrices generate rotations of quantum states. Torsion measures how fast the curve twists out of its osculating plane. A curve with $\kappa > 0$ is planar if and only if $\tau \equiv 0$ (Item 2.34). For computation,

$$\tau = \frac{(\mathbf{r}' \times \mathbf{r}'') \cdot \mathbf{r}'''}{|\mathbf{r}' \times \mathbf{r}''|^2}. \tag{2.1}$$

Example 2.22 (Frenet frame of the helix). For $\mathbf{r} = \langle a\cos t, a\sin t, bt\rangle$ with $c = \sqrt{a^2 + b^2}$:

$$\hat{\mathbf{T}} = \tfrac{1}{c}\langle -a\sin t,\ a\cos t,\ b\rangle, \qquad \hat{\mathbf{N}} = \langle -\cos t,\ -\sin t,\ 0\rangle,$$
$$\hat{\mathbf{B}} = \tfrac{1}{c}\langle b\sin t,\ -b\cos t,\ a\rangle.$$

The normal points horizontally toward the axis. With $\mathbf{r}''' = \langle a\sin t, -a\cos t, 0\rangle$ and $(\mathbf{r}' \times \mathbf{r}'') \cdot \mathbf{r}''' = a^2 b$, (2.1) gives $\tau = a^2 b/(a^2 c^2) = b/(a^2 + b^2)$. Curvature and torsion are both constant, and in fact the helices are the only curves with that property.

Theorem 2.23 (Fundamental theorem of space curves). *Given continuous functions $\kappa(s) > 0$ and $\tau(s)$ on an interval, there is a unit-speed curve with that curvature and torsion, and it is unique up to a rigid motion of space.*

The proof applies the existence and uniqueness theorem for linear systems of ordinary differential equations (Chapter 10) to the Frenet–Serret system, and we omit it. The theorem says that κ and τ are a complete description of the shape of a curve.

2.6 Motion in Space

For a particle at position $\mathbf{r}(t)$, the *velocity* is $\mathbf{v} = \mathbf{r}'$, the *speed* is $v = |\mathbf{v}|$, and the *acceleration* is $\mathbf{a} = \mathbf{r}''$. Newton's second law reads $m\mathbf{r}''(t) = \mathbf{F}$.

Theorem 2.24 (Tangential and normal components of acceleration).

$$\mathbf{a} = a_T\,\hat{\mathbf{T}} + a_N\,\hat{\mathbf{N}}, \qquad a_T = \frac{dv}{dt} = \frac{\mathbf{r}'\cdot\mathbf{r}''}{|\mathbf{r}'|}, \qquad a_N = \kappa v^2 = \frac{|\mathbf{r}'\times\mathbf{r}''|}{|\mathbf{r}'|}.$$

In particular the acceleration always lies in the osculating plane, with no $\hat{\mathbf{B}}$ component.

Proof. From $\mathbf{v} = v\hat{\mathbf{T}}$, $\mathbf{a} = v'\hat{\mathbf{T}} + v\hat{\mathbf{T}}' = v'\hat{\mathbf{T}} + v\,(\kappa v)\hat{\mathbf{N}}$, since $\hat{\mathbf{T}}'(t) = \kappa\, s'(t)\,\hat{\mathbf{N}} = \kappa v\hat{\mathbf{N}}$. Dotting with $\mathbf{v}$ and taking cross products with $\mathbf{v}$ give the two computational formulas. ∎

A driver feels a_T as pressing the accelerator or brake, and $a_N = v^2/\rho$ as the push toward the outside of a turn. Doubling the speed through a curve quadruples the normal force required. This is why highway curves are banked and speed limits drop on tight curves.

Example 2.25 (Uniform circular motion). $\mathbf{r} = \langle R\cos\omega t, R\sin\omega t, 0\rangle$ has $v = R\omega$ constant, so $a_T = 0$ and $a_N = \kappa v^2 = R\omega^2 = v^2/R$. The acceleration $\mathbf{a} = -\omega^2\mathbf{r}$ points to the center: the *centripetal* acceleration.

Example 2.26 (Projectile motion). A projectile is launched from the origin with speed v_0 at elevation angle α, under gravity $\mathbf{a} = -g\hat{\mathbf{j}}$ and with no air resistance. Find its trajectory, range, and maximum height.

Solution. Integrating twice with $\mathbf{v}(0) = v_0\langle\cos\alpha, \sin\alpha\rangle$ and $\mathbf{r}(0) = \mathbf{0}$:

$$\mathbf{r}(t) = \big\langle v_0 t\cos\alpha,\; v_0 t\sin\alpha - \tfrac{1}{2}gt^2\big\rangle.$$

The projectile lands when $y = 0$, at $t = 2v_0\sin\alpha/g$. The range is $R =$

$v_0^2 \sin 2\alpha / g$, maximized at $\alpha = 45°$. The peak occurs at $y' = 0$, giving $H = v_0^2 \sin^2\alpha/(2g)$. Eliminating t shows that the path is the parabola $y = x\tan\alpha - \dfrac{g x^2}{2v_0^2\cos^2\alpha}$.

2.6.1 CENTRAL FORCES AND KEPLER'S LAWS

A force is *central* if it always points along $\mathbf{r}$: $\mathbf{F} = F(r)\,\hat{\mathbf{r}}$. Gravity and the Coulomb force are central.

Theorem 2.27 (Conservation of angular momentum). *Under a central force, $\mathbf{L} = \mathbf{r} \times m\mathbf{v}$ is constant. Consequently the motion lies in a fixed plane through the origin, and the position vector sweeps out equal areas in equal times (Kepler's second law).*

Proof. $\mathrm{d}\mathbf{L}/\mathrm{d}t = \mathbf{v}\times m\mathbf{v} + \mathbf{r}\times m\mathbf{a} = \mathbf{0} + \mathbf{r}\times\mathbf{F} = \mathbf{0}$, since $\mathbf{F} \parallel \mathbf{r}$. Since $\mathbf{r} \perp \mathbf{L}$ at all times, $\mathbf{r}$ stays in the plane through O normal to the fixed vector $\mathbf{L}$. In polar coordinates (r,θ) in that plane, $\mathbf{r} = r\hat{\mathbf{r}}$ and $\mathbf{v} = \dot r\,\hat{\mathbf{r}} + r\dot\theta\,\hat{\boldsymbol{\theta}}$, so $|\mathbf{L}| = mr^2\dot\theta$. The area swept is $A(t) = \tfrac12\int r^2\,d\theta$, so $\dot A = \tfrac12 r^2\dot\theta = |\mathbf{L}|/2m$, which is constant. ∎

Theorem 2.28 (Kepler's first law). *A body moving under the inverse-square attraction $\mathbf{F} = -\dfrac{GMm}{r^2}\hat{\mathbf{r}}$ with $\mathbf{L} \neq \mathbf{0}$ travels on a conic section with a focus at the origin:*

$$r = \frac{h^2/GM}{1 + e\cos\theta}, \qquad h = |\mathbf{r}\times\mathbf{v}|,$$

where $e \geq 0$ is the eccentricity and θ is measured from the direction of closest approach. Bound orbits ($e < 1$) are ellipses.

Proof. Let $\mathbf{h} = \mathbf{r}\times\mathbf{v}$, constant by Theorem 2.27. With $\mathbf{a} = -GM\hat{\mathbf{r}}/r^2$ and $\mathbf{r} = r\hat{\mathbf{r}}$,

$$\mathbf{a}\times\mathbf{h} = -\frac{GM}{r^2}\,\hat{\mathbf{r}}\times(r\hat{\mathbf{r}}\times\mathbf{v}) = -\frac{GM}{r}\big[(\hat{\mathbf{r}}\cdot\mathbf{v})\hat{\mathbf{r}} - \mathbf{v}\big],$$

by Theorem 1.26. Now $\hat{\mathbf{r}}\cdot\mathbf{v} = \dot r$ and $\mathbf{v} = \dot r\hat{\mathbf{r}} + r\,\mathrm{d}\hat{\mathbf{r}}/\mathrm{d}t$, so the bracket is $-r\,\mathrm{d}\hat{\mathbf{r}}/\mathrm{d}t$ and $\mathbf{a}\times\mathbf{h} = GM\,\mathrm{d}\hat{\mathbf{r}}/\mathrm{d}t$. Since $\mathbf{h}$ is constant, the left side is $\mathrm{d}(\mathbf{v}\times\mathbf{h})/\mathrm{d}t$. Integrating,

$$\mathbf{v}\times\mathbf{h} = GM\,\hat{\mathbf{r}} + \mathbf{c}$$

for a constant vector $\mathbf{c}$ in the orbital plane. Dot with $\mathbf{r}$. On the left, $\mathbf{r}\cdot(\mathbf{v}\times\mathbf{h}) = (\mathbf{r}\times\mathbf{v})\cdot\mathbf{h} = h^2$, by cyclicity of the triple product. On the right we get $GMr +$

$rc\cos\theta$, where θ is the angle from $\mathbf{c}$ to $\mathbf{r}$. Solving for r gives the formula with $e = c/GM$, which is the polar equation of a conic with focus at the origin. ■

The constant vector $\mathbf{A} = \mathbf{v} \times \mathbf{h} - GM\hat{\mathbf{r}}$ is the *Laplace–Runge–Lenz vector*. It points toward perihelion, and its conservation is special to the inverse-square law. For any other central force, orbits precess.

LOOKING AHEAD. The Coulomb attraction between a proton and an electron is also inverse-square, so classically an electron should orbit on a Kepler ellipse. But an accelerating charge radiates, so the classical electron would spiral into the nucleus in about 10^{-11} seconds. Atoms are stable, and the failure of this calculation is one of the crises that produced quantum mechanics (Chapter 12). Remarkably, the quantum Runge–Lenz vector explains why hydrogen energy levels depend only on the principal quantum number n (Chapter 16).

Theorem 2.29 (Kepler's third law). *For an elliptical orbit with semi-major axis a, the period satisfies $T^2 = \dfrac{4\pi^2}{GM}\,a^3$.*

Proof. The area of the ellipse is πab, swept at the constant rate $\dot{A} = h/2$, so $T = 2\pi ab/h$. For the conic of Theorem 2.28, the semi-latus rectum is $\ell = h^2/GM$, and for an ellipse $\ell = b^2/a$. Hence $T^2 = 4\pi^2a^2b^2/h^2 = 4\pi^2a^2b^2/(GM\,b^2/a) = 4\pi^2a^3/GM$. ■

2.7 Worked Problems

Each problem below is solved in full. The helix and Kepler's laws were treated in the text, so these problems concentrate on curves where the formulas must be applied without shortcuts.

Problem 2.1 The full Frenet apparatus of a cubic curve

For $\mathbf{r}(t) = \langle t,\ t^2,\ \tfrac{2}{3}t^3\rangle$, find the speed, κ, and τ as functions of t, and the frame $\hat{\mathbf{T}}, \hat{\mathbf{N}}, \hat{\mathbf{B}}$ at $t = 1$.

Solution. Differentiate three times: $\mathbf{r}' = \langle 1, 2t, 2t^2\rangle$, $\mathbf{r}'' = \langle 0, 2, 4t\rangle$, $\mathbf{r}''' = \langle 0, 0, 4\rangle$. The speed is a perfect square,

$$|\mathbf{r}'| = \sqrt{1 + 4t^2 + 4t^4} = 1 + 2t^2,$$

which is why this curve was chosen: its arc length $s = t + \frac{2}{3}t^3$ is elementary. Next,

$$\mathbf{r}' \times \mathbf{r}'' = \langle 2t \cdot 4t - 2t^2 \cdot 2,\ 2t^2 \cdot 0 - 1 \cdot 4t,\ 1 \cdot 2 - 2t \cdot 0 \rangle = \langle 4t^2,\ -4t,\ 2 \rangle,$$

with length $2\sqrt{4t^4 + 4t^2 + 1} = 2(1 + 2t^2)$, and $(\mathbf{r}' \times \mathbf{r}'') \cdot \mathbf{r}''' = 8$. By Theorem 2.16 and (2.1),

$$\kappa = \frac{2(1 + 2t^2)}{(1 + 2t^2)^3} = \frac{2}{(1 + 2t^2)^2}, \qquad \tau = \frac{8}{4(1 + 2t^2)^2} = \frac{2}{(1 + 2t^2)^2}.$$

At $t = 1$, $\kappa = \tau = \frac{2}{9}$. The frame: $\hat{\mathbf{T}} = \frac{1}{3}\langle 1, 2, 2 \rangle$, and $\hat{\mathbf{B}}$ is the unit vector along $\mathbf{r}' \times \mathbf{r}'' = \langle 4, -4, 2 \rangle$, so $\hat{\mathbf{B}} = \frac{1}{3}\langle 2, -2, 1 \rangle$. Then

$$\begin{aligned} \hat{\mathbf{N}} = \hat{\mathbf{B}} \times \hat{\mathbf{T}} &= \frac{1}{9}\langle (-2)(2) - (1)(2),\ (1)(1) - (2)(2),\ (2)(2) - (-2)(1) \rangle \\ &= \frac{1}{3}\langle -2, -1, 2 \rangle. \end{aligned}$$

Checks: the three vectors have length 1 and are mutually orthogonal, and $\hat{\mathbf{N}} \cdot \mathbf{r}'' = \frac{1}{3}(0 - 2 + 8) > 0$, so $\hat{\mathbf{N}}$ points toward the side to which the curve bends.

The equality $\kappa = \tau$ holds for every t, not just $t = 1$. Curves with constant ratio τ/κ are studied in Problem 2.5: they are the curves whose tangent makes a constant angle with a fixed direction.

Problem 2.2 Tangential and normal acceleration

A particle moves along $\mathbf{r}(t) = e^t \langle \cos t, \sin t, 1 \rangle$. Find its speed, the tangential and normal components of acceleration, the curvature of its path, and the arc length from $t = 0$ to $t = T$.

Solution. By the product rule,

$$\mathbf{v} = e^t \langle \cos t - \sin t,\ \sin t + \cos t,\ 1 \rangle, \qquad \mathbf{a} = e^t \langle -2\sin t,\ 2\cos t,\ 1 \rangle.$$

Since $(\cos t - \sin t)^2 + (\sin t + \cos t)^2 = 2$, the speed is $v = e^t\sqrt{3}$, and $|\mathbf{a}| = e^t\sqrt{4 + 1} = e^t\sqrt{5}$. By Theorem 2.24,

$$a_T = \frac{\mathrm{d}v}{\mathrm{d}t} = \sqrt{3}\,e^t, \qquad a_N = \sqrt{|\mathbf{a}|^2 - a_T^2} = \sqrt{5 - 3}\,e^t = \sqrt{2}\,e^t,$$

$$\kappa = \frac{a_N}{v^2} = \frac{\sqrt{2}e^t}{3e^{2t}} = \frac{\sqrt{2}}{3}e^{-t}.$$

The arc length is $\int_0^T \sqrt{3}e^t\,dt = \sqrt{3}(e^T - 1)$. The curvature decreases as the particle spirals up and outward on the cone $z^2 = x^2 + y^2$. This is a *conical spiral*: it makes a constant angle with the cone's generators, since $\mathbf{v}\cdot\mathbf{r}/(|\mathbf{v}||\mathbf{r}|) = \frac{2e^{2t}}{\sqrt{3}e^t\sqrt{2}e^t} = \sqrt{2/3}$ for all t.

Problem 2.3 Range on an inclined plane

A projectile is launched with speed v_0 from the bottom of a slope of inclination β, at angle $\alpha > \beta$ above the horizontal. Ignoring air resistance, find the range R measured along the slope, the angle that maximizes it, and the maximum range.

Solution. From Example 2.26, $x = v_0\cos\alpha\, t$ and $y = v_0\sin\alpha\, t - \frac{1}{2}gt^2$. The projectile lands when $y = x\tan\beta$. Dividing out the root $t = 0$,

$$v_0\sin\alpha - \tfrac{1}{2}gt = v_0\cos\alpha\tan\beta,$$
$$t_f = \frac{2v_0(\sin\alpha\cos\beta - \cos\alpha\sin\beta)}{g\cos\beta} = \frac{2v_0\sin(\alpha-\beta)}{g\cos\beta}.$$

The landing point is a horizontal distance $x(t_f)$ from the launch, so the distance along the slope is

$$R = \frac{x(t_f)}{\cos\beta} = \frac{2v_0^2\cos\alpha\sin(\alpha-\beta)}{g\cos^2\beta} = \frac{v_0^2\big[\sin(2\alpha-\beta) - \sin\beta\big]}{g\cos^2\beta},$$

using $2\cos\alpha\sin(\alpha-\beta) = \sin(2\alpha-\beta) - \sin\beta$. Now only one term depends on α, and it is maximized when $2\alpha - \beta = \pi/2$, that is, $\alpha^* = \frac{\pi}{4} + \frac{\beta}{2}$: the optimal launch direction bisects the angle between the slope and the vertical. Then, using $\cos^2\beta = (1-\sin\beta)(1+\sin\beta)$,

$$R_{\max} = \frac{v_0^2(1-\sin\beta)}{g\cos^2\beta} = \frac{v_0^2}{g(1+\sin\beta)}.$$

Checks: $\beta = 0$ gives the familiar $\alpha^* = 45^\circ$ and $R_{\max} = v_0^2/g$. Firing downhill ($\beta < 0$) gives a larger range, and as $\beta \to -\pi/2$ the range grows without bound, since the projectile falls forever along a vertical cliff.

Problem 2.4 The osculating circle of a parabola

For $y = x^2$, find the curvature and the center of curvature at $(1,1)$. Then find the center of curvature at a general point, the curve it traces (the *evolute*), and the curvature at the vertex.

Solution. For a plane graph $\kappa = \dfrac{|f''|}{(1+f'^2)^{3/2}} = \dfrac{2}{(1+4x^2)^{3/2}}$. At $x = 1$, $\kappa = 2/5^{3/2}$ and the radius of curvature is $\rho = 5\sqrt{5}/2 \approx 5.59$. The parabola bends

upward, so the unit normal points up and to the left, perpendicular to the tangent $\langle 1,2\rangle$: $\hat{\mathbf{N}} = \frac{1}{\sqrt{5}}\langle -2,1\rangle$. The center of curvature is

$$(1,1) + \rho\hat{\mathbf{N}} = (1,1) + \frac{5\sqrt{5}}{2}\cdot\frac{\langle -2,1\rangle}{\sqrt{5}} = \Big(-4,\ \frac{7}{2}\Big).$$

At a general point (x,x^2) the normal is $\langle -2x,1\rangle/\sqrt{1+4x^2}$ and $\rho = (1+4x^2)^{3/2}/2$, so the center is

$$(x,x^2) + \frac{1+4x^2}{2}\langle -2x,1\rangle = \Big(-4x^3,\ \frac{1}{2}+3x^2\Big).$$

At $x=1$ this gives $(-4,\frac{7}{2})$, as before. Eliminating x: $X=-4x^3$, $Y-\frac{1}{2}=3x^2$, so $\big(Y-\frac{1}{2}\big)^3 = 27x^6 = \frac{27}{16}X^2$. The evolute is a *semicubical parabola* with a cusp at $(0,\frac{1}{2})$. At the vertex $\kappa=2$, so the osculating circle there has radius $\frac{1}{2}$ and center $(0,\frac{1}{2})$, the cusp of the evolute. Note that the vertex is where the curvature is maximal, and the evolute has a cusp exactly there.

Problem 2.5 A curve with $\kappa = \tau$

Let $\mathbf{r}(t) = \langle \cosh t, \sinh t, t\rangle$.

(a) Reparametrize by arc length from $t=0$.

(b) Compute κ and τ, and show $\kappa=\tau$ everywhere.

(c) Show that $\hat{\mathbf{T}}$ makes a constant angle with a fixed direction, and find the direction and the angle.

Solution. (a) $\mathbf{r}' = \langle \sinh t, \cosh t, 1\rangle$, so $|\mathbf{r}'|^2 = \sinh^2 t + \cosh^2 t + 1 = 2\cosh^2 t$ and $|\mathbf{r}'| = \sqrt{2}\cosh t$. Then $s = \int_0^t \sqrt{2}\cosh u\,du = \sqrt{2}\sinh t$, which inverts to $t = \operatorname{arcsinh}(s/\sqrt{2})$. Using $\cosh t = \sqrt{1+\sinh^2 t}$,

$$\mathbf{r}(s) = \Big(\sqrt{1+s^2/2},\ \frac{s}{\sqrt{2}},\ \operatorname{arcsinh}\frac{s}{\sqrt{2}}\Big).$$

(b) $\mathbf{r}'' = \langle \cosh t, \sinh t, 0\rangle$, $\mathbf{r}''' = \langle \sinh t, \cosh t, 0\rangle$, and

$$\begin{aligned}\mathbf{r}'\times\mathbf{r}'' &= \big\langle \cosh t\cdot 0 - 1\cdot\sinh t,\ 1\cdot\cosh t - \sinh t\cdot 0,\ \sinh^2 t - \cosh^2 t\big\rangle \\ &= \langle -\sinh t, \cosh t, -1\rangle .\end{aligned}$$

Its length is again $\sqrt{2}\cosh t$, and $(\mathbf{r}'\times\mathbf{r}'')\cdot\mathbf{r}''' = \cosh^2 t - \sinh^2 t = 1$. Hence

$$\kappa = \frac{\sqrt{2}\cosh t}{(\sqrt{2}\cosh t)^3} = \frac{1}{2\cosh^2 t}, \qquad \tau = \frac{1}{2\cosh^2 t}.$$

(c) Lancret's theorem says that a curve is a *generalized helix*, with tangent at a constant angle to a fixed direction, if and only if τ/κ is constant. Here $\tau/\kappa = 1$. The fixed direction is that of the Darboux vector $\tau\hat{\mathbf{T}} + \kappa\hat{\mathbf{B}}$, which is constant whenever τ/κ is, because by the Frenet equations $(\tau\hat{\mathbf{T}} + \kappa\hat{\mathbf{B}})' = \tau'\hat{\mathbf{T}} + \kappa'\hat{\mathbf{B}} + (\tau\kappa - \kappa\tau)\hat{\mathbf{N}}$, and this is parallel to $\tau\hat{\mathbf{T}} + \kappa\hat{\mathbf{B}}$ exactly when $\tau'/\tau = \kappa'/\kappa$. Here it is proportional to $\hat{\mathbf{T}} + \hat{\mathbf{B}}$. With $\hat{\mathbf{T}} = \frac{\langle \sinh t, \cosh t, 1\rangle}{\sqrt{2}\cosh t}$ and $\hat{\mathbf{B}} = \frac{\langle -\sinh t, \cosh t, -1\rangle}{\sqrt{2}\cosh t}$,

$$\hat{\mathbf{T}} + \hat{\mathbf{B}} = \frac{\langle 0,\, 2\cosh t,\, 0\rangle}{\sqrt{2}\cosh t} = \left\langle 0, \sqrt{2}, 0 \right\rangle.$$

So the axis is the y-axis, and $\hat{\mathbf{T}} \cdot \hat{\jmath} = \dfrac{\cosh t}{\sqrt{2}\cosh t} = \dfrac{1}{\sqrt{2}}$: the tangent makes a constant angle of 45° with the y-axis. In hindsight this is visible from $\mathbf{r}'$ directly, whose y-component $\cosh t$ is always $1/\sqrt{2}$ of its length. The general theory tells us where to look when no such shortcut is apparent.

Problem 2.6 Plane curves are exactly the curves with zero torsion

Let $\mathbf{r}(s)$ be a unit-speed curve with $\kappa > 0$. Prove that $\mathbf{r}$ lies in a plane if and only if $\tau \equiv 0$.

Solution. Suppose $\mathbf{r}$ lies in the plane $\mathbf{n} \cdot \mathbf{x} = c$ with $\mathbf{n}$ a unit vector. Differentiating $\mathbf{n} \cdot \mathbf{r}(s) = c$ gives $\mathbf{n} \cdot \hat{\mathbf{T}} = 0$, and again, $\mathbf{n} \cdot \kappa\hat{\mathbf{N}} = 0$, so $\mathbf{n} \cdot \hat{\mathbf{N}} = 0$ since $\kappa > 0$. Then $\mathbf{n}$ is orthogonal to $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$, so $\hat{\mathbf{B}} = \pm\mathbf{n}$ is constant, and $\tau\hat{\mathbf{N}} = -\hat{\mathbf{B}}' = \mathbf{0}$ by the Frenet equations (Theorem 2.21). So $\tau = 0$.

Conversely, suppose $\tau \equiv 0$. Then $\hat{\mathbf{B}}' = -\tau\hat{\mathbf{N}} = \mathbf{0}$, so $\hat{\mathbf{B}} = \hat{\mathbf{B}}_0$ is constant. Consider $f(s) = \hat{\mathbf{B}}_0 \cdot (\mathbf{r}(s) - \mathbf{r}(0))$. Then $f(0) = 0$ and $f'(s) = \hat{\mathbf{B}}_0 \cdot \hat{\mathbf{T}}(s) = 0$, so $f \equiv 0$: every point of the curve lies in the plane through $\mathbf{r}(0)$ with normal $\hat{\mathbf{B}}_0$.

The hypothesis $\kappa > 0$ is needed for $\hat{\mathbf{N}}$ and $\hat{\mathbf{B}}$ to be defined. A curve made of two planar arcs in different planes, joined at a point where $\kappa = 0$ and the arcs are tangent to a common line, has $\tau = 0$ wherever τ is defined, but is not planar.

EXERCISES

A. COMPUTATION

2.1 Find the domain of $\mathbf{r}(t) = \left\langle \ln(4 - t^2),\, \sqrt{t+1},\, e^{-t} \right\rangle$ and compute $\lim_{t\to 0} \mathbf{r}(t)$.

2.2 Evaluate $\lim\limits_{t\to 0} \left\langle \dfrac{\sin 3t}{t},\, \dfrac{1-\cos t}{t^2},\, \dfrac{e^{2t}-1}{t} \right\rangle$.

2.3 Find a vector function for the curve of intersection of the paraboloid $z = x^2 + y^2$ and the plane $z = 4x$.

2.4 Show that $\mathbf{r}(t) = \langle \sin t,\, \cos t,\, \sin^2 t \rangle$ lies on the parabolic cylinder $z = x^2$ and on the circular cylinder $x^2 + y^2 = 1$, and describe the curve.

2.5 For $\mathbf{r}(t) = \langle te^t,\, \ln(1+t^2),\, \arctan t \rangle$, find $\mathbf{r}'(0)$ and the unit tangent vector at $t = 0$.

2.6 Find the tangent line to $\mathbf{r}(t) = \langle 2\cos t,\, 2\sin t,\, 4\cos 2t \rangle$ at $t = \pi/3$.

2.7 Evaluate $\displaystyle\int_0^1 \left\langle 4t^3,\, \frac{1}{1+t^2},\, te^{t^2} \right\rangle dt$.

2.8 Find $\mathbf{r}(t)$ if $\mathbf{r}'(t) = \left\langle 2t,\, 3t^2,\, \sqrt{t} \right\rangle$ and $\mathbf{r}(1) = \langle 1, 1, 0 \rangle$.

2.9 Find the length of $\mathbf{r}(t) = \langle 2t,\, t^2,\, \tfrac{1}{3}t^3 \rangle$ from $t = 0$ to $t = 3$.

2.10 Find the length of $\mathbf{r}(t) = \langle \cos^3 t,\, \sin^3 t,\, 0 \rangle$ for $0 \le t \le 2\pi$ (the astroid).

2.11 Find the length of one turn of the helix $\langle 3\cos t, 3\sin t, 4t \rangle$, and reparametrize it by arc length from $t = 0$.

2.12 Find the length of $\mathbf{r}(t) = \langle e^t \cos t,\, e^t \sin t,\, e^t \rangle$ from $t = 0$ to $t = \ln 2$.

2.13 Find the curvature of $\mathbf{r}(t) = \langle t,\, t^2,\, t^3 \rangle$ at $t = 0$ and at $t = 1$.

2.14 Find the curvature of $y = \ln x$, and the point at which it is largest.

2.15 Find the curvature of $y = \sin x$ at $x = \pi/2$ and the osculating circle there.

2.16 Find the curvature of the ellipse $\mathbf{r}(t) = \langle a\cos t, b\sin t \rangle$ at $t = 0$ and at $t = \pi/2$, with $a > b > 0$.

2.17 Find $\hat{\mathbf{T}}$, $\hat{\mathbf{N}}$, and $\hat{\mathbf{B}}$ for $\mathbf{r}(t) = \langle 2\sin 3t,\, t,\, 2\cos 3t \rangle$.

2.18 Find the torsion of the twisted cubic $\mathbf{r}(t) = \langle t, t^2, t^3 \rangle$ at $t = 0$.

2.19 Find the equation of the osculating plane of $\mathbf{r}(t) = \langle t, t^2, t^3 \rangle$ at $(1, 1, 1)$.

2.20 Find a_T and a_N for $\mathbf{r}(t) = \langle t,\, t^2,\, \tfrac{2}{3}t^3 \rangle$ at $t = 1$.

2.21 Find a_T and a_N for $\mathbf{r}(t) = \left\langle e^t,\, \sqrt{2}\,t,\, e^{-t} \right\rangle$.

2.22 A projectile is fired at $300\,\text{m/s}$ at 30° above horizontal from ground level. With $g = 9.8\,\text{m/s}^2$, find the range, maximum height, and impact speed.

2.23 A ball is thrown from the top of a $20\,\text{m}$ building at $15\,\text{m/s}$, 45° above horizontal. Where does it land, using $g = 9.8\,\text{m/s}^2$?

2.24 A car of mass $1200\,\text{kg}$ rounds an unbanked curve of radius $80\,\text{m}$ at $25\,\text{m/s}$. Find the frictional force required.

2.25 A particle has $\mathbf{a}(t) = \langle 6t,\ 12t^2,\ -6t\rangle$, $\mathbf{v}(0) = \langle 1, -1, 3\rangle$, and $\mathbf{r}(0) = \mathbf{0}$. Find $\mathbf{r}(t)$.

2.26 The Moon's orbit has semi-major axis $3.84 \times 10^8\,\text{m}$ and period 27.3 days. Estimate the mass of the Earth.

2.27 A satellite in circular orbit $400\,\text{km}$ above the Earth's surface: find its period, using $R_E = 6.371 \times 10^6\,\text{m}$ and $GM_E = 3.986 \times 10^{14}\,\text{m}^3/\text{s}^2$.

2.28 An orbit has perihelion distance r_p and aphelion distance r_a. Express the eccentricity and the semi-latus rectum in terms of r_p and r_a.

B. THEORY AND PROOF

2.29 Prove parts (ii) and (iv) of Theorem 2.7.

2.30 Prove that if $\mathbf{r}'(t) = \mathbf{0}$ for all t in an interval, then $\mathbf{r}$ is constant there.

2.31 Prove that $\frac{\mathrm{d}}{\mathrm{d}t}\big[\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\big] = \mathbf{u}'\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{v}'\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}')$.

2.32 Show that if $\mathbf{r}\times\mathbf{r}' = \mathbf{0}$ for all t and $\mathbf{r}\neq\mathbf{0}$, then $\mathbf{r}(t)$ has constant direction.

2.33 Derive the torsion formula (2.1). (Hint: express $\mathbf{r}'''$ in the Frenet frame.)

2.34 Prove that a curve with $\kappa > 0$ lies in a plane if and only if $\tau \equiv 0$.

2.35 Show that a curve with constant $\kappa > 0$ and $\tau = 0$ is a circle of radius $1/\kappa$.

2.36 Prove that the *Darboux vector* $\boldsymbol{\omega} = \tau\hat{\mathbf{T}} + \kappa\hat{\mathbf{B}}$ satisfies $\hat{\mathbf{X}}' = \boldsymbol{\omega}\times\hat{\mathbf{X}}$ for each $\hat{\mathbf{X}} \in \{\hat{\mathbf{T}}, \hat{\mathbf{N}}, \hat{\mathbf{B}}\}$. Interpret $\boldsymbol{\omega}$ as the angular velocity of the frame.

2.37 Show that the Laplace–Runge–Lenz vector $\mathbf{A} = \mathbf{v}\times\mathbf{h} - GM\hat{\mathbf{r}}$ is conserved, and that $|\mathbf{A}| = GMe$.

2.38 Show that for a central force $F(r)\hat{\mathbf{r}}$ the energy $E = \tfrac{1}{2}mv^2 + U(r)$, where $U'(r) = -F(r)$, is conserved.

C. CHALLENGE

2.39 Show that the length of the ellipse $\langle a\cos t, b\sin t\rangle$ is $4a\int_0^{\pi/2}\sqrt{1 - e^2\sin^2 t}\,dt$, where $e^2 = 1 - b^2/a^2$. (This is a complete elliptic integral and has no elementary closed form.)

2.40 A bead slides without friction down a helical wire $\langle a\cos t, a\sin t, -bt\rangle$ under gravity, starting from rest. Find its speed as a function of t and the normal force the wire must supply. (Use energy conservation.)

2.41 Prove that a curve all of whose normal lines pass through a fixed point is (part of) a circle.

2.42 For the orbit of Theorem 2.28 show that the total energy is $E = -GMm/(2a)$ for an ellipse. Deduce that bound orbits have $E < 0$, parabolic orbits $E = 0$, and hyperbolic orbits $E > 0$.

CHAPTER 3

Partial Derivatives and the Gradient

LEARNING OBJECTIVES

- Describe functions of several variables through graphs, level curves, and level surfaces.
- Work with limits and continuity in several variables, including proofs that limits fail to exist.
- Compute partial derivatives of all orders and state Clairaut's theorem on mixed partials.
- Define differentiability, use linear approximation, and prove that continuous partials imply differentiability.
- Apply the multivariable chain rule in matrix form, including changes of variables in differential operators.
- Compute directional derivatives and gradients, and use the gradient's geometry: steepest ascent and normals to level sets.
- State the implicit and inverse function theorems.

Temperature in a room depends on three coordinates and time. The potential energy of a molecule depends on the positions of all its atoms. A quantum wavefunction $\psi(x, y, z, t)$ depends on four variables. Calculus in one variable is not enough for any of these. This chapter builds differential calculus for functions of several variables. Its central object is the *gradient*, and the operator ∇ that produces it reappears in the Schrödinger equation as the momentum operator $-i\hbar\nabla$.

3.1 Functions of Several Variables

Definition 3.1 (Function of several variables). A *real-valued function of n variables* is a rule $f : D \to \mathbb{R}$ that assigns a real number $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ to each point $\mathbf{x}$ of a set $D \subseteq \mathbb{R}^n$, the *domain*. The *range* is $\{f(\mathbf{x}) : \mathbf{x} \in D\}$. Unless stated otherwise, the domain is the largest set on which the formula makes sense.

For $n = 2$, the *graph* $\{(x, y, z) : z = f(x, y)\}$ is a surface in $\mathbb{R}^3$. Graphs are hard to draw by hand, and for $n \geq 3$ impossible, so we use *level sets* instead. The *level curve* of $f(x, y)$ at height k is $\{(x, y) : f(x, y) = k\}$, and a collection of level curves for evenly spaced k is a *contour map*. Where the curves crowd together, the function changes rapidly. For $f(x, y, z)$ the level sets $f = k$ are *level surfaces*: the equipotential surfaces of electrostatics, or the isotherms of a temperature field.

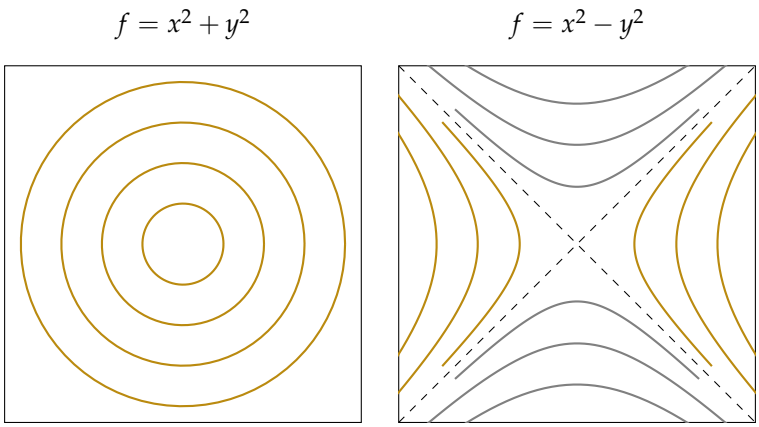


Figure 3.1: Contour maps of a bowl (left) and a saddle (right). In the saddle, gold curves are positive levels, gray curves are negative levels, and the dashed lines are the zero level.

Example 3.2 (Domain and level curves). Find the domain of $f(x, y) = \sqrt{9 - x^2 - y^2}$, its range, and its level curves.

Solution. The domain is the closed disk $x^2 + y^2 \leq 9$ and the range is $[0, 3]$. The graph is the upper hemisphere of radius 3. For $0 \leq k \leq 3$ the level curve $f = k$ is the circle $x^2 + y^2 = 9 - k^2$. The curves crowd together near the boundary, where the hemisphere is steep.

3.2 Limits and Continuity

In one variable, x can approach a from only two sides. In the plane, (x, y) can approach (a, b) along infinitely many paths, and a limit must exist along all of them simultaneously. This makes multivariable limits both subtler and more interesting.

Definition 3.3 (Limit). Let f be defined on a set containing points arbitrarily close to $\mathbf{a} \in \mathbb{R}^n$. We write $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = L$ if for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$0 < |\mathbf{x} - \mathbf{a}| < \delta,\ \mathbf{x} \in D \quad \Longrightarrow \quad |f(\mathbf{x}) - L| < \varepsilon.$$

f is *continuous* at $\mathbf{a} \in D$ if $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$, and continuous on D if it is continuous at every point of D.

Sums, products, quotients (where the denominator is nonzero), and compositions of continuous functions are continuous; the proofs are identical to those in one variable. Polynomials in $x_1, \dots, x_n$ are continuous everywhere, and rational functions are continuous on their domains.

Theorem 3.4 (Two-path test). *If $f(\mathbf{x}) \to L_1$ as $\mathbf{x} \to \mathbf{a}$ along one path and $f(\mathbf{x}) \to L_2 \neq L_1$ along another, then $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x})$ does not exist.*

Proof. If the limit were L, the definition would force every path limit to equal L. ∎

Example 3.5 (A limit that fails). Show that $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ does not exist.

Solution. Along the line $y = mx$ the function equals $\dfrac{mx^2}{x^2(1+m^2)} = \dfrac{m}{1+m^2}$. This depends on m: it is 0 along the axes and $\tfrac{1}{2}$ along $y = x$.

Example 3.6 (Lines are not enough). Let $f(x,y) = \dfrac{x^2 y}{x^4+y^2}$. Along every line $y = mx$, $f = \dfrac{mx}{x^2+m^2} \to 0$, and along the y-axis $f = 0$. Yet along the parabola $y = x^2$, $f = \tfrac{1}{2}$. The limit at the origin does not exist, even though every straight-line approach gives 0.

COMMON ERROR (proving existence). Checking finitely many paths can only ever *disprove* the existence of a limit. To prove a limit exists, you need an estimate that controls $|f(\mathbf{x}) - L|$ for all $\mathbf{x}$ near $\mathbf{a}$ at once. Squeeze arguments and polar coordinates are the usual tools.

Example 3.7 (Proving a limit by squeezing). Show that $\lim_{(x,y)\to(0,0)} \dfrac{3x^2y}{x^2+y^2} = 0$.

Solution. Since $x^2 \le x^2 + y^2$, we have $\left|\dfrac{3x^2y}{x^2+y^2}\right| \le 3|y| \le 3\sqrt{x^2+y^2}$. Given $\varepsilon > 0$, take $\delta = \varepsilon/3$. In polar coordinates the same estimate reads $|3r\cos^2\theta\sin\theta| \le 3r \to 0$, uniformly in θ.

The word *uniformly* matters. A bound of the form $|f - L| \le g(r)$ with $g(r) \to 0$ and g independent of θ proves the limit. In Example 3.6, by contrast, $f = r\cos^2\theta\sin\theta/(r^2\cos^4\theta + \sin^2\theta)$ tends to 0 for each fixed θ, but not uniformly: near $\theta = 0$ the denominator is small.

3.3 Partial Derivatives

Definition 3.8 (Partial derivative). The *partial derivatives* of $f(x,y)$ at (a,b) are

$$f_x(a,b) = \frac{\partial f}{\partial x} = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h},$$
$$f_y(a,b) = \frac{\partial f}{\partial y} = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h}.$$

For $f(x_1,\dots,x_n)$, $\partial f/\partial x_i$ is the derivative with respect to x_i with all other variables held fixed.

To compute f_x, treat every variable except x as a constant and differentiate. Geometrically, $f_x(a,b)$ is the slope of the curve cut from the graph by the vertical plane $y = b$, and $f_y(a,b)$ is the slope in the plane $x = a$.

Example 3.9 (Computing partials). For $f(x,y,z) = x^2y^3z + e^{xz}$: $f_x = 2xy^3z + ze^{xz}$, $f_y = 3x^2y^2z$, $f_z = x^2y^3 + xe^{xz}$.

3.3.1 HIGHER DERIVATIVES AND CLAIRAUT'S THEOREM

Partial derivatives are themselves functions, and can be differentiated again. We write

$$f_{xx} = \frac{\partial^2 f}{\partial x^2}, \qquad f_{xy} = (f_x)_y = \frac{\partial^2 f}{\partial y\,\partial x}, \qquad f_{yx} = (f_y)_x = \frac{\partial^2 f}{\partial x\,\partial y}.$$

Note the opposite order of the subscripts and the denominators.

Theorem 3.10 (Clairaut–Schwarz). *If f_{xy} and f_{yx} are both continuous on an open set containing (a,b), then $f_{xy}(a,b) = f_{yx}(a,b)$.*

Proof. For small $h \neq 0$ consider the second difference

$$\Delta(h) = f(a+h,b+h) - f(a+h,b) - f(a,b+h) + f(a,b).$$

With $g(x) = f(x,b+h) - f(x,b)$, $\Delta = g(a+h) - g(a) = hg'(c_1) = h\,[f_x(c_1,b+h) - f_x(c_1,b)]$ by the mean value theorem, for some c_1 between a and $a+h$. Applying the mean value theorem again, in y, gives $\Delta = h^2 f_{xy}(c_1,d_1)$. Grouping the four terms the other way round gives likewise $\Delta = h^2 f_{yx}(c_2,d_2)$. Thus $f_{xy}(c_1,d_1) = f_{yx}(c_2,d_2)$ with both points within distance $\sqrt{2}|h|$ of (a,b). Let $h \to 0$ and use continuity. ∎

The continuity hypothesis cannot be dropped. For $f(x,y) = xy(x^2-y^2)/(x^2+y^2)$ with $f(0,0) = 0$, one computes $f_{xy}(0,0) = -1$ but $f_{yx}(0,0) = 1$ (Item 3.34).

3.3.2 PARTIAL DIFFERENTIAL EQUATIONS

An equation relating a function to its partial derivatives is a *partial differential equation* (PDE). Three PDEs recur throughout physics and throughout this book:

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \qquad \text{(Laplace)}$$

$$\frac{\partial u}{\partial t} = k\,\nabla^2 u \qquad \text{(heat)}$$

$$\frac{\partial^2 u}{\partial t^2} = c^2\,\nabla^2 u \qquad \text{(wave)}$$

Laplace's equation governs electrostatics and steady heat flow; the heat equation governs diffusion; the wave equation governs sound, light, and vibrating strings. The operator ∇^2 is the *Laplacian*. The time-dependent Schrödinger equation, $i\hbar\,\partial\psi/\partial t = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi$, looks like a heat equation with an imaginary

diffusion constant. That single factor of i changes the behavior completely, turning diffusion into waves (Chapter 13).

Example 3.11 (Checking solutions). (a) $u = e^{-k\pi^2 t}\sin(\pi x)$ satisfies the one-dimensional heat equation $u_t = k u_{xx}$: both sides equal $-k\pi^2 e^{-k\pi^2 t}\sin \pi x$. (b) For any twice-differentiable F, G, $u = F(x-ct) + G(x+ct)$ satisfies $u_{tt} = c^2 u_{xx}$: the general solution of the one-dimensional wave equation is a pair of shapes travelling left and right at speed c (d'Alembert). (c) $u = \ln(x^2+y^2)$ satisfies $u_{xx} + u_{yy} = 0$ away from the origin.

3.4 Differentiability and Linear Approximation

In one variable, differentiability at a means that f is well approximated near a by a line. In several variables the right notion is approximation by a linear function, whose graph is a plane. The mere existence of partial derivatives is not enough to guarantee this, because partials only see two directions.

Definition 3.12 (Differentiability). $f : \mathbb{R}^n \to \mathbb{R}$ is *differentiable* at $\mathbf{a}$ if there is a vector $\mathbf{g}$ such that

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - \mathbf{g}\cdot\mathbf{h}}{|\mathbf{h}|} = 0.$$

In that case $\mathbf{g}$ is unique, its components are the partial derivatives, and we call it the *gradient* $\nabla f(\mathbf{a}) = \langle f_{x_1}(\mathbf{a}), \dots, f_{x_n}(\mathbf{a})\rangle$.

Taking $\mathbf{h} = h\hat{\mathbf{e}}_i$ in the definition shows $g_i = \partial f/\partial x_i$, which proves the uniqueness claim. In words, $f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot\mathbf{h} + o(|\mathbf{h}|)$: the error in the linear approximation goes to zero faster than $|\mathbf{h}|$.

Theorem 3.13 (Differentiable implies continuous). *If f is differentiable at $\mathbf{a}$, it is continuous at $\mathbf{a}$.*

Proof. $f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \nabla f(\mathbf{a})\cdot\mathbf{h} + |\mathbf{h}|\,\epsilon(\mathbf{h})$ with $\epsilon(\mathbf{h}) \to 0$. Both terms tend to 0 as $\mathbf{h} \to \mathbf{0}$. ∎

Example 3.14 (Partials without differentiability). Let $f(x,y) = xy/(x^2+y^2)$ for $(x,y) \neq (0,0)$ and $f(0,0) = 0$. Since $f = 0$ on both axes, $f_x(0,0) = f_y(0,0) = 0$. But f is not even continuous at the origin (Example 3.5), so it is not differentiable there.

Theorem 3.15 (A sufficient condition). *If all partial derivatives of f exist near $\mathbf{a}$ and are continuous at $\mathbf{a}$, then f is differentiable at $\mathbf{a}$. Such functions are called* continuously differentiable, *or C^1.*

Proof. We give the proof for $n = 2$. Write $\mathbf{h} = \langle h, k\rangle$ and move from (a,b) to $(a+h, b+k)$ in two steps:

$$f(a+h,b+k) - f(a,b) = [f(a+h,b+k) - f(a,b+k)] + [f(a,b+k) - f(a,b)].$$

By the mean value theorem, the first bracket is $f_x(a+\theta_1 h,\, b+k)\, h$ and the second is $f_y(a,\, b+\theta_2 k)\, k$, for some $\theta_1, \theta_2 \in (0,1)$. Subtracting $f_x(a,b)h + f_y(a,b)k$ leaves

$$\big[f_x(a+\theta_1 h, b+k) - f_x(a,b)\big]h + \big[f_y(a,b+\theta_2 k) - f_y(a,b)\big]k.$$

Dividing by $|\mathbf{h}|$ and using $|h|, |k| \le |\mathbf{h}|$, the result is bounded by the sum of the two brackets, each of which tends to 0 by continuity of the partials. ∎

When $f(x,y)$ is differentiable at (a,b), the plane

$$z = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$$

is the *tangent plane* to the graph. The approximation $f(x,y) \approx L(x,y)$ by this linear function is the *linearization*. In differential notation, $dz = f_x\,dx + f_y\,dy$ is the *total differential*.

Example 3.16 (Linear approximation). Estimate $\sqrt{(3.02)^2 + (3.97)^2}$.

Solution. Let $f(x,y) = \sqrt{x^2+y^2}$ at $(3,4)$: $f = 5$, $f_x = 3/5$, $f_y = 4/5$. Then $f \approx 5 + \frac{3}{5}(0.02) + \frac{4}{5}(-0.03) = 4.988$. The true value is $4.98812\ldots$.

Example 3.17 (Propagation of error). The period of a pendulum is $T = 2\pi\sqrt{L/g}$. If L is measured with relative error at most 1% and g with relative error at most 0.5%, bound the relative error in T.

Solution. $\ln T = \ln 2\pi + \frac{1}{2}\ln L - \frac{1}{2}\ln g$, so $dT/T = \frac{1}{2}\,dL/L - \frac{1}{2}\,dg/g$. The worst case is $|dT/T| \le \frac{1}{2}(0.01) + \frac{1}{2}(0.005) = 0.0075$, a relative error of

0.75%. Logarithmic differentiation is the standard way to propagate relative errors through products and powers.

3.5 The Chain Rule

Theorem 3.18 (Chain rule). *Suppose $x_1, \dots, x_n$ are differentiable functions of $t_1, \dots, t_m$ and $f(x_1, \dots, x_n)$ is differentiable. Then $u = f(\mathbf{x}(\mathbf{t}))$ is differentiable, and for each j*

$$\frac{\partial u}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\,\frac{\partial x_i}{\partial t_j}.$$

In matrix form, the row vector of derivatives of u with respect to the t_j is $\nabla f^{\,T}$ times the Jacobian matrix $J = \left[\partial x_i / \partial t_j\right]$.

Proof. It suffices to treat one variable t. Write $\Delta\mathbf{x} = \mathbf{x}(t+\Delta t) - \mathbf{x}(t)$. By differentiability of f,

$$\frac{\Delta u}{\Delta t} = \nabla f \cdot \frac{\Delta \mathbf{x}}{\Delta t} + \epsilon(\Delta \mathbf{x})\,\frac{|\Delta \mathbf{x}|}{\Delta t},$$

where $\epsilon(\Delta\mathbf{x}) \to 0$ as $\Delta\mathbf{x} \to \mathbf{0}$. As $\Delta t \to 0$, $\Delta\mathbf{x}/\Delta t \to \mathbf{x}'(t)$, $|\Delta\mathbf{x}|/|\Delta t|$ stays bounded, and $\Delta\mathbf{x} \to \mathbf{0}$. The second term therefore vanishes. ∎

Example 3.19 (Polar coordinates). If $u = f(x,y)$ with $x = r\cos\theta$, $y = r\sin\theta$, then

$$u_r = f_x\cos\theta + f_y\sin\theta, \qquad u_\theta = -f_x\, r\sin\theta + f_y\, r\cos\theta.$$

Squaring and adding gives the useful identity $f_x^2 + f_y^2 = u_r^2 + \dfrac{1}{r^2}u_\theta^2$.

Example 3.20 (The Laplacian in polar coordinates). Show that $u_{xx} + u_{yy} = u_{rr} + \dfrac{1}{r}u_r + \dfrac{1}{r^2}u_{\theta\theta}$.

Solution. Differentiate the expressions of Example 3.19 again, treating f_x, f_y as functions of (x,y) and applying the chain rule to each:

$$\begin{aligned} u_{rr} &= f_{xx}\cos^2\theta + 2f_{xy}\sin\theta\cos\theta + f_{yy}\sin^2\theta,\\ u_{\theta\theta} &= r^2\big(f_{xx}\sin^2\theta - 2f_{xy}\sin\theta\cos\theta + f_{yy}\cos^2\theta\big) - r\big(f_x\cos\theta + f_y\sin\theta\big). \end{aligned}$$

The last bracket in $u_{\theta\theta}$ is u_r. Hence $u_{rr} + r^{-2}u_{\theta\theta} = f_{xx} + f_{yy} - r^{-1}u_r$, which rearranges to the claim. The analogous computation in spherical coordinates, carried out in Chapter 7, gives the operator at the heart of the hydrogen atom.

Example 3.21 (Implicit differentiation). If $F(x,y) = 0$ defines y implicitly as a differentiable function of x, differentiating $F(x, y(x)) = 0$ gives $F_x + F_y\, y' = 0$, so $\mathrm{d}y/\mathrm{d}x = -F_x/F_y$ wherever $F_y \neq 0$. Likewise, if $F(x,y,z) = 0$ defines $z(x,y)$, then $z_x = -F_x/F_z$ and $z_y = -F_y/F_z$. For $x^3 + y^3 + z^3 + 6xyz = 1$: $z_x = -(x^2 + 2yz)/(z^2 + 2xy)$.

Example 3.22 (Thermodynamic identity). If $F(x,y,z) = 0$ relates three variables, each can be regarded as a function of the other two. Then $\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1$: by Example 3.21 the product is $(-F_y/F_x)(-F_z/F_y)(-F_x/F_z) = -1$. The subscripts record which variable is held fixed. In thermodynamics, where the same quantity is a function of different sets of variables, omitting them leads to genuine errors. For an ideal gas $PV = nRT$ the identity can be checked directly.

3.6 Directional Derivatives and the Gradient

Definition 3.23 (Directional derivative). The *directional derivative* of f at $\mathbf{a}$ in the direction of a unit vector $\hat{\mathbf{u}}$ is

$$D_{\hat{\mathbf{u}}} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\hat{\mathbf{u}}) - f(\mathbf{a})}{h} = \frac{\mathrm{d}}{\mathrm{d}h} f(\mathbf{a} + h\hat{\mathbf{u}})\Big|_{h=0}.$$

Partial derivatives are the special cases $\hat{\mathbf{u}} = \hat{\mathbf{e}}_i$.

Theorem 3.24 (Directional derivatives from the gradient). *If f is differentiable at $\mathbf{a}$, then $D_{\hat{\mathbf{u}}} f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \hat{\mathbf{u}}$.*

Proof. Apply the chain rule to $g(h) = f(\mathbf{a} + h\hat{\mathbf{u}})$: $g'(0) = \nabla f(\mathbf{a}) \cdot \hat{\mathbf{u}}$. ∎

Theorem 3.25 (Geometry of the gradient). *Let f be differentiable at $\mathbf{a}$ with $\nabla f(\mathbf{a}) \neq \mathbf{0}$.*

(a) *The maximum of $D_{\hat{\mathbf{u}}} f(\mathbf{a})$ over unit vectors $\hat{\mathbf{u}}$ is $|\nabla f(\mathbf{a})|$, attained exactly when $\hat{\mathbf{u}}$ points along $\nabla f(\mathbf{a})$. The minimum is $-|\nabla f(\mathbf{a})|$, in the opposite direction.*

(b) *$\nabla f(\mathbf{a})$ is orthogonal to the level set of f through $\mathbf{a}$: if $\mathbf{r}(t)$ is any differentiable curve in the level set with $\mathbf{r}(0) = \mathbf{a}$, then $\nabla f(\mathbf{a}) \cdot \mathbf{r}'(0) = 0$.*

Proof. (a) $D_{\hat{\mathbf{u}}} f = |\nabla f| \cos\theta$, where θ is the angle between ∇f and $\hat{\mathbf{u}}$; this is largest at $\theta = 0$. Alternatively, apply Cauchy–Schwarz. (b) Differentiate $f(\mathbf{r}(t)) = k$ at $t = 0$ using the chain rule. ∎

So the gradient points uphill, as steeply as possible, and cuts every level set at a right angle. On a topographic map, water flows along $-\nabla h$, perpendicular to the contours.

Example 3.26 (Steepest ascent). For $f(x, y) = 4 - x^2 - 2y^2$ at $(1, 1)$: $\nabla f = \langle -2, -4 \rangle$, so the steepest ascent is in direction $\langle -1, -2 \rangle / \sqrt{5}$, with rate $|\nabla f| = 2\sqrt{5}$. In the direction $\langle 3, 4 \rangle / 5$ the rate of change is $(-6 - 16)/5 = -4.4$.

Definition 3.27 (Tangent plane to a level surface). If $F(x, y, z)$ is differentiable at a point P_0 of the level surface $F = k$ and $\nabla F(P_0) \neq \mathbf{0}$, the *tangent plane* at P_0 is $\nabla F(P_0) \cdot (\mathbf{r} - \mathbf{r}_0) = 0$, and the *normal line* is $\mathbf{r}_0 + t\nabla F(P_0)$.

A graph $z = f(x, y)$ is the level surface $F = f(x, y) - z = 0$, and with $\nabla F = \langle f_x, f_y, -1 \rangle$ this recovers the tangent plane of the previous section.

Example 3.28 (Tangent plane to an ellipsoid). Find the tangent plane to $\frac{x^2}{4} + y^2 + \frac{z^2}{9} = 3$ at $(-2, 1, -3)$.

Solution. $\nabla F = \langle x/2, 2y, 2z/9 \rangle = \langle -1, 2, -\frac{2}{3} \rangle$. The plane is $-(x + 2) + 2(y - 1) - \frac{2}{3}(z + 3) = 0$, or $3x - 6y + 2z + 18 = 0$.

LOOKING AHEAD. In quantum mechanics the momentum of a particle is represented by the operator $\hat{\mathbf{p}} = -i\hbar\nabla$, acting on wavefunctions. A plane wave $\psi = e^{i\mathbf{k}\cdot\mathbf{r}}$ satisfies $-i\hbar\nabla\psi = \hbar\mathbf{k}\,\psi$: it is a state of definite momentum

$\hbar\mathbf{k}$, travelling perpendicular to its wavefronts $\mathbf{k}\cdot\mathbf{r} = \text{constant}$. This is Theorem 3.25(b) in quantum dress.

3.7 The Implicit and Inverse Function Theorems

When does an equation $F(x,y) = 0$ actually define y as a function of x? The circle $x^2+y^2=1$ does so near $(0,1)$, as $y=\sqrt{1-x^2}$. It fails near $(1,0)$, where the circle is vertical and $F_y = 2y = 0$. The general principle is that a smooth equation can be solved for a variable wherever its derivative with respect to that variable is nonzero.

Theorem 3.29 (Implicit function theorem). *Let $F:\mathbb{R}^{n+1}\to\mathbb{R}$ be C^1 near $(\mathbf{a},b)$, with $F(\mathbf{a},b)=0$ and $\partial F/\partial y(\mathbf{a},b)\neq 0$. Then there are neighborhoods $U\ni\mathbf{a}$ and $W\ni b$ and a unique C^1 function $g:U\to W$ such that, for $\mathbf{x}\in U$ and $y\in W$, $F(\mathbf{x},y)=0$ if and only if $y=g(\mathbf{x})$. Moreover $\partial g/\partial x_i = -F_{x_i}/F_y$.*

Theorem 3.30 (Inverse function theorem). *Let $\mathbf{F}:\mathbb{R}^n\to\mathbb{R}^n$ be C^1 near $\mathbf{a}$ with Jacobian determinant $\det J_{\mathbf{F}}(\mathbf{a})\neq 0$. Then $\mathbf{F}$ has a C^1 inverse near $\mathbf{F}(\mathbf{a})$, and $J_{\mathbf{F}^{-1}} = J_{\mathbf{F}}^{-1}$.*

Both theorems are proved in courses on real analysis by a contraction-mapping argument; we use them freely. The inverse function theorem explains why polar coordinates fail at the origin. There $\det \partial(x,y)/\partial(r,\theta) = r$, which vanishes at $r=0$, and indeed every θ maps to the same point. The same Jacobian determinant will reappear in Chapter 5 as the scale factor in a change of variables for integrals.

Example 3.31 (Jacobian of spherical coordinates). For $(x,y,z) = (r\sin\theta\cos\phi,\ r\sin\theta\sin\phi,\ r\cos\theta)$,

$$\det\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix} = r^2\sin\theta.$$

It vanishes on the z-axis, where $\theta = 0$ or π and ϕ is undefined.

3.8 WORKED PROBLEMS

The first two problems probe the definition of differentiability from both sides. The rest exercise the chain rule, the gradient, and the implicit and inverse function theorems in combination.

Problem 3.1 Continuous, with every directional derivative, yet not differentiable

Let $g(x,y) = \dfrac{x^2 y}{x^2+y^2}$ for $(x,y) \neq (0,0)$ and $g(0,0) = 0$.

(a) Show that g is continuous at the origin.

(b) Show that the directional derivative $D_{\hat{\mathbf{u}}} g(0,0)$ exists for every unit vector $\hat{\mathbf{u}}$, and compute it.

(c) Show that g is not differentiable at the origin.

Solution. (a) Since $x^2 \leq x^2+y^2$, $|g(x,y)| \leq |y| \to 0$. By the squeeze theorem $g \to 0 = g(0,0)$.

(b) With $\hat{\mathbf{u}} = \langle u_1, u_2 \rangle$ and $u_1^2 + u_2^2 = 1$,

$$D_{\hat{\mathbf{u}}} g(0,0) = \lim_{h \to 0} \frac{g(hu_1, hu_2) - 0}{h} = \lim_{h \to 0} \frac{h^3 u_1^2 u_2}{h \cdot h^2} = u_1^2 u_2.$$

In particular $g_x(0,0) = D_{\hat{\imath}} g = 0$ and $g_y(0,0) = D_{\hat{\jmath}} g = 0$.

(c) If g were differentiable at the origin, Theorem 3.24 would give $D_{\hat{\mathbf{u}}} g(0,0) = \nabla g(0,0) \cdot \hat{\mathbf{u}} = 0$ for every $\hat{\mathbf{u}}$. But along $\hat{\mathbf{u}} = \langle 1,1 \rangle / \sqrt{2}$ the directional derivative is $\frac{1}{2} \cdot \frac{1}{\sqrt{2}} = \frac{1}{2\sqrt{2}} \neq 0$. So g is not differentiable. Geometrically, the graph of g is a cone-like surface made of straight lines through the origin (since g is homogeneous of degree 1), and these lines do not lie in a single plane. There is no tangent plane.

The lesson: existence of all directional derivatives is not enough. Differentiability requires that they depend *linearly* on the direction.

Problem 3.2 Differentiable, with discontinuous partial derivatives

Let $f(x,y) = (x^2+y^2)\sin \dfrac{1}{\sqrt{x^2+y^2}}$ for $(x,y) \neq (0,0)$ and $f(0,0) = 0$. Show that f is differentiable at the origin, but that f_x is not continuous there. Conclude that the converse of Theorem 3.15 is false.

Solution. Write $r = \sqrt{x^2+y^2}$. The candidate linear map is zero, since $f_x(0,0) = \lim_{h\to 0} h^2 \sin(1/|h|)/h = 0$, and likewise $f_y(0,0) = 0$. Differentiability at the origin requires

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - 0 - 0}{r} = \lim_{r\to 0} r \sin\frac{1}{r} = 0,$$

which holds since $|r\sin(1/r)| \le r$. So f is differentiable at the origin with $\nabla f(0,0) = \mathbf{0}$.

Away from the origin,

$$f_x = 2x\sin\frac{1}{r} + r^2\cos\frac{1}{r}\cdot\Big(-\frac{1}{r^2}\Big)\frac{x}{r} = 2x\sin\frac{1}{r} - \frac{x}{r}\cos\frac{1}{r}.$$

On the positive x-axis this is $2x\sin(1/x) - \cos(1/x)$. As $x \to 0^+$ the first term tends to 0, but the second oscillates between -1 and 1. Taking $x = 1/(2n\pi)$ gives $f_x = -1$, and $x = 1/((2n+1)\pi)$ gives $f_x = +1$. So f_x has no limit at the origin.

Thus continuity of the partials is *sufficient* for differentiability, but not necessary. The function oscillates infinitely often near the origin, but with amplitude r^2 that is crushed fast enough to leave a tangent plane.

Problem 3.3 d'Alembert's solution of the wave equation

Show that every C^2 solution of the one-dimensional wave equation $u_{tt} = c^2 u_{xx}$ on $\mathbb{R}^2$ has the form $u(x,t) = F(x-ct) + G(x+ct)$.

Solution. Change variables to $\xi = x - ct$, $\eta = x + ct$, and regard u as a function $U(\xi,\eta)$. This change is invertible, with $x = \frac{1}{2}(\xi+\eta)$ and $t = \frac{1}{2c}(\eta-\xi)$. By the chain rule (Theorem 3.18),

$$u_x = U_\xi + U_\eta, \qquad u_t = -cU_\xi + cU_\eta.$$

Applying the same rule again, and using $U_{\xi\eta} = U_{\eta\xi}$ by Clairaut's theorem,

$$u_{xx} = U_{\xi\xi} + 2U_{\xi\eta} + U_{\eta\eta}, \qquad u_{tt} = c^2\big(U_{\xi\xi} - 2U_{\xi\eta} + U_{\eta\eta}\big).$$

Therefore $u_{tt} - c^2 u_{xx} = -4c^2 U_{\xi\eta}$, and the wave equation becomes $U_{\xi\eta} = 0$. Then U_ξ does not depend on η: $U_\xi = f(\xi)$. Integrating in ξ, $U = F(\xi) + G(\eta)$ with $F' = f$, where the "constant" of integration may depend on η. Returning to x and t, $u = F(x-ct) + G(x+ct)$.

Every solution is the superposition of a profile moving right and a profile moving left, each at speed c, without change of shape. Given initial data $u(x,0) = \phi(x)$

and $u_t(x,0) = \psi(x)$, one finds $u = \frac{1}{2}[\phi(x-ct)+\phi(x+ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct}\psi(s)\,ds$. The variables $x \pm ct$ are the *characteristics* of the equation, and the computation above is the simplest case of the method of characteristics.

Problem 3.4 Second derivatives of an implicit function

The equation $e^z + xz + y = 2$ is satisfied at $(1,1,0)$.

(a) Show that it defines z as a smooth function of (x,y) near $(1,1)$.

(b) Find z_x, z_y, z_{xx}, z_{xy}, and z_{yy} at $(1,1)$.

(c) Write the second-order Taylor polynomial of z about $(1,1)$.

Solution. (a) With $F = e^z + xz + y - 2$, $F_z = e^z + x = 2 \neq 0$ at the point. By the implicit function theorem (Theorem 3.29), $z = z(x,y)$ exists and is smooth near $(1,1)$, with $z(1,1) = 0$.

(b) Differentiate the identity $e^z + xz + y = 2$, treating z as a function of x and y:

$$\partial_x: \quad (e^z + x)z_x + z = 0, \qquad\qquad \partial_y: \quad (e^z + x)z_y + 1 = 0.$$

At the point, $e^z + x = 2$ and $z = 0$, so $z_x = 0$ and $z_y = -\frac{1}{2}$. Differentiate the first identity with respect to x and to y, and the second with respect to y:

$$\begin{aligned}
\partial_x\partial_x: &\quad e^z z_x^2 + (e^z + x)z_{xx} + 2z_x = 0,\\
\partial_y\partial_x: &\quad e^z z_y z_x + (e^z + x)z_{xy} + z_y = 0,\\
\partial_y\partial_y: &\quad e^z z_y^2 + (e^z + x)z_{yy} = 0.
\end{aligned}$$

Substituting $z = 0$, $z_x = 0$, $z_y = -\frac{1}{2}$: $2z_{xx} = 0$, $2z_{xy} - \frac{1}{2} = 0$, and $\frac{1}{4} + 2z_{yy} = 0$. So $z_{xx} = 0$, $z_{xy} = \frac{1}{4}$, $z_{yy} = -\frac{1}{8}$.

(c) With $h = x - 1$ and $k = y - 1$,

$$z \approx -\frac{1}{2}k + \frac{1}{2}\left(0\cdot h^2 + 2\cdot\frac{1}{4}hk - \frac{1}{8}k^2\right) = -\frac{k}{2} + \frac{hk}{4} - \frac{k^2}{16}.$$

A numerical check, solving $e^z + xz + y = 2$ by bisection at $(1.1, 1.1)$, gives $z = -0.048163$, against the Taylor value $-0.05 + 0.0025 - 0.000625 = -0.048125$. The error, about 4×10^{-5}, is of third order in the step 0.1, as it should be.

Problem 3.5 Tangent planes with a prescribed normal

Find all points of the ellipsoid $x^2 + 2y^2 + 3z^2 = 6$ at which the tangent plane is parallel to the plane $x + y + z = 0$, and the equations of those tangent planes.

Solution. The normal at (x,y,z) is $\nabla F = \langle 2x, 4y, 6z\rangle$, and it must be parallel to $\langle 1,1,1\rangle$: $2x = 4y = 6z = \lambda$. Then $x = \lambda/2$, $y = \lambda/4$, $z = \lambda/6$, and substituting into the ellipsoid,

$$\frac{\lambda^2}{4} + \frac{\lambda^2}{8} + \frac{\lambda^2}{12} = \frac{11\lambda^2}{24} = 6, \qquad \lambda = \pm\frac{12}{\sqrt{11}}.$$

The points are $\pm\frac{1}{\sqrt{11}}(6,3,2)$, and the tangent planes are $x+y+z = \pm\frac{11}{\sqrt{11}} = \pm\sqrt{11}$.

These are the two planes with normal $\langle 1,1,1\rangle$ that just touch the ellipsoid. So $\sqrt{11}$ is the maximum of $x+y+z$ on the ellipsoid, and $-\sqrt{11}$ the minimum. This is the method of Lagrange multipliers of Chapter 4 in disguise: the condition "$\nabla F \parallel \nabla(x+y+z)$" is the Lagrange condition. Cauchy–Schwarz confirms the value: $x+y+z = 1\cdot x + \frac{1}{\sqrt{2}}(\sqrt{2}y) + \frac{1}{\sqrt{3}}(\sqrt{3}z) \le \sqrt{1+\frac{1}{2}+\frac{1}{3}}\cdot\sqrt{6} = \sqrt{11}$.

Problem 3.6 The path of steepest ascent

The temperature on a plate is $T(x,y) = 100e^{-x^2-2y^2}$. A heat-seeking particle starts at $(1,1)$ and always moves in the direction of ∇T. Find its path. At the start, what is the maximum rate of increase of T?

Solution. $\nabla T = -100e^{-x^2-2y^2}\,\langle 2x, 4y\rangle$. At $(1,1)$ this is $-100e^{-3}\,\langle 2,4\rangle$, with magnitude $|\nabla T| = 200\sqrt{5}\,e^{-3} \approx 22.3$ degrees per unit length. That is the maximum rate (Theorem 3.25).

The path $\mathbf{r}(t)$ satisfies $\mathbf{r}'(t) \parallel \nabla T$. Only the direction matters, and ∇T points along $-\langle 2x, 4y\rangle$. So the path satisfies

$$\frac{dy}{dx} = \frac{4y}{2x} = \frac{2y}{x} \quad\Longrightarrow\quad \ln|y| = 2\ln|x| + C \quad\Longrightarrow\quad y = Cx^2.$$

Through $(1,1)$: $y = x^2$. The particle moves along the parabola $y = x^2$ toward the origin, the hottest point.

Note that the path is *not* the straight line to the origin, because the temperature falls off faster in y than in x. The gradient is perpendicular to the level curves, the ellipses $x^2+2y^2 = c$, and the parabolas $y = Cx^2$ are their orthogonal trajectories. Check: on $y = x^2$ the tangent is $\langle 1, 2x\rangle$, and the ellipse $x^2+2y^2 = c$ has normal $\langle 2x, 4y\rangle = \langle 2x, 4x^2\rangle = 2x\,\langle 1, 2x\rangle$, which is parallel to it.

Problem 3.7 A map that is locally but not globally invertible

Let $\mathbf{F}(x,y) = (u,v) = (x^2 - y^2,\ 2xy)$.

(a) Find where the inverse function theorem applies.

(b) Near $(x,y)=(1,1)$, find $\partial x/\partial u$, $\partial x/\partial v$, $\partial y/\partial u$, $\partial y/\partial v$ at $(u,v)=(0,2)$.

(c) Show that $\mathbf{F}$ is not injective on $\mathbb{R}^2\setminus\{\mathbf{0}\}$, even though its Jacobian never vanishes there.

Solution. (a) $D\mathbf{F}=\begin{pmatrix}2x & -2y\\ 2y & 2x\end{pmatrix}$, with determinant $4(x^2+y^2)$. This is nonzero except at the origin, so by Theorem 3.30 $\mathbf{F}$ has a smooth local inverse near every other point.

(b) At $(1,1)$, $D\mathbf{F}=\begin{pmatrix}2 & -2\\ 2 & 2\end{pmatrix}$, with determinant 8. The derivative of the inverse is the inverse matrix:

$$\begin{pmatrix}x_u & x_v\\ y_u & y_v\end{pmatrix}=\frac{1}{8}\begin{pmatrix}2 & 2\\ -2 & 2\end{pmatrix}=\begin{pmatrix}\frac{1}{4} & \frac{1}{4}\\ -\frac{1}{4} & \frac{1}{4}\end{pmatrix}.$$

Note that $x_u=\frac{1}{4}$ is *not* $1/u_x=\frac{1}{2}$. A partial derivative of the inverse requires inverting the whole matrix, because x_u holds v fixed, while u_x holds y fixed.

(c) $\mathbf{F}(-x,-y)=\mathbf{F}(x,y)$, so every value is taken at least twice. In complex notation (Chapter 8), $\mathbf{F}$ is $z\mapsto z^2$, which doubles angles. Each local inverse is a branch of $\sqrt{w}$, but no single continuous inverse exists on all of $\mathbb{R}^2\setminus\{\mathbf{0}\}$: going once around the origin in the (u,v)-plane takes a branch of $\sqrt{w}$ to its negative. The inverse function theorem is strictly local.

Problem 3.8 The reflection property of the ellipse, by gradients

Let F_1 and F_2 be points in the plane, and let $d_i(P)=|P-F_i|$.

(a) Show that $\nabla d_i=\hat{\mathbf{u}}_i$, the unit vector pointing from F_i to P, for $P\neq F_i$.

(b) Show that at every point of the ellipse $d_1+d_2=2a$, the normal line bisects the angle F_1PF_2. Conclude that a light ray leaving F_1 reflects off the ellipse to pass through F_2.

(c) State and prove the analogous property of the hyperbola $d_1-d_2=\pm 2a$.

Solution. (a) With $\mathbf{r}=P$ and $\mathbf{f}=F_i$, $d_i=|\mathbf{r}-\mathbf{f}|=\big((x-f_1)^2+(y-f_2)^2\big)^{1/2}$. Then $\partial_x d_i=(x-f_1)/d_i$ and $\partial_y d_i=(y-f_2)/d_i$, so $\nabla d_i=(\mathbf{r}-\mathbf{f})/|\mathbf{r}-\mathbf{f}|=\hat{\mathbf{u}}_i$. Geometrically: the distance from a fixed point increases fastest when moving directly away from it, at unit rate.

(b) The ellipse is a level curve of $g = d_1 + d_2$, so its normal at P is $\nabla g = \hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2$ (Theorem 3.25). This is nonzero unless $\hat{\mathbf{u}}_1 = -\hat{\mathbf{u}}_2$, which happens only on the segment between the foci, not on the ellipse when a exceeds half the focal distance. The sum of two unit vectors bisects the angle between them, since $(\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2) \cdot \hat{\mathbf{u}}_1 = 1 + \hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_2 = (\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2) \cdot \hat{\mathbf{u}}_2$. So the normal makes equal angles with $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$, the directions from the foci to P.

A ray from F_1 arrives at P traveling along $\hat{\mathbf{u}}_1$. The law of reflection (angle of incidence equals angle of reflection, measured from the normal) sends it back along the mirror image of $-\hat{\mathbf{u}}_1$ across the normal, which by the bisection property is $-\hat{\mathbf{u}}_2$: the ray heads straight for F_2. Whispering galleries and lithotripters, which focus shock waves from one focus onto a kidney stone at the other, use this property.

(c) For the hyperbola, the level set of $d_1 - d_2$ has normal $\nabla(d_1 - d_2) = \hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2$. Since $(\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2) \cdot (\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2) = 1 - 1 = 0$, the normal is perpendicular to the bisector $\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2$. So the *tangent* line bisects the angle F_1PF_2. In optical terms, light converging toward one focus is redirected by a hyperbolic mirror toward the other focus, the principle of the Cassegrain telescope. The whole argument used nothing but the gradient of a distance function. The classical synthetic proofs require separate constructions for each conic.

EXERCISES

A. COMPUTATION

3.1 Find and sketch the domain of $f(x, y) = \ln(9 - x^2 - 9y^2)$.

3.2 Find the domain and range of $f(x, y, z) = \sqrt{1 - x^2 - y^2 - z^2}$ and describe its level surfaces.

3.3 Sketch a contour map of $f(x, y) = y - x^2$ with levels $-2, -1, 0, 1, 2$.

3.4 Describe the level surfaces of $f(x, y, z) = x^2 + y^2 - z^2$ for $k > 0$, $k = 0$, $k < 0$.

3.5 Evaluate $\displaystyle\lim_{(x,y)\to(1,2)} \frac{x^2y - xy^2}{x^2 + y^2}$.

3.6 Show that $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{x^2 + y^2}$ does not exist.

3.7 Evaluate $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2 + y^2}$ or show it does not exist.

3.8 Evaluate $\lim_{(x,y)\to(0,0)} \dfrac{xy^2}{x^2+y^4}$ or show it does not exist.

3.9 Evaluate $\lim_{(x,y)\to(0,0)} \dfrac{\sin(x^2+y^2)}{x^2+y^2}$.

3.10 Evaluate $\lim_{(x,y,z)\to(0,0,0)} \dfrac{xyz}{x^2+y^2+z^2}$ or show it does not exist.

3.11 Find all first partial derivatives of $f(x,y) = x^3y^2 - 4xy + \arctan(y/x)$.

3.12 Find f_x, f_y, f_z for $f(x,y,z) = xe^{yz} + \ln(x^2+z^2)$.

3.13 Find all second partial derivatives of $f(x,y) = \sin(xy)$ and verify Clairaut's theorem.

3.14 Find f_{xyz} for $f(x,y,z) = e^{xyz}$.

3.15 Show that $u = e^{-\alpha^2k^2t}\cos(kx)$ satisfies $u_t = \alpha^2 u_{xx}$.

3.16 Show that $u = 1/\sqrt{x^2+y^2+z^2}$ satisfies Laplace's equation away from the origin.

3.17 Find the tangent plane to $z = x^2 + xy + 3y^2$ at $(1,1,5)$.

3.18 Find the linearization of $f(x,y) = \sqrt{20 - x^2 - 7y^2}$ at $(2,1)$ and use it to estimate $f(1.95, 1.08)$.

3.19 The dimensions of a closed box are 80, 60, 50 cm, each with possible error 0.2 cm. Use differentials to estimate the maximum error in its computed volume.

3.20 Use differentials to estimate the maximum percentage error in the resistance $R = \rho L/(\pi r^2)$ if L and r have errors of at most 1% and 2%.

3.21 If $z = x^2y + 3xy^4$ with $x = \sin 2t$ and $y = \cos t$, find $\mathrm{d}z/\mathrm{d}t$ at $t = 0$.

3.22 If $u = x^2 + yz$, $x = pr\cos\theta$, $y = pr\sin\theta$, $z = p + r$, find $\partial u/\partial p$ at $(p, r, \theta) = (2, 3, 0)$.

3.23 Find $\partial z/\partial x$ and $\partial z/\partial y$ if $xyz = \cos(x + y + z)$.

3.24 The radius of a cylinder grows at 1.2 cm/s while its height shrinks at 3 cm/s. How fast is the volume changing when $r = 80$ cm and $h = 150$ cm?

3.25 Find the directional derivative of $f(x,y) = x^2y^3 - 4y$ at $(2,-1)$ in the direction of $\langle 2, 5\rangle$.

3.26 Find the directional derivative of $f(x,y,z) = xe^{y} + ye^{z} + ze^{x}$ at $(0,0,0)$ in the direction of $\langle 5,1,-2\rangle$.

3.27 Find the maximum rate of change of $f(x,y,z) = x\ln(yz)$ at $(1,2,\tfrac{1}{2})$ and the direction in which it occurs.

3.28 The temperature in a region is $T(x,y,z) = 200e^{-x^2-3y^2-9z^2}$. In which direction does it increase fastest at $(2,-1,2)$? Find the maximum rate.

3.29 Find the tangent plane and normal line to $x^2 + 2y^2 + 3z^2 = 21$ at $(4,-1,1)$.

3.30 Find the points on the hyperboloid $x^2 + y^2 - z^2 = 1$ where the tangent plane is parallel to the plane $x + y - z = 0$.

3.31 Find the Jacobian determinant $\partial(x,y)/\partial(u,v)$ for $x = u^2 - v^2$, $y = 2uv$, and say where the map is locally invertible.

3.32 Find $\mathrm{d}y/\mathrm{d}x$ for $y^5 + x^2y^3 = 1 + ye^{x^2}$ at the point $(0,1)$.

B. THEORY AND PROOF

3.33 Prove that $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^3+y^3}{x^2+y^2} = 0$ directly from the definition.

3.34 For $f(x,y) = xy(x^2-y^2)/(x^2+y^2)$, $f(0,0)=0$, show $f_{xy}(0,0) = -1$ and $f_{yx}(0,0) = 1$. Which hypothesis of Theorem 3.10 fails?

3.35 Show that $f(x,y) = \sqrt{|xy|}$ has $f_x(0,0) = f_y(0,0) = 0$ but is not differentiable at the origin.

3.36 Prove that if f is differentiable and $\nabla f \equiv \mathbf{0}$ on an open connected set, then f is constant there.

3.37 A function is *homogeneous of degree n* if $f(t\mathbf{x}) = t^n f(\mathbf{x})$ for $t>0$. Prove *Euler's theorem*: $\mathbf{x}\cdot\nabla f = n f$.

3.38 Show that in polar coordinates $\nabla f = f_r\,\hat{\mathbf{r}} + \dfrac{1}{r} f_\theta\,\hat{\boldsymbol{\theta}}$.

3.39 Show that if $u(x,y)$ satisfies Laplace's equation, so does $v(x,y) = u\!\left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2}\right)$ away from the origin. (Use Example 3.20.)

3.40 Prove that the wave equation $u_{tt} = c^2 u_{xx}$ becomes $u_{\xi\eta} = 0$ under $\xi = x - ct$, $\eta = x + ct$. Deduce d'Alembert's general solution.

3.41 Prove the product rule for gradients, $\nabla(fg) = f\nabla g + g\nabla f$, and compute $\nabla(r^n)$ where $r = |\mathbf{r}|$.

C. CHALLENGE

3.42 Show that $u(x,t) = \dfrac{1}{\sqrt{4\pi k t}} e^{-x^2/(4kt)}$ satisfies the heat equation for $t > 0$, that $\int_{-\infty}^{\infty} u\,dx = 1$ for every $t > 0$ (use Chapter 5 or assume $\int e^{-x^2}dx = \sqrt{\pi}$), and describe what happens as $t \to 0^+$. This is the *heat kernel*; its quantum analog is the free-particle propagator.

3.43 Compute the Laplacian of $u(r)$, a function of $r = |\mathbf{r}|$ only, in $\mathbb{R}^n$. Show that $\nabla^2 u = u'' + \frac{n-1}{r}u'$ and find all radial harmonic functions in $\mathbb{R}^2$ and $\mathbb{R}^3$.

3.44 Let $f(x,y) = x^2 y/(x^4+y^2)$ for $(x,y) \neq (0,0)$ and $f(0,0) = 0$. Show that every directional derivative $D_{\hat{\mathbf{u}}} f(0,0)$ exists, and that f is nevertheless not continuous at the origin.

CHAPTER 4

Optimization in Several Variables

LEARNING OBJECTIVES

- State and use Taylor's theorem to second order for functions of several variables.
- Classify quadratic forms as definite, semidefinite, or indefinite, using principal minors or eigenvalues.
- Find and classify critical points with the second derivative test, and recognize when it is inconclusive.
- Find absolute extrema on closed, bounded regions.
- Derive and apply the method of Lagrange multipliers with one or several constraints, and interpret the multiplier.
- Apply these tools to least squares, Fermat's principle, statistical mechanics, and the Rayleigh quotient.

Nature optimizes. Light takes the path of least time, a soap film minimizes its area, a system in thermal equilibrium maximizes its entropy, and the ground state of a quantum system minimizes its energy. Engineering and economics optimize on purpose. This chapter develops the mathematics of finding maxima and minima of functions of several variables, both freely and subject to constraints. The second-derivative analysis rests on *quadratic forms*, whose classification by eigenvalues is a first glimpse of the spectral theory that dominates Chapters 9 and 15.

4.1 Taylor's Theorem in Several Variables

Recall Taylor's theorem in one variable. If g is twice continuously differentiable on an interval containing 0 and 1,

$$g(1) = g(0) + g'(0) + \tfrac{1}{2}g''(c) \quad \text{for some } c \in (0,1).$$

To extend this to $f : \mathbb{R}^n \to \mathbb{R}$, restrict f to a line. Fix $\mathbf{a}$ and a displacement $\mathbf{h}$, and set $g(t) = f(\mathbf{a} + t\mathbf{h})$. By the chain rule,

$$g'(t) = \sum_i h_i \, \partial_i f(\mathbf{a} + t\mathbf{h}), \qquad g''(t) = \sum_{i,j} h_i h_j \, \partial_i \partial_j f(\mathbf{a} + t\mathbf{h}),$$

where $\partial_i = \partial/\partial x_i$. The second derivatives assemble into a symmetric matrix (by Clairaut's theorem).

Definition 4.1 (Hessian). The *Hessian matrix* of a C^2 function f at $\mathbf{a}$ is the $n \times n$ symmetric matrix $H_f(\mathbf{a}) = \big[\partial_i \partial_j f(\mathbf{a})\big]$. For $n = 2$, $H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}$.

Theorem 4.2 (Second-order Taylor formula). *If f is C^2 on an open convex set containing $\mathbf{a}$ and $\mathbf{a} + \mathbf{h}$, then*

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h} + \tfrac{1}{2}\, \mathbf{h}^{T} H_f(\mathbf{c})\, \mathbf{h}$$

for some point $\mathbf{c}$ on the segment from $\mathbf{a}$ to $\mathbf{a} + \mathbf{h}$. Consequently $f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h} + \tfrac{1}{2}\mathbf{h}^{T} H_f(\mathbf{a})\mathbf{h} + o(|\mathbf{h}|^2)$.

Proof. Apply the one-variable theorem to $g(t) = f(\mathbf{a} + t\mathbf{h})$, with $g'(0) = \nabla f(\mathbf{a}) \cdot \mathbf{h}$ and $g''(c) = \mathbf{h}^{T} H_f(\mathbf{a} + c\mathbf{h})\mathbf{h}$. For the second statement, write $H_f(\mathbf{c}) = H_f(\mathbf{a}) + E$. Continuity of the second partials makes every entry of E tend to 0 as $\mathbf{h} \to \mathbf{0}$, and $|\mathbf{h}^{T} E \mathbf{h}| \le \|E\|\,|\mathbf{h}|^2$. ∎

Example 4.3 (A second-order expansion). Expand $f(x,y) = e^{x}\cos y$ to second order about $(0,0)$.

Solution. $f = 1$, $f_x = 1$, $f_y = 0$, $f_{xx} = 1$, $f_{xy} = 0$, $f_{yy} = -1$ at the origin. So $e^{x}\cos y \approx 1 + x + \tfrac{1}{2}(x^2 - y^2)$. This agrees with multiplying the one-variable series $(1 + x + \tfrac{x^2}{2})(1 - \tfrac{y^2}{2})$ and discarding terms of degree above 2.

Higher-order terms follow the same pattern. The kth-order term is

$$\frac{1}{k!}\sum h_{i_1}\cdots h_{i_k}\,\partial_{i_1}\cdots\partial_{i_k} f(\mathbf{a}) = \frac{1}{k!}(\mathbf{h}\cdot\nabla)^k f(\mathbf{a}).$$

Formally,

$$f(\mathbf{a}+\mathbf{h}) = \sum_{k=0}^{\infty}\frac{1}{k!}(\mathbf{h}\cdot\nabla)^k f(\mathbf{a}) = e^{\mathbf{h}\cdot\nabla} f(\mathbf{a}), \tag{4.1}$$

valid for analytic f. The operator $e^{\mathbf{h}\cdot\nabla}$ translates a function by $\mathbf{h}$.

LOOKING AHEAD. Equation (4.1) says that ∇ *generates translations*. In quantum mechanics, where $\hat{\mathbf{p}} = -i\hbar\nabla$, the translation operator becomes $e^{-i\mathbf{h}\cdot\hat{\mathbf{p}}/\hbar}$. This is the precise sense in which momentum is "the generator of translations," and it is why conservation of momentum follows from translation symmetry (Chapter 15).

4.2 Quadratic Forms

Definition 4.4 (Quadratic form and definiteness). A *quadratic form* on $\mathbb{R}^n$ is $Q(\mathbf{h}) = \mathbf{h}^T A\mathbf{h} = \sum_{i,j} a_{ij}h_i h_j$ with A symmetric. It is

- *positive definite* if $Q(\mathbf{h}) > 0$ for all $\mathbf{h} \neq \mathbf{0}$;
- *negative definite* if $Q(\mathbf{h}) < 0$ for all $\mathbf{h} \neq \mathbf{0}$;
- *indefinite* if Q takes both positive and negative values;
- *positive (negative) semidefinite* if $Q \geq 0$ (≤ 0) for all $\mathbf{h}$, but Q is not definite.

Theorem 4.5 (Two-variable criterion). *Let $Q(h,k) = ah^2 + 2bhk + ck^2$, with matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ and $D = \det A = ac - b^2$. Then Q is positive definite iff $a > 0$ and $D > 0$; negative definite iff $a < 0$ and $D > 0$; and indefinite iff $D < 0$.*

Proof. If $a \neq 0$, completing the square gives

$$Q = a\left(h + \frac{b}{a}k\right)^2 + \frac{D}{a}k^2.$$

If $D > 0$, both coefficients a and D/a have the sign of a, and Q vanishes only when $k = 0$ and $h = 0$. If $D < 0$, the two coefficients have opposite signs, and

choosing $k=0$ or $h=-bk/a$ produces values of both signs. If $a=0$ and $D<0$, then $b\neq 0$ and $Q=k(2bh+ck)$ clearly changes sign. ∎

For $n\times n$ matrices the corresponding test uses the *leading principal minors* $\Delta_1=a_{11}$, $\Delta_2=\det\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$, …, $\Delta_n=\det A$.

Theorem 4.6 (Sylvester's criterion and the eigenvalue test). *For symmetric A, the following are equivalent: (i) Q is positive definite; (ii) every eigenvalue of A is positive; (iii) every leading principal minor Δ_k is positive. Further, Q is negative definite iff $(-1)^k\Delta_k>0$ for all k, iff every eigenvalue is negative. If A has eigenvalues of both signs, Q is indefinite.*

The equivalence (i) $\iff$ (ii) follows from the spectral theorem (Theorem 9.27). A symmetric matrix can be diagonalized by a rotation, $A=R\,\mathrm{diag}(\lambda_1,\dots,\lambda_n)R^{T}$, so $Q=\sum\lambda_i u_i^2$ in rotated coordinates $\mathbf{u}=R^{T}\mathbf{h}$. We prove that theorem in Chapter 9 and use the result now. This rotation is exactly the change of axes that brings a quadric surface to the standard forms of Table 1.1.

Example 4.7 (Classifying a form). $Q=2x^2+2y^2+2z^2+2xy+2yz$ has matrix $\begin{pmatrix}2&1&0\\1&2&1\\0&1&2\end{pmatrix}$ with minors $2,3,4$, all positive, so Q is positive definite. Its eigenvalues are $2-\sqrt{2}$, 2, and $2+\sqrt{2}$, all positive, as Sylvester's criterion predicts.

4.3 Local Extrema

Definition 4.8 (Local and absolute extrema). f has a *local maximum* at $\mathbf{a}$ if $f(\mathbf{x})\le f(\mathbf{a})$ for all $\mathbf{x}$ in some neighborhood of $\mathbf{a}$, and an *absolute maximum* if the inequality holds on the whole domain. Minima are defined likewise. A point where $\nabla f=\mathbf{0}$, or where ∇f fails to exist, is a *critical point*.

Theorem 4.9 (Fermat). *If f has a local extremum at an interior point $\mathbf{a}$ and is differentiable there, then $\nabla f(\mathbf{a})=\mathbf{0}$.*

Proof. Each function $g_i(t)=f(\mathbf{a}+t\hat{\mathbf{e}}_i)$ has a local extremum at $t=0$, so $g_i'(0)=$

$\partial_i f(\mathbf{a}) = 0$. ∎

A critical point need not be an extremum. A point where $\nabla f = \mathbf{0}$ but which is neither a local maximum nor a local minimum is a *saddle point*; $f = x^2 - y^2$ at the origin is the model case.

Theorem 4.10 (Second derivative test). *Let f be C^2 near a critical point $\mathbf{a}$ with $\nabla f(\mathbf{a}) = \mathbf{0}$, and let $H = H_f(\mathbf{a})$.*

(a) *If H is positive definite, f has a strict local minimum at $\mathbf{a}$.*

(b) *If H is negative definite, f has a strict local maximum at $\mathbf{a}$.*

(c) *If H is indefinite, $\mathbf{a}$ is a saddle point.*

(d) *If H is only semidefinite, the test is inconclusive.*

For $n = 2$, with $D = f_{xx}f_{yy} - f_{xy}^2$: a minimum if $D > 0$ and $f_{xx} > 0$; a maximum if $D > 0$ and $f_{xx} < 0$; a saddle if $D < 0$; no conclusion if $D = 0$.

Proof. (a) A continuous function on the unit sphere attains a minimum, so $m = \min_{|\mathbf{u}|=1} \mathbf{u}^T H \mathbf{u} > 0$, and by scaling $\mathbf{h}^T H \mathbf{h} \geq m|\mathbf{h}|^2$. By Theorem 4.2,

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \tfrac{1}{2}\mathbf{h}^T H \mathbf{h} + o(|\mathbf{h}|^2) \geq \tfrac{1}{2} m|\mathbf{h}|^2 + o(|\mathbf{h}|^2),$$

which is positive for all small $\mathbf{h} \neq \mathbf{0}$. (b) Apply (a) to $-f$. (c) Choose $\mathbf{u}$ and $\mathbf{w}$ with $\mathbf{u}^T H \mathbf{u} > 0 > \mathbf{w}^T H \mathbf{w}$. Along the line $\mathbf{a} + t\mathbf{u}$, f exceeds $f(\mathbf{a})$ for small $t \neq 0$; along $\mathbf{a} + t\mathbf{w}$ it falls below. (d) The functions $x^2 + y^4$, $x^2 - y^4$, and x^2 all have $H = \operatorname{diag}(2, 0)$ at the origin, and they have respectively a strict minimum, a saddle, and a non-strict minimum. ∎

Example 4.11 (Classifying critical points). Find and classify the critical points of $f(x, y) = x^3 - 3xy + y^3$.

Solution. $f_x = 3x^2 - 3y = 0$ and $f_y = -3x + 3y^2 = 0$ give $y = x^2$ and $x = y^2$, so $x = x^4$: $x = 0$ or $x = 1$. The critical points are $(0,0)$ and $(1,1)$. With $f_{xx} = 6x$, $f_{xy} = -3$, $f_{yy} = 6y$, we get $D = 36xy - 9$. At $(0,0)$, $D = -9 < 0$: a saddle. At $(1,1)$, $D = 27 > 0$ and $f_{xx} = 6 > 0$: a local minimum with $f(1,1) = -1$. It is not an absolute minimum, since $f(x, 0) = x^3 \to -\infty$.

Example 4.12 (A degenerate critical point). The *monkey saddle* $f = x^3 - 3xy^2$ has a single critical point at the origin with $H = 0$. In polar coordinates $f = r^3 \cos 3\theta$, which changes sign six times around the origin: three "valleys" for the legs, and the tail. Third-order information is needed to see this.

Example 4.13 (Nearest point). Find the point of the plane $x + 2y + z = 4$ nearest $(1, 0, -2)$.

Solution. Minimize the squared distance $d^2 = (x-1)^2 + y^2 + (z+2)^2$ with $z = 4 - x - 2y$: $g(x,y) = (x-1)^2 + y^2 + (6 - x - 2y)^2$. Setting $g_x = 2(x-1) - 2(6-x-2y) = 0$ and $g_y = 2y - 4(6-x-2y) = 0$ gives $x = \frac{11}{6}$, $y = \frac{5}{3}$, so $z = -\frac{7}{6}$. The nearest point is $(\frac{11}{6}, \frac{5}{3}, -\frac{7}{6})$, at distance $\sqrt{25/6} = 5/\sqrt{6}$. This agrees with Theorem 1.33: $|1 + 0 - 2 - 4|/\sqrt{6} = 5/\sqrt{6}$.

4.4 Absolute Extrema

Theorem 4.14 (Extreme value theorem). *A continuous function on a closed, bounded subset of $\mathbb{R}^n$ attains an absolute maximum and an absolute minimum.*

A set is *closed* if it contains its boundary points and *bounded* if it fits in some ball. The proof uses the Bolzano–Weierstrass theorem and belongs to analysis. Both hypotheses are needed: $f(x,y) = x$ has no maximum on the open disk, and none on the closed half-plane $x \geq 0$.

To find the absolute extrema of a continuous f on a closed bounded region D: (1) find the values of f at critical points in the interior of D; (2) find the extreme values of f on the boundary, usually by parametrizing it or by Lagrange multipliers; (3) the largest and smallest of these values are the absolute maximum and minimum.

Example 4.15 (Extrema on a triangle). Find the extreme values of $f(x,y) = x^2 + y^2 - xy - x - y$ on the triangle with vertices $(0,0)$, $(3,0)$, $(0,3)$.

Solution. Interior: $f_x = 2x - y - 1 = 0$ and $f_y = 2y - x - 1 = 0$ give $(1,1)$, with $f = -1$. On the edge $y = 0$, $0 \leq x \leq 3$: $f = x^2 - x$, with values 0 at $x = 0$, $-\frac{1}{4}$ at $x = \frac{1}{2}$, and 6 at $x = 3$. The edge $x = 0$ is the same by symmetry. On the edge $x + y = 3$, set $y = 3 - x$: $f = 3x^2 - 9x + 6$, with minimum $-\frac{3}{4}$ at $x = \frac{3}{2}$ and value 6 at the ends. So the absolute maximum is 6, at $(3,0)$

and $(0,3)$, and the absolute minimum is -1, at $(1,1)$.

4.5 LAGRANGE MULTIPLIERS

Often the variables are not free but must satisfy a constraint $g(\mathbf{x}) = c$. At a constrained extremum, f cannot increase to first order in any direction tangent to the constraint surface. Hence ∇f has no component tangent to the surface. That is, ∇f must be normal to it, parallel to ∇g.

Theorem 4.16 (Lagrange multipliers). *Let f, g be C^1 on an open set in $\mathbb{R}^n$, and suppose f restricted to the level set $S = \{g = c\}$ has a local extremum at $\mathbf{a}$, where $\nabla g(\mathbf{a}) \neq \mathbf{0}$. Then there is a scalar λ, the* Lagrange multiplier, *such that*

$$\nabla f(\mathbf{a}) = \lambda \nabla g(\mathbf{a}).$$

Proof. Let $\mathbf{r}(t)$ be any C^1 curve in S with $\mathbf{r}(0) = \mathbf{a}$. Then $f(\mathbf{r}(t))$ has a local extremum at $t = 0$, so $\nabla f(\mathbf{a}) \cdot \mathbf{r}'(0) = 0$. By the implicit function theorem (Theorem 3.29), near $\mathbf{a}$ the set S is the graph of a C^1 function of $n-1$ of the variables. Therefore the velocities $\mathbf{r}'(0)$ of curves in S through $\mathbf{a}$ fill out the entire tangent space $T = \{\mathbf{v} : \nabla g(\mathbf{a}) \cdot \mathbf{v} = 0\}$, which has dimension $n-1$. So $\nabla f(\mathbf{a})$ is orthogonal to the hyperplane T, whose orthogonal complement is spanned by $\nabla g(\mathbf{a})$. ∎

The method: solve the $n+1$ equations $\nabla f = \lambda \nabla g$, $g = c$ for the $n+1$ unknowns $(\mathbf{x}, \lambda)$, then compare the values of f at the solutions, together with any points where $\nabla g = \mathbf{0}$. Equivalently, find the critical points of the *Lagrangian* $\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda\,(g(\mathbf{x}) - c)$.

Example 4.17 (Largest box in an ellipsoid). Find the volume of the largest rectangular box, with edges parallel to the axes, inscribed in $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Solution. Maximize $V = 8xyz$ $(x, y, z > 0)$ subject to the constraint. The Lagrange conditions are $8yz = 2\lambda x/a^2$, $8xz = 2\lambda y/b^2$, $8xy = 2\lambda z/c^2$. Multiplying them by x, y, z respectively shows $\frac{x^2}{a^2} = \frac{y^2}{b^2} = \frac{z^2}{c^2}$, and the constraint makes each equal to $\frac{1}{3}$. So $V_{\max} = 8 \cdot \frac{abc}{3\sqrt{3}} = \frac{8abc}{3\sqrt{3}}$. For a sphere of radius R this is a cube of volume $8R^3/(3\sqrt{3})$, about 37% of the sphere's volume.

Example 4.18 (Cauchy–Schwarz from Lagrange). Maximize $f(\mathbf{x}) = \mathbf{c}\cdot\mathbf{x}$ subject to $|\mathbf{x}|^2 = 1$, for fixed $\mathbf{c} \neq \mathbf{0}$.

Solution. $\nabla f = \mathbf{c}$ and $\nabla g = 2\mathbf{x}$, so $\mathbf{c} = 2\lambda\mathbf{x}$ and $\mathbf{x} = \pm\mathbf{c}/|\mathbf{c}|$. The maximum is $|\mathbf{c}|$. Hence $\mathbf{c}\cdot\mathbf{x} \le |\mathbf{c}||\mathbf{x}|$ for all unit $\mathbf{x}$, and by scaling for all $\mathbf{x}$. This reproves Theorem 1.13.

4.5.1 THE MEANING OF THE MULTIPLIER

Theorem 4.19 (Sensitivity). *Let $M(c)$ be the extreme value of f on $\{g = c\}$, attained at $\mathbf{a}(c)$ with multiplier $\lambda(c)$, all depending differentiably on c. Then $M'(c) = \lambda(c)$.*

Proof. $M(c) = f(\mathbf{a}(c))$, so $M' = \nabla f\cdot\mathbf{a}' = \lambda\nabla g\cdot\mathbf{a}'$. Differentiating $g(\mathbf{a}(c)) = c$ gives $\nabla g\cdot\mathbf{a}' = 1$. ∎

The multiplier is the rate at which the optimal value changes as the constraint is relaxed. In economics it is the *shadow price* of a resource. In statistical mechanics it is the inverse temperature, as the next example shows.

Example 4.20 (The Boltzmann distribution). A system can occupy states $1,\dots,N$ with energies $E_1,\dots,E_N$. Find the probabilities p_i that maximize the entropy $S = -\sum p_i \ln p_i$ subject to $\sum p_i = 1$ and a fixed mean energy $\sum p_i E_i = U$.

Solution. With two constraints there are two multipliers (see the next subsection). The condition is $\nabla S = \alpha\nabla(\sum p_i) + \beta\nabla(\sum p_i E_i)$, that is,

$$-\ln p_i - 1 = \alpha + \beta E_i \qquad\Longrightarrow\qquad p_i = e^{-1-\alpha}e^{-\beta E_i}.$$

Normalization gives $p_i = e^{-\beta E_i}/Z$ with $Z = \sum_j e^{-\beta E_j}$, the *partition function*. Physically $\beta = 1/k_B T$. Thermal equilibrium is the maximum-entropy distribution, and temperature is the Lagrange multiplier that enforces fixed energy. Planck's resolution of the blackbody problem (Chapter 12) starts from exactly this distribution, applied to quantized energies $E_n = nh\nu$.

4.5.2 SEVERAL CONSTRAINTS

Theorem 4.21 (Lagrange multipliers, k constraints). *If f has a local extremum at $\mathbf{a}$ on $\{g_1 = c_1, \dots, g_k = c_k\}$, and $\nabla g_1(\mathbf{a}), \dots, \nabla g_k(\mathbf{a})$ are linearly independent, then $\nabla f(\mathbf{a}) = \sum_{j=1}^{k} \lambda_j \nabla g_j(\mathbf{a})$ for some scalars λ_j.*

Example 4.22 (Two constraints). Find the extreme values of $f = x + 2y + 3z$ on the curve where the plane $x - y + z = 1$ meets the cylinder $x^2 + y^2 = 1$.

Solution. $\nabla f = \lambda \nabla g_1 + \mu \nabla g_2$ reads $1 = \lambda + 2\mu x$, $2 = -\lambda + 2\mu y$, $3 = \lambda$. So $\lambda = 3$, $x = -1/\mu$, $y = 5/(2\mu)$. The cylinder gives $\mu^2 = 29/4$, so $\mu = \pm\sqrt{29}/2$ and $(x, y) = \mp(2, -5)/\sqrt{29}$, with $z = 1 - x + y$. Then $f = x + 2y + 3(1 - x + y) = 3 - 2x + 5y = 3 \pm \sqrt{29}$. The maximum is $3 + \sqrt{29}$ and the minimum is $3 - \sqrt{29}$.

4.6 APPLICATIONS

4.6.1 LEAST SQUARES

Given data $(x_1, y_1), \dots, (x_n, y_n)$, the *least squares line* $y = mx + b$ minimizes $E(m, b) = \sum_{i=1}^{n} (y_i - mx_i - b)^2$. Setting $E_m = E_b = 0$ gives the *normal equations*

$$m \sum x_i^2 + b \sum x_i = \sum x_i y_i, \qquad m \sum x_i + nb = \sum y_i.$$

The Hessian is $2\begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix}$. Its determinant is $4\big(n \sum x_i^2 - (\sum x_i)^2\big) > 0$ unless all the x_i are equal (this is Cauchy–Schwarz again), so the critical point is a minimum. In matrix form, the best-fitting $\mathbf{c}$ in $X\mathbf{c} \approx \mathbf{y}$ satisfies $X^T X \mathbf{c} = X^T \mathbf{y}$. Geometrically, $X\mathbf{c}$ is the orthogonal projection of $\mathbf{y}$ onto the column space of X: least squares is the projection of Theorem 1.18 in n dimensions.

Example 4.23 (Fitting a line). For the points $(0,1)$, $(1,3)$, $(2,4)$, $(3,4)$: $\sum x = 6$, $\sum x^2 = 14$, $\sum y = 12$, $\sum xy = 23$, $n = 4$. The normal equations $14m + 6b = 23$ and $6m + 4b = 12$ give $m = 1$ and $b = 1.5$, so the line is $y = x + 1.5$.

4.6.2 FERMAT'S PRINCIPLE AND SNELL'S LAW

Example 4.24 (Refraction). Light travels from $A = (0, a)$ in a medium with speed v_1 to $B = (d, -b)$ in a medium with speed v_2, crossing the interface $y = 0$ at $(x, 0)$. Fermat's principle says that light takes the path of least time. Show that this implies Snell's law.

Solution. The travel time is $T(x) = \sqrt{a^2 + x^2}/v_1 + \sqrt{b^2 + (d-x)^2}/v_2$. Setting $T'(x) = 0$:

$$\frac{x}{v_1\sqrt{a^2 + x^2}} = \frac{d - x}{v_2\sqrt{b^2 + (d-x)^2}} \quad \Longleftrightarrow \quad \frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2},$$

where θ_1, θ_2 are the angles from the normal. With $n = c/v$ this is $n_1 \sin\theta_1 = n_2 \sin\theta_2$. Since $T'' > 0$, the critical point is the minimum.

Fermat's principle is the optical prototype of the *principle of least action* in mechanics, and Feynman's path-integral formulation of quantum mechanics explains why both hold. A quantum particle explores every path, and the amplitudes of paths near the stationary one add constructively.

4.6.3 THE RAYLEIGH QUOTIENT

Theorem 4.25 (Extremes of a quadratic form). *For a symmetric matrix A, the maximum and minimum of $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ on the unit sphere $|\mathbf{x}| = 1$ are the largest and smallest eigenvalues of A, attained at corresponding unit eigenvectors. Equivalently, $\lambda_{\min} \le \dfrac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \le \lambda_{\max}$ for all $\mathbf{x} \neq \mathbf{0}$.*

Proof. The sphere is closed and bounded, so the extrema exist. Lagrange's condition with $g = \mathbf{x}^T\mathbf{x}$ is $2A\mathbf{x} = 2\lambda\mathbf{x}$, so every constrained critical point is a unit eigenvector, and there $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \lambda\mathbf{x}^T\mathbf{x} = \lambda$. Hence the extreme values of Q are eigenvalues, and every eigenvalue occurs as a critical value. The ratio $\mathbf{x}^T A\mathbf{x}/\mathbf{x}^T\mathbf{x}$ is the *Rayleigh quotient*. ∎

Example 4.26 (Principal axes). Find the maximum and minimum of $Q = 5x^2 + 4xy + 2y^2$ on $x^2 + y^2 = 1$.

Solution. $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$ has characteristic polynomial $\lambda^2 - 7\lambda + 6 =$

$(\lambda-1)(\lambda-6)$. The maximum is 6, at $\pm\frac{1}{\sqrt{5}}\langle 2,1\rangle$; the minimum is 1, at $\pm\frac{1}{\sqrt{5}}\langle 1,-2\rangle$. The level curve $Q=1$ is an ellipse with semi-axes $1/\sqrt{6}$ and 1 along these orthogonal directions.

LOOKING AHEAD. Theorem 4.25 is the finite-dimensional form of the *variational principle* of quantum mechanics. For the Hamiltonian operator $\hat{H}$ and any normalized trial wavefunction ψ, $\langle\psi|\hat{H}|\psi\rangle \geq E_0$, the ground-state energy, with equality only for the ground state itself. Minimizing over a family of trial functions gives upper bounds on E_0, and this is how most of quantum chemistry estimates molecular energies (Chapter 15).

4.7 Worked Problems

These problems cover the classification of critical points in two and three variables, absolute extrema on a compact set, and constrained optimization, including a case where the Lagrange condition finds a point that is not the extremum sought.

Problem 4.1 The folium function

Find and classify the critical points of $f(x,y)=x^3+y^3-3xy$. Does f have a global minimum?

Solution. The critical points satisfy $f_x=3x^2-3y=0$ and $f_y=3y^2-3x=0$. So $y=x^2$ and $x=y^2=x^4$, giving $x(x^3-1)=0$. The only real solutions are $x=0$ and $x=1$, so the critical points are $(0,0)$ and $(1,1)$. The Hessian is

$$H=\begin{pmatrix}6x & -3\\ -3 & 6y\end{pmatrix}.$$

At $(0,0)$, $\det H=-9<0$: a saddle, with $f=0$. At $(1,1)$, $\det H=36-9=27>0$ and $f_{xx}=6>0$: a local minimum, with $f=-1$.

The local minimum is not global: $f(x,0)=x^3\to-\infty$ as $x\to-\infty$. The zero level set $x^3+y^3=3xy$ is the *folium of Descartes*, which crosses itself at the saddle point $(0,0)$. That is typical: at a nondegenerate saddle, the level curve through it consists of two curves crossing transversally, along the directions where the quadratic form of H vanishes, here $-6xy=0$, the coordinate axes. The folium's loop lies in the first quadrant and encloses the minimum at $(1,1)$.

Problem 4.2 Classification in three variables

Find and classify the critical points of $g(x,y,z) = x^3 + y^2 + z^2 + 12xy + 2z$.

Solution. Set the gradient to zero: $3x^2 + 12y = 0$, $2y + 12x = 0$, $2z + 2 = 0$. The second gives $y = -6x$, and then the first gives $3x^2 - 72x = 0$, so $x = 0$ or $x = 24$. The critical points are $(0,0,-1)$ and $(24,-144,-1)$. The Hessian is

$$H = \begin{pmatrix} 6x & 12 & 0 \\ 12 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad \Delta_1 = 6x, \quad \Delta_2 = 12x - 144, \quad \Delta_3 = 2\Delta_2.$$

At $(0,0,-1)$: $\Delta_1 = 0$ and $\Delta_2 = -144$. A negative Δ_2 means the upper-left 2×2 block, a quadratic form in (x,y), is indefinite, so H is indefinite and the point is a *saddle*. (Here Sylvester's criterion, Theorem 4.6, cannot be applied mechanically since $\Delta_1 = 0$, but an indefinite principal block settles it.) The eigenvalues are 2 and $1 \pm \sqrt{145}$, of both signs, confirming this.

At $(24,-144,-1)$: $\Delta_1 = 144$, $\Delta_2 = 144$, $\Delta_3 = 288$, all positive, so H is positive definite and the point is a *local minimum*, with $g = -6913$. It is not a global minimum, since $g(x,0,0) = x^3 \to -\infty$ as $x \to -\infty$. A cubic function has no global extrema, whatever its local structure.

Problem 4.3 Absolute extrema on a triangle

Find the absolute maximum and minimum of $g(x,y) = x^2 + y^2 - 2x - 4y$ on the closed triangle T with vertices $(0,0)$, $(4,0)$, $(0,4)$.

Solution. T is closed and bounded and g is continuous, so both extrema exist. Candidates are interior critical points, critical points of g restricted to each edge, and the vertices.

Interior. $g_x = 2x - 2 = 0$ and $g_y = 2y - 4 = 0$ give $(1,2)$, which is inside T since $1 + 2 < 4$. There $g = -5$.

Edges. On $y = 0$: $g = x^2 - 2x$, critical at $x = 1$, giving $g(1,0) = -1$. On $x = 0$: $g = y^2 - 4y$, critical at $y = 2$, giving $g(0,2) = -4$. On $x + y = 4$: $g = x^2 + (4-x)^2 - 2x - 4(4-x) = 2x^2 - 6x$, critical at $x = \frac{3}{2}$, giving $g(\frac{3}{2}, \frac{5}{2}) = -\frac{9}{2}$.

Vertices. $g(0,0) = 0$, $g(4,0) = 8$, $g(0,4) = 0$.

The absolute minimum is -5 at $(1,2)$, and the absolute maximum is 8 at $(4,0)$. A geometric reading confirms this: $g = (x-1)^2 + (y-2)^2 - 5$, so $g + 5$ is the squared distance from $(1,2)$. The minimum is at that point, and the maximum is at the point of T farthest from it, which must be a vertex. Checking distances, $(4,0)$ at $\sqrt{13}$ is farther than $(0,4)$ at $\sqrt{5}$ or $(0,0)$ at $\sqrt{5}$.

Problem 4.4 Nearest and farthest points on a tilted ellipse

Find the points of the ellipse $x^2 + xy + y^2 = 3$ nearest to and farthest from the origin.

Solution. Extremize $f = x^2 + y^2$ subject to $g = x^2 + xy + y^2 = 3$. The Lagrange condition $\nabla f = \lambda \nabla g$ reads

$$2x = \lambda(2x + y), \qquad 2y = \lambda(x + 2y).$$

In matrix form this says $\mathbf{x}$ is an eigenvector of $A = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$, the matrix of the quadratic form g, with eigenvalue $\mu = 1/\lambda$. The eigenvalues are $\frac{3}{2}$, with eigenvector $\langle 1, 1 \rangle$, and $\frac{1}{2}$, with eigenvector $\langle 1, -1 \rangle$.

On an eigenvector with eigenvalue μ, $g = \mu|\mathbf{x}|^2 = 3$, so $|\mathbf{x}|^2 = 3/\mu$. Along $\langle 1, 1 \rangle$: $|\mathbf{x}|^2 = 2$, at the points $\pm(1, 1)$. Along $\langle 1, -1 \rangle$: $|\mathbf{x}|^2 = 6$, at the points $\pm(\sqrt{3}, -\sqrt{3})$. The nearest points are $\pm(1, 1)$, at distance $\sqrt{2}$, and the farthest are $\pm(\sqrt{3}, -\sqrt{3})$, at distance $\sqrt{6}$.

This is Theorem 4.25 at work. The semi-axes of the ellipse $\mathbf{x}^T A \mathbf{x} = 3$ point along the eigenvectors of A and have lengths $\sqrt{3/\mu}$. The same calculation, in infinite dimensions, will give the ground state of a quantum system as the minimum of an energy functional (Chapter 15).

Problem 4.5 A weighted product

Maximize $P = xy^2z^3$ subject to $x + y + z = 6$, $x, y, z > 0$.

Solution. Since $P > 0$ on the constraint set, it is equivalent and easier to maximize $\ln P = \ln x + 2 \ln y + 3 \ln z$. The Lagrange conditions $\nabla \ln P = \lambda \nabla (x + y + z)$ are

$$\frac{1}{x} = \lambda, \qquad \frac{2}{y} = \lambda, \qquad \frac{3}{z} = \lambda.$$

So $x : y : z = 1 : 2 : 3$, and with $x + y + z = 6$, $(x, y, z) = (1, 2, 3)$, where $P = 1 \cdot 4 \cdot 27 = 108$.

It is a maximum: $\ln P \to -\infty$ as any variable tends to 0, so the maximum over the open triangle $x + y + z = 6$, $x, y, z > 0$ is attained in its interior, at a Lagrange point, and there is only one. In general, $x^a y^b z^c$ on $x + y + z = s$ is maximized when each variable is proportional to its exponent: $x = \frac{as}{a+b+c}$, and so on. This is the weighted AM–GM inequality,

$$\left(\frac{x}{1}\right)^1 \left(\frac{y}{2}\right)^2 \left(\frac{z}{3}\right)^3 \leq \left(\frac{x + y + z}{6}\right)^6 = 1,$$

which gives $xy^2z^3 \le 1 \cdot 4 \cdot 27$ directly. Taking logarithms to turn products into sums is the same move that produces the Boltzmann distribution in Example 4.20.

Problem 4.6 When Lagrange finds the wrong extremum

Find the maximum and minimum of $f = x + y$ on the curve $xy = 1$, $x, y > 0$.

Solution. The condition $\nabla f = \lambda \nabla (xy)$ gives $1 = \lambda y$, $1 = \lambda x$, so $x = y = 1$ and $f = 2$. Is it the maximum or the minimum? On the curve, $f = x + \frac{1}{x}$, which tends to ∞ as $x \to 0^+$ or $x \to \infty$. So f has no maximum, and $f = 2$ is the minimum, by AM–GM or since $x + \frac{1}{x} - 2 = \frac{(x-1)^2}{x} \ge 0$.

The Lagrange condition is necessary, not sufficient. It locates candidates, and the problem's global structure must decide which are extrema, if any. Here the constraint set is not compact, and the extreme value theorem does not apply. A blind reading of "the Lagrange point is the maximum" gives a wrong answer.

The multiplier still has its meaning (Theorem 4.19). For the constraint $xy = c$ the minimum is $2\sqrt{c}$, and $\frac{\mathrm{d}}{\mathrm{d}c}\left(2\sqrt{c}\right) = 1/\sqrt{c}$, which at $c = 1$ is $1 = \lambda$.

Problem 4.7 Fitting a plane by least squares

Fit $z = a + bx + cy$ to the data $(0,0,1)$, $(1,0,3)$, $(0,1,2)$, $(1,1,5)$ by minimizing the sum of squared vertical errors.

Solution. Minimize $E(a,b,c) = \sum_i (a + bx_i + cy_i - z_i)^2$. Setting $\partial E/\partial a$, $\partial E/\partial b$, $\partial E/\partial c$ to zero gives the *normal equations* $X^T X \mathbf{fi} = X^T \mathbf{z}$, where X has rows $(1, x_i, y_i)$:

$$X^T X = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}, \qquad X^T \mathbf{z} = \begin{pmatrix} 11 \\ 8 \\ 7 \end{pmatrix}.$$

Solving: subtracting the third equation from the second gives $b - c = 1$. Then the first equation, $4a + 2b + 2c = 11$, and the second, $2a + 2b + c = 8$, with $b = c + 1$, become $4a + 4c = 9$ and $2a + 3c = 6$. Hence $c = \frac{3}{2}$, $b = \frac{5}{2}$, $a = \frac{3}{4}$:

$$z = \frac{3}{4} + \frac{5}{2}x + \frac{3}{2}y.$$

The residuals are $-\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, -\frac{1}{4}$. They sum to zero, as the first normal equation requires, and are orthogonal to the columns x_i and y_i, as the others require. $E = \frac{1}{4}$ is the minimum: E is a convex quadratic, since its Hessian $2X^T X$ is positive definite (its leading principal minors are 8, 16, 32). This is the projection of $\mathbf{z}$ onto the column space of X, a theme taken up in Chapter 9.

EXERCISES

A. COMPUTATION

4.1 Find the second-order Taylor polynomial of $f(x,y) = \sin x \sin y$ about $(0,0)$.

4.2 Find the second-order Taylor polynomial of $f(x,y) = \sqrt{1+x+2y}$ about $(0,0)$.

4.3 Find the second-order Taylor polynomial of $f(x,y) = e^{x-y}$ about $(1,1)$.

4.4 Classify $Q(x,y) = 3x^2 - 4xy + y^2$ and $Q(x,y) = x^2 + xy + y^2$.

4.5 Classify $Q = x^2 + 2y^2 + 3z^2 + 2xy$ using leading principal minors.

4.6 Find the eigenvalues of $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ and classify the associated quadratic form.

4.7 Find and classify the critical points of $f = x^2 + xy + y^2 + 3x - 3y + 4$.

4.8 Find and classify the critical points of $f = x^4 + y^4 - 4xy + 1$.

4.9 Find and classify the critical points of $f = xy(1 - x - y)$.

4.10 Find and classify the critical points of $f = e^x \cos y$.

4.11 Find and classify the critical points of $f = x^3 + y^3 - 3x - 12y + 20$.

4.12 Find and classify the critical points of $f = 2x^3 + xy^2 + 5x^2 + y^2$.

4.13 Find the shortest distance from $(2,0,-3)$ to the plane $x + y + z = 1$ by minimizing a function of two variables.

4.14 Find three positive numbers with sum 100 whose product is maximal.

4.15 Find the absolute extrema of $f = x^2 + y^2 - 2x$ on the closed triangle with vertices $(2,0)$, $(0,2)$, $(0,-2)$.

4.16 Find the absolute extrema of $f = x^2 + 2y^2 - x$ on the closed disk $x^2 + y^2 \le 1$.

4.17 Find the absolute extrema of $f = xy^2$ on $D = \{x \ge 0,\ y \ge 0,\ x^2 + y^2 \le 3\}$.

4.18 Use Lagrange multipliers to find the extreme values of $f = x^2 y$ on $x^2 + 2y^2 = 6$.

4.19 Use Lagrange multipliers to find the extreme values of $f = x + y + z$ on the sphere $x^2 + y^2 + z^2 = 12$.

4.20 Find the points on the surface $z^2 = xy + 1$ closest to the origin.

4.21 Find the dimensions of the open-top box of volume $32\,\text{m}^3$ with least surface area.

4.22 Find the maximum of $f = xyz$ on the plane $x + 2y + 3z = 6$ in the first octant.

4.23 A rectangle is inscribed in the ellipse $x^2/a^2 + y^2/b^2 = 1$ with sides parallel to the axes. Find its maximum area.

4.24 Find the extreme values of $f = x^2 + y^2 + z^2$ subject to $x - y = 1$ and $y^2 - z^2 = 1$.

4.25 Find the maximum and minimum of $Q = 3x^2 + 2xy + 3y^2$ on $x^2 + y^2 = 1$ and identify the principal axes.

4.26 Find the least squares line for $(1,2)$, $(2,3)$, $(3,5)$, $(4,7)$.

4.27 A Cobb–Douglas production function $P = bL^{\alpha}K^{1-\alpha}$ is maximized subject to the budget $mL + nK = p$. Show that $L = \alpha p/m$ and $K = (1-\alpha)p/n$.

4.28 For three energy levels $0, \varepsilon, 2\varepsilon$ in thermal equilibrium, compute the Boltzmann probabilities at $\beta\varepsilon = \ln 2$.

B. THEORY AND PROOF

4.29 Prove that a symmetric 2×2 matrix with $a > 0$ and $ac - b^2 > 0$ has two positive eigenvalues.

4.30 Show by example that the second derivative test fails when $D = 0$: find three functions with $D = 0$ at the origin having a minimum, a maximum, and a saddle there.

4.31 Prove that if f is C^2 and convex in the sense that H_f is positive semidefinite everywhere on $\mathbb{R}^n$, then every critical point of f is an absolute minimum.

4.32 Prove the AM–GM inequality $\sqrt[n]{x_1 \cdots x_n} \le \dfrac{x_1 + \cdots + x_n}{n}$ for positive x_i by maximizing $x_1 \cdots x_n$ subject to $\sum x_i = S$.

4.33 Show that among all triangles with given perimeter, the equilateral triangle has the largest area. (Use Heron's formula.)

4.34 Prove that the least squares normal equations $X^T X\mathbf{c} = X^T\mathbf{y}$ characterize the $\mathbf{c}$ for which $\mathbf{y} - X\mathbf{c}$ is orthogonal to every column of X.

4.35 Show that the Boltzmann distribution maximizes entropy, not merely makes it stationary, by showing S is a concave function of $(p_1, \dots, p_N)$.

4.36 Prove Theorem 4.19 for k constraints: $\partial M/\partial c_j = \lambda_j$.

C. CHALLENGE

4.37 (Heron's problem and reflection.) Find the shortest path from A to B, on the same side of a line, that touches the line. Show that the angle of incidence equals the angle of reflection, and relate the result to Fermat's principle.

4.38 Show that the Hessian of $f(x,y) = \ln(e^x + e^y)$ is positive semidefinite everywhere but never positive definite, and explain why in terms of the behavior of f along the direction $\langle 1, 1\rangle$.

4.39 A particle in a box with sides a, b, c has ground-state energy $E = \dfrac{\pi^2\hbar^2}{2m}\left(\dfrac{1}{a^2} + \dfrac{1}{b^2} + \dfrac{1}{c^2}\right)$. Among boxes of fixed volume V, which minimizes E? Find the minimal energy.

4.40 (Variational estimate.) For the quadratic form with matrix $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$, evaluate the Rayleigh quotient at $\mathbf{x} = \langle 1, 1, 1\rangle$ and at $\mathbf{x} = \left\langle 1, \sqrt{2}, 1\right\rangle$. Compare with the exact smallest and largest eigenvalues $2 \mp \sqrt{2}$ and comment.

CHAPTER 5

Multiple Integrals

LEARNING OBJECTIVES

- Define double and triple integrals as limits of Riemann sums and evaluate them as iterated integrals using Fubini's theorem.
- Integrate over general regions, choosing and reversing the order of integration.
- Evaluate integrals in polar, cylindrical, and spherical coordinates, and compute the Gaussian integral exactly.
- Prove and apply the change-of-variables formula with the Jacobian determinant.
- Compute mass, centers of mass, moments of inertia, probabilities, and expectation values.
- Use the Gamma function, Gaussian moments, and differentiation under the integral sign.

Integration in one variable adds up a quantity spread along a line. Here we add up quantities spread over regions of the plane and of space: the mass of a plate with variable density, the charge in a cloud, the probability that a particle is found in some region. That last application is the one this book is building toward. In quantum mechanics, $|\psi(\mathbf{r})|^2$ is a probability density, and almost every physical prediction is an integral of it, often in spherical coordinates and often involving a Gaussian. This chapter develops all the integration technique Part III needs.

5.1 Double Integrals over Rectangles

Let f be bounded on a closed rectangle $R = [a, b] \times [c, d]$. Partition $[a, b]$ into m subintervals of width $\Delta x = (b - a)/m$ and $[c, d]$ into n subintervals of width

$\Delta y = (d-c)/n$. This divides R into mn subrectangles R_{ij} of area $\Delta A = \Delta x\,\Delta y$. Choose a *sample point* (x_{ij}^*, y_{ij}^*) in each.

Definition 5.1 (Double integral). The *double integral* of f over R is

$$\iint_R f(x,y)\,dA = \lim_{m,n\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n} f(x_{ij}^*, y_{ij}^*)\,\Delta A,$$

if the limit exists and is the same for every choice of sample points. Then f is *integrable* on R.

If $f \geq 0$, each term $f(x_{ij}^*, y_{ij}^*)\Delta A$ is the volume of a thin column, and the integral is the volume of the solid under the graph. Every continuous function on R is integrable. So is every bounded function whose discontinuities lie on finitely many smooth curves. The integral is linear, monotone ($f \leq g \Rightarrow \iint f \leq \iint g$), and additive over regions that overlap only along their boundaries.

The definition is rarely used for computation. Instead we slice. Fix x and integrate in y to get the area $A(x) = \int_c^d f(x,y)\,dy$ of a cross-section; then integrate the cross-sectional areas.

Theorem 5.2 (Fubini). *If f is continuous on $R = [a,b]\times[c,d]$, then*

$$\iint_R f\,dA = \int_a^b\!\int_c^d f(x,y)\,dy\,dx = \int_c^d\!\int_a^b f(x,y)\,dx\,dy.$$

More generally, the conclusion holds whenever f is bounded on R, its discontinuities lie on finitely many smooth curves, and the iterated integrals exist.

Example 5.3 (An iterated integral). $\iint_{[0,2]\times[1,2]} (x - 3y^2)\,dA = \int_0^2 \Big[xy - y^3\Big]_{y=1}^{2} dx = \int_0^2 (x-7)\,dx = 2 - 14 = -12$. Integrating in the other order gives the same result.

Example 5.4 (Separable integrands). If $f(x,y) = g(x)h(y)$, then $\iint_R f\,dA = \left(\int_a^b g\,dx\right)\left(\int_c^d h\,dy\right)$. For example, $\iint_{[0,\pi/2]^2} \sin x\cos y\,dA = 1\cdot 1 = 1$. Integrals of products of single-variable functions over rectangles factor. The Gaussian integral in Section 5.3.1 exploits exactly this.

COMMON ERROR (when Fubini fails). The hypotheses matter. For $f(x,y) = \dfrac{x^2-y^2}{(x^2+y^2)^2}$ on $(0,1]^2$, one computes $\int_0^1\int_0^1 f\,dy\,dx = \frac{\pi}{4}$ but $\int_0^1\int_0^1 f\,dx\,dy = -\frac{\pi}{4}$ (Item 5.34). The function is unbounded near the origin, and $\iint |f|\,dA = \infty$. When an integrand is unbounded or a region is infinite, check absolute integrability before swapping orders. Fubini's theorem for Lebesgue integrals, often called the Fubini–Tonelli theorem, makes this precise.

5.2 Double Integrals over General Regions

For a bounded region D, enclose D in a rectangle R and define $\iint_D f\,dA = \iint_R F\,dA$, where $F = f$ on D and $F = 0$ outside. Two kinds of regions are especially convenient.

- *Type I*: $D = \{a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$. Then $\displaystyle\iint_D f\,dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f\,dy\,dx.$
- *Type II*: $D = \{c \le y \le d,\ h_1(y) \le x \le h_2(y)\}$. Then $\displaystyle\iint_D f\,dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f\,dx\,dy.$

The inner limits depend on the outer variable, and the outer limits are constants. With $f = 1$, $\iint_D 1\,dA$ is the area of D.

Example 5.5 (Region between curves). Evaluate $\iint_D (x+2y)\,dA$ where D is bounded by $y = 2x^2$ and $y = 1 + x^2$.

Solution. The curves meet where $2x^2 = 1 + x^2$, at $x = \pm 1$, and $1 + x^2 \ge 2x^2$ in between. So

$$\begin{aligned}\int_{-1}^{1}\int_{2x^2}^{1+x^2} (x+2y)\,dy\,dx &= \int_{-1}^{1} \left(-3x^4 - x^3 + 2x^2 + x + 1\right) dx \\ &= -\tfrac{6}{5} + \tfrac{4}{3} + 2 = \tfrac{32}{15}.\end{aligned}$$

Example 5.6 (Reversing the order). Evaluate $\displaystyle\int_0^1 \int_x^1 \sin\left(y^2\right) dy\,dx.$

Solution. The inner integral has no elementary antiderivative. The region is $0 \le x \le y \le 1$, which as a type II region is $0 \le y \le 1$, $0 \le x \le y$. Then

$$\int_0^1 \int_0^y \sin(y^2)\,dx\,dy = \int_0^1 y\sin(y^2)\,dy = \tfrac{1}{2}(1-\cos 1).$$

Reversing the order converts an impossible integral into a substitution.

Example 5.7 (Volume of a tetrahedron). Find the volume under the plane $x + 2y + z = 2$ in the first octant.

Solution. Over $D = \{0 \le x \le 2,\ 0 \le y \le 1 - x/2\}$, the height is $z = 2 - x - 2y$: $V = \int_0^2 \int_0^{1-x/2} (2 - x - 2y)\,dy\,dx = \int_0^2 \left(1 - x + \tfrac{x^2}{4}\right)dx = \tfrac{2}{3}$. As a check, this is $\tfrac{1}{3}\cdot$base area $\cdot$ height $= \tfrac{1}{3}\cdot 1 \cdot 2$.

Theorem 5.8 (Mean value theorem for double integrals). *If f is continuous on a closed, bounded, connected region D of area $A(D) > 0$, then $\iint_D f\,dA = f(\mathbf{p})\,A(D)$ for some $\mathbf{p} \in D$.*

Proof. Let m and M be the minimum and maximum of f on D. Then $mA \le \iint_D f\,dA \le MA$, so the average $\frac{1}{A}\iint_D f\,dA$ lies in $[m, M]$, and the intermediate value theorem on the connected set D gives $\mathbf{p}$. ∎

5.3 Polar Coordinates

Regions bounded by circles and rays are awkward in Cartesian coordinates and natural in polar ones. A *polar rectangle* $\{r_1 \le r \le r_2,\ \theta_1 \le \theta \le \theta_2\}$ is an annular sector. A small one with sides Δr and $\Delta\theta$ has area

$$\tfrac{1}{2}(r + \Delta r)^2\Delta\theta - \tfrac{1}{2}r^2\Delta\theta = r\,\Delta r\,\Delta\theta + \tfrac{1}{2}(\Delta r)^2\Delta\theta,$$

and to first order that is $r\,\Delta r\,\Delta\theta$. The factor r is the Jacobian of Section 5.7.

Theorem 5.9 (Polar coordinates). *If f is continuous on $D = \{\alpha \le \theta \le \beta,\ h_1(\theta) \le r \le h_2(\theta)\}$ with $0 \le \beta - \alpha \le 2\pi$, then*

$$\iint_D f(x,y)\,dA = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta.$$

Example 5.10 (Volume of a paraboloid cap). Find the volume below $z = 1 - x^2 - y^2$ and above $z = 0$.

Solution. $V = \int_0^{2\pi}\int_0^1 (1-r^2)\, r\, dr\, d\theta = 2\pi\left[\frac{r^2}{2} - \frac{r^4}{4}\right]_0^1 = \frac{\pi}{2}$.

Example 5.11 (Inside one curve, outside another). Find the area inside the cardioid $r = 1 + \cos\theta$ and outside the circle $r = 1$.

Solution. The curves meet where $\cos\theta = 0$, so the region is $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$, $1 \le r \le 1 + \cos\theta$:

$$A = \int_{-\pi/2}^{\pi/2} \tfrac{1}{2}\left[(1+\cos\theta)^2 - 1\right] d\theta = \int_{-\pi/2}^{\pi/2} \left(\cos\theta + \tfrac{1}{2}\cos^2\theta\right) d\theta = 2 + \frac{\pi}{4}.$$

5.3.1 THE GAUSSIAN INTEGRAL

The function e^{-x^2} has no elementary antiderivative, yet its integral over the whole line can be computed exactly. The trick is to square it and pass to two dimensions. This integral is the most important definite integral in physics. It governs the normal distribution, the heat kernel, the ground state of the harmonic oscillator, and every wave packet in Chapters 11 and 13.

Theorem 5.12 (Gaussian integral). *$\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$, and more generally $\displaystyle\int_{-\infty}^{\infty} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}$ for $a > 0$.*

Proof. Let $I = \int_{-\infty}^{\infty} e^{-x^2}dx$, which converges by comparison with $e^{-|x|}$ for $|x| \ge 1$. Since the integrand is positive, Fubini's theorem applies on the plane, and by Example 5.4,

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-x^2}dx \int_{-\infty}^{\infty} e^{-y^2}dy = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dA \\ &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\, r\, dr\, d\theta = 2\pi \cdot \tfrac{1}{2} = \pi. \end{aligned}$$

The factor r from the polar area element makes the inner integral elementary. Since $I > 0$, $I = \sqrt{\pi}$. The general case follows from the substitution $x = u/\sqrt{a}$. ∎

To be fully rigorous about the improper double integral, compare the integral over the square $[-R,R]^2$ with the integrals over the inscribed disk of radius R and the circumscribed disk of radius $R\sqrt{2}$. These equal $\pi(1-e^{-R^2})$ and $\pi(1-e^{-2R^2})$, and both tend to π.

Corollary 5.13 (The normal distribution). $\dfrac{1}{\sigma\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)}\,dx = 1$ *for all* $\mu \in \mathbb{R}$, $\sigma > 0$.

5.4 Applications of Double Integrals

5.4.1 Mass, Center of Mass, and Moments

A thin plate (*lamina*) occupying D with area density $\rho(x,y)$ has

$$m = \iint_D \rho\,dA, \qquad \bar{x} = \frac{1}{m}\iint_D x\rho\,dA, \qquad \bar{y} = \frac{1}{m}\iint_D y\rho\,dA.$$

Its *moments of inertia* about the x-axis, the y-axis, and the origin are

$$I_x = \iint_D y^2\rho\,dA, \qquad I_y = \iint_D x^2\rho\,dA, \qquad I_0 = I_x + I_y = \iint_D (x^2+y^2)\rho\,dA.$$

The moment of inertia measures resistance to angular acceleration, playing the role of mass in rotational motion: the rotational kinetic energy is $\tfrac{1}{2}I\omega^2$.

Example 5.14 (Center of mass of a half-disk). A half-disk $x^2+y^2 \le a^2$, $y \ge 0$, has density proportional to the distance from the origin, $\rho = kr$. Find its center of mass.

Solution. $m = \int_0^{\pi}\int_0^{a} kr\cdot r\,dr\,d\theta = \frac{k\pi a^3}{3}$. By symmetry $\bar{x} = 0$, and $\iint y\rho\,dA = \int_0^{\pi}\int_0^{a}(r\sin\theta)(kr)r\,dr\,d\theta = k\cdot 2\cdot\frac{a^4}{4} = \frac{ka^4}{2}$. So $\bar{y} = \frac{3a}{2\pi} \approx 0.477a$. For uniform density the answer is $\frac{4a}{3\pi} \approx 0.424a$. Weighting the rim more heavily moves the center outward.

5.4.2 Probability

Definition 5.15 (Joint density). A pair of random variables (X,Y) has *joint probability density* $f(x,y) \ge 0$ if $P\big((X,Y)\in D\big) = \iint_D f\,dA$ for every region D, and so $\iint_{\mathbb{R}^2} f\,dA = 1$. The *expected value* of a function $g(X,Y)$ is $E[g] = \iint_{\mathbb{R}^2} g(x,y)f(x,y)\,dA$.

Example 5.16 (Independent exponentials). Waiting times X and Y are independent with density $f(x,y) = \lambda\mu e^{-\lambda x - \mu y}$ for $x, y \geq 0$. Find $P(X < Y)$.

Solution. $P(X < Y) = \int_0^\infty \int_x^\infty \lambda\mu e^{-\lambda x} e^{-\mu y}\, dy\, dx = \int_0^\infty \lambda e^{-(\lambda+\mu)x} dx = \dfrac{\lambda}{\lambda+\mu}$.

LOOKING AHEAD. For a particle in two dimensions with wavefunction $\psi(x,y)$, the *Born rule* says $|\psi(x,y)|^2$ is the joint probability density of its position. Normalization is $\iint |\psi|^2\, dA = 1$, and the expected value of an observable such as x^2 is $\iint x^2 |\psi|^2\, dA$. Every normalization integral in Part III is an integral of exactly this kind.

5.5 Triple Integrals

Triple integrals over boxes $B = [a,b] \times [c,d] \times [r,s]$ are defined by Riemann sums exactly as before, and Fubini's theorem allows evaluation as iterated integrals in any of the six orders. For a general bounded solid E, the most common description is

$$E = \{(x,y,z) : (x,y) \in D,\ u_1(x,y) \leq z \leq u_2(x,y)\},$$
$$\iiint_E f\, dV = \iint_D \Big[\int_{u_1(x,y)}^{u_2(x,y)} f\, dz \Big] dA,$$

where D is the projection of E onto the xy-plane. Solids may equally be projected onto the yz- or xz-planes. The volume of E is $\iiint_E 1\, dV$. Mass, center of mass, and moments of inertia are defined as for laminas, with

$$I_z = \iiint_E (x^2 + y^2)\rho\, dV, \quad \text{and likewise } I_x, I_y.$$

Example 5.17 (A tetrahedron). Evaluate $\iiint_E z\, dV$ where E is the tetrahedron bounded by the coordinate planes and $x + y + z = 1$.

Solution. $\displaystyle\int_0^1 \int_0^{1-x} \int_0^{1-x-y} z\, dz\, dy\, dx = \int_0^1 \int_0^{1-x} \tfrac{1}{2}(1-x-y)^2\, dy\, dx = \int_0^1 \tfrac{1}{6}(1-x)^3 dx = \tfrac{1}{24}$. Since the volume is $\tfrac{1}{6}$, the centroid has $\bar{z} = \frac{1/24}{1/6} = \frac{1}{4}$.

Example 5.18 (Changing the order). Rewrite $\int_0^1 \int_0^{x^2} \int_0^y f\,dz\,dy\,dx$ in the order $dx\,dz\,dy$.

Solution. The solid is $0 \le z \le y \le x^2 \le 1$ with $0 \le x \le 1$. In terms of y, the condition $y \le x^2$ means $x \ge \sqrt{y}$. So it is $\{0 \le y \le 1,\ 0 \le z \le y,\ \sqrt{y} \le x \le 1\}$, and the integral is $\int_0^1 \int_0^y \int_{\sqrt{y}}^1 f\,dx\,dz\,dy$. Sketching the projections onto each coordinate plane is the reliable way to get such limits right.

5.6 Cylindrical and Spherical Coordinates

Theorem 5.19 (Volume elements). *In cylindrical coordinates* $dV = \rho\,d\rho\,d\phi\,dz$. *In spherical coordinates (physics convention)*

$$dV = r^2 \sin\theta\,dr\,d\theta\,d\phi.$$

The cylindrical factor is the polar factor from before. The spherical factor is the Jacobian computed in Example 3.31. Geometrically, a small spherical box has edges dr, $r\,d\theta$ (along a meridian), and $r\sin\theta\,d\phi$ (along a circle of latitude, whose radius is $r\sin\theta$).

Example 5.20 (Volume of a ball). $V = \displaystyle\int_0^{2\pi}\int_0^{\pi}\int_0^{a} r^2\sin\theta\,dr\,d\theta\,d\phi = \frac{a^3}{3}\cdot 2\cdot 2\pi = \frac{4}{3}\pi a^3$. The θ-integral $\int_0^\pi \sin\theta\,d\theta = 2$ and the ϕ-integral 2π together give the *solid angle* of a full sphere, 4π steradians.

Example 5.21 (Moment of inertia of a solid ball). A uniform ball of mass M and radius a rotates about a diameter. Find its moment of inertia.

Solution. The distance from the z-axis is $r\sin\theta$, and $\rho = M/(\tfrac{4}{3}\pi a^3)$. So

$$I_z = \rho\int_0^{2\pi}\int_0^{\pi}\int_0^{a} (r\sin\theta)^2 r^2\sin\theta\,dr\,d\theta\,d\phi = \rho\cdot\frac{a^5}{5}\cdot\frac{4}{3}\cdot 2\pi = \frac{2}{5}Ma^2,$$

using $\int_0^\pi \sin^3\theta\,d\theta = \tfrac{4}{3}$.

Example 5.22 (Ice-cream cone). Find the volume of the solid above the cone $z = \sqrt{x^2+y^2}$ and inside the sphere $x^2+y^2+z^2 = z$.

Solution. The sphere is $r = \cos\theta$ and the cone is $\theta = \pi/4$. So

$$V = \int_0^{2\pi}\int_0^{\pi/4}\int_0^{\cos\theta} r^2 \sin\theta \, dr\, d\theta\, d\phi = 2\pi \int_0^{\pi/4} \frac{\cos^3\theta}{3} \sin\theta \, d\theta$$
$$= \frac{2\pi}{3}\cdot\frac{1}{4}\left(1-\frac{1}{4}\right) = \frac{\pi}{8}.$$

Example 5.23 (Radial integrals: a preview of hydrogen). The ground state of hydrogen has $|\psi|^2 = Ce^{-2r/a}$, where a is the Bohr radius. Find C so that $\iiint_{\mathbb{R}^3} |\psi|^2\, dV = 1$, and the expected distance $\langle r\rangle$ of the electron from the nucleus.

Solution. The angular integrals give 4π. With $\int_0^\infty r^n e^{-\lambda r} dr = n!/\lambda^{n+1}$ (Theorem 5.30),

$$1 = 4\pi C\int_0^\infty r^2 e^{-2r/a} dr = 4\pi C\cdot\frac{2!}{(2/a)^3} = \pi a^3 C, \qquad C = \frac{1}{\pi a^3}.$$

Then $\langle r\rangle = \iiint r|\psi|^2 dV = \frac{4}{a^3}\int_0^\infty r^3 e^{-2r/a} dr = \frac{4}{a^3}\cdot\frac{3!}{(2/a)^4} = \frac{3}{2}a$.

5.7 Change of Variables

Polar, cylindrical, and spherical coordinates are special cases of a general principle. A transformation $T:(u,v)\mapsto(x,y)$ distorts small rectangles into small curvilinear parallelograms, and the integral must account for how much areas are stretched. The stretching factor is the absolute value of the Jacobian determinant.

Lemma 5.24 (Linear maps scale area by $|\det|$). *A linear map $L(\mathbf{u}) = A\mathbf{u}$ on $\mathbb{R}^2$ maps the unit square to a parallelogram of area $|\det A|$, and hence maps any region of area S to a region of area $|\det A|\,S$.*

Proof. The image of the unit square is the parallelogram spanned by the columns $\mathbf{a}_1, \mathbf{a}_2$ of A. Regard them as vectors in $\mathbb{R}^3$ with zero third component. Its area is $|\mathbf{a}_1\times\mathbf{a}_2| = |a_{11}a_{22}-a_{21}a_{12}| = |\det A|$. A general region is approximated by squares, each of which is scaled by the same factor. ∎

A differentiable map is approximately linear on small scales, with linear part the Jacobian matrix (Theorem 3.18). Applying the lemma to each small piece suggests the following theorem, whose full proof belongs to analysis.

Theorem 5.25 (Change of variables). *Let $T : S \to D$ be a C^1 bijection between bounded regions in $\mathbb{R}^n$ whose Jacobian determinant $J_T = \det\left[\partial x_i/\partial u_j\right]$ is nonzero on the interior of S. Then for continuous f on D,*

$$\int_D f(\mathbf{x})\,d\mathbf{x} = \int_S f\big(T(\mathbf{u})\big)\,\big|J_T(\mathbf{u})\big|\,d\mathbf{u}.$$

In two variables, $dx\,dy = \left|\dfrac{\partial(x,y)}{\partial(u,v)}\right| du\,dv$.

The hypothesis may fail on a boundary set of zero area, as it does for polar coordinates at $r = 0$, without affecting the conclusion. It is often easier to compute $\partial(u,v)/\partial(x,y)$ and invert: $\dfrac{\partial(x,y)}{\partial(u,v)} = \left[\dfrac{\partial(u,v)}{\partial(x,y)}\right]^{-1}$ by the inverse function theorem.

Example 5.26 (Area of an ellipse). With $x = au$, $y = bv$, the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1$ corresponds to the unit disk, and $J = ab$. So the area is $ab \cdot \pi = \pi ab$. Likewise the ellipsoid of Table 1.1 has volume $\frac{4}{3}\pi abc$.

Example 5.27 (Adapted coordinates). Evaluate $\iint_R \cos\!\left(\dfrac{y-x}{y+x}\right) dA$, where R is the trapezoid with vertices $(1,0)$, $(2,0)$, $(0,2)$, $(0,1)$.

Solution. Let $u = y - x$ and $v = y + x$. The region becomes $1 \le v \le 2$, $-v \le u \le v$, and $\partial(u,v)/\partial(x,y) = -2$, so $dA = \frac{1}{2}\,du\,dv$. Then

$$\int_1^2 \int_{-v}^{v} \cos(u/v)\,\tfrac{1}{2}\,du\,dv = \int_1^2 v \sin 1\,dv = \tfrac{3}{2}\sin 1.$$

Example 5.28 (A region between hyperbolas). Find the area of the region in the first quadrant bounded by $xy = 1$, $xy = 3$, $y = x$, and $y = 4x$.

Solution. Let $u = xy \in [1,3]$ and $v = y/x \in [1,4]$. Then $\dfrac{\partial(u,v)}{\partial(x,y)} =$

$\det\begin{pmatrix} y & x \\ -y/x^2 & 1/x \end{pmatrix} = \dfrac{2y}{x} = 2v$, so $dA = \dfrac{du\,dv}{2v}$ and the area is $\int_1^3 \int_1^4 \dfrac{dv}{2v}\,du = 2 \cdot \tfrac{1}{2}\ln 4 = \ln 4$.

5.8 The Gamma Function and Gaussian Moments

Two families of one-dimensional integrals appear so often in physics that they deserve names and a table of values.

Definition 5.29 (Gamma function). For $s > 0$, $\Gamma(s) = \displaystyle\int_0^\infty t^{s-1}e^{-t}\,dt$.

Theorem 5.30 (Properties of Γ). *(a)* $\Gamma(s+1) = s\,\Gamma(s)$, *and* $\Gamma(n+1) = n!$ *for integers* $n \geq 0$.

(b) $\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$, *and hence* $\Gamma(\tfrac{3}{2}) = \tfrac{\sqrt{\pi}}{2}$, $\Gamma(\tfrac{5}{2}) = \tfrac{3\sqrt{\pi}}{4}$, *and so on.*

(c) $\displaystyle\int_0^\infty r^n e^{-\lambda r}\,dr = \frac{n!}{\lambda^{n+1}}$ *for* $\lambda > 0$.

Proof. (a) Integrate by parts: $\int_0^\infty t^s e^{-t}dt = \big[-t^s e^{-t}\big]_0^\infty + s\int_0^\infty t^{s-1}e^{-t}dt$. Since $\Gamma(1) = 1$, induction gives $n!$. (b) Substitute $t = x^2$: $\Gamma(\tfrac{1}{2}) = \int_0^\infty t^{-1/2}e^{-t}dt = 2\int_0^\infty e^{-x^2}dx = \sqrt{\pi}$. (c) Substitute $t = \lambda r$. ∎

Theorem 5.31 (Gaussian moments). *For* $a > 0$ *and integers* $n \geq 0$,

$$\int_{-\infty}^{\infty} x^{2n}e^{-ax^2}dx = \frac{(2n-1)!!}{(2a)^n}\sqrt{\frac{\pi}{a}}, \qquad \int_{-\infty}^{\infty} x^{2n+1}e^{-ax^2}dx = 0,$$

where $(2n-1)!! = 1\cdot 3\cdot 5\cdots(2n-1)$ *and* $(-1)!! = 1$. *In particular* $\int x^2e^{-ax^2}dx = \tfrac{1}{2a}\sqrt{\pi/a}$ *and* $\int x^4e^{-ax^2}dx = \tfrac{3}{4a^2}\sqrt{\pi/a}$.

Proof. Odd moments vanish by symmetry. For even moments, differentiate $G(a) = \int e^{-ax^2}dx = \sqrt{\pi}\,a^{-1/2}$ with respect to a, n times, under the integral sign (justified by the next theorem). Each derivative brings down a factor of $-x^2$:

$$(-1)^n\int x^{2n}e^{-ax^2}dx = G^{(n)}(a) = \sqrt{\pi}\Big(-\tfrac{1}{2}\Big)\Big(-\tfrac{3}{2}\Big)\cdots\Big(-\tfrac{2n-1}{2}\Big)a^{-n-1/2}.$$ ∎

Theorem 5.32 (Differentiation under the integral sign). *If $f(x,t)$ and $\partial f/\partial t$ are continuous on $[a,b]\times[c,d]$, then for $t\in(c,d)$*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b f(x,t)\,dx=\int_a^b \frac{\partial f}{\partial t}\,dx.$$

The same holds on an infinite interval provided $|\partial f/\partial t(x,t)|\le g(x)$ for an integrable g independent of t.

Proof. For the finite case, write the difference quotient of $F(t)=\int_a^b f\,dx$ as $\int_a^b \frac{f(x,t+h)-f(x,t)}{h}dx$. By the mean value theorem the integrand is $f_t(x,t+\theta h)$ for some $\theta=\theta(x,h)\in(0,1)$. Since f_t is uniformly continuous on the compact rectangle, this converges uniformly in x to $f_t(x,t)$, and the integral of a uniformly convergent family converges. ∎

Example 5.33 (Feynman's trick). Evaluate $I=\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$.

Solution. Introduce $F(t)=\int_0^\infty \frac{\sin x}{x}e^{-tx}dx$ for $t>0$. Differentiating under the integral sign, $F'(t)=-\int_0^\infty \sin x\,e^{-tx}dx=-\frac{1}{1+t^2}$. Since $F(t)\to 0$ as $t\to\infty$, $F(t)=\frac{\pi}{2}-\arctan t$. Letting $t\to 0^+$ (which requires an extra argument at the endpoint) gives $I=\frac{\pi}{2}$. Introducing a parameter and differentiating is a powerful method, which Feynman popularized.

Example 5.34 (Width of a Gaussian). For the probability density $p(x)=\sqrt{a/\pi}\,e^{-ax^2}$, the mean is 0 and the variance is $\langle x^2\rangle=\sqrt{a/\pi}\cdot\frac{1}{2a}\sqrt{\pi/a}=\frac{1}{2a}$. So the standard deviation is $\sigma=1/\sqrt{2a}$. In Chapter 11 we show that the Fourier transform of a Gaussian of width σ is a Gaussian of width $1/\sigma$. That reciprocal relationship is the mathematical core of the Heisenberg uncertainty principle.

Example 5.35 (Volume of the unit ball in $\mathbb{R}^n$). Compute $\int_{\mathbb{R}^n} e^{-|\mathbf{x}|^2}d\mathbf{x}$ in two ways. As a product of n Gaussian integrals it is $\pi^{n/2}$. Using *spherical shells* it is $\int_0^\infty e^{-r^2}S_{n-1}r^{n-1}dr=\frac{1}{2}S_{n-1}\Gamma(n/2)$, where S_{n-1} is the surface area of the unit sphere in $\mathbb{R}^n$. Therefore

$$S_{n-1}=\frac{2\pi^{n/2}}{\Gamma(n/2)},\qquad V_n=\frac{S_{n-1}}{n}=\frac{\pi^{n/2}}{\Gamma(\tfrac{n}{2}+1)}.$$

Check: $V_2 = \pi$ and $V_3 = \pi^{3/2}/\Gamma(\frac{5}{2}) = \frac{4}{3}\pi$. Surprisingly, $V_n \to 0$ as $n \to \infty$. Most of a high-dimensional cube lies outside its inscribed ball.

5.9 Worked Problems

The problems below practice choosing an order of integration and a coordinate system adapted to the region, and end with two integrals that recur in physics: convergence of radial integrals and the Dirichlet integral over a simplex. Spherical coordinates follow the physics convention (r, θ, ϕ) of Chapter 1.

Problem 5.1 Reversing the order to make an integral elementary

Evaluate $\displaystyle\int_0^1 \int_{\sqrt{y}}^1 \sqrt{1+x^3}\, dx\, dy$.

Solution. The inner integral $\int \sqrt{1+x^3}\, dx$ is not elementary (it is an elliptic integral), so we reverse the order. The region is $0 \le y \le 1$, $\sqrt{y} \le x \le 1$, which is the set $0 \le y \le x^2$, $0 \le x \le 1$: the region under the parabola $y = x^2$. Then

$$\int_0^1 \int_0^{x^2} \sqrt{1+x^3}\, dy\, dx = \int_0^1 x^2\sqrt{1+x^3}\, dx = \Big[\frac{2}{9}(1+x^3)^{3/2}\Big]_0^1$$
$$= \frac{2}{9}(2\sqrt{2}-1) \approx 0.4063.$$

The reversal produced exactly the factor x^2 needed for the substitution $w = 1 + x^3$. This is typical. When an iterated integral has a nonelementary inner integral and a variable limit, the limit often supplies the missing derivative after reversal.

Problem 5.2 Volume between a paraboloid and a tilted plane

Find the volume of the solid between the paraboloid $z = x^2 + y^2$ and the plane $z = 2y$.

Solution. The surfaces meet where $x^2 + y^2 = 2y$, that is, on the cylinder over the circle $x^2 + (y-1)^2 = 1$. Inside this circle the plane is above the paraboloid. In polar coordinates the circle is $r^2 = 2r\sin\theta$, or $r = 2\sin\theta$ for $0 \le \theta \le \pi$. Then

$$V = \int_0^{\pi}\int_0^{2\sin\theta} (2r\sin\theta - r^2)\, r\, dr\, d\theta = \int_0^{\pi} \Big(\frac{2\sin\theta}{3}(2\sin\theta)^3 - \frac{(2\sin\theta)^4}{4}\Big) d\theta$$
$$= \int_0^{\pi} \frac{4}{3}\sin^4\theta\, d\theta = \frac{4}{3}\cdot\frac{3\pi}{8} = \frac{\pi}{2}.$$

A cleaner route: in the shifted variables $X = x$, $Y = y - 1$ the height is $2y - x^2 - y^2 = 1 - X^2 - Y^2$, so $V = \iint_{X^2+Y^2\le 1}(1 - X^2 - Y^2)\,dA = 2\pi\int_0^1 (1 - r^2)r\,dr = \frac{\pi}{2}$. Indeed the volume between *any* paraboloid $z = x^2 + y^2$ and any non-vertical plane depends only on the area of the shadow disk: for a disk of radius R it is $\pi R^4/2$.

Problem 5.3 Centroid and moment of inertia of a solid cone

A uniform solid cone of mass M has height h and base radius a. Find its centroid and its moment of inertia about its axis.

Solution. Place the apex at the origin with the axis along z, so the cone is $\rho \le az/h$, $0 \le z \le h$, in cylindrical coordinates (ρ, ϕ, z). Its volume is

$$V = \int_0^h \int_0^{2\pi} \int_0^{az/h} \rho\,d\rho\,d\phi\,dz = 2\pi \int_0^h \frac{a^2 z^2}{2h^2}\,dz = \frac{\pi a^2 h}{3}.$$

By symmetry the centroid is on the axis. Then

$$\iiint z\,dV = 2\pi \int_0^h z\,\frac{a^2 z^2}{2h^2}\,dz = \frac{\pi a^2 h^2}{4}, \qquad \bar{z} = \frac{\pi a^2 h^2/4}{\pi a^2 h/3} = \frac{3}{4}h.$$

So the centroid is three quarters of the way from the apex to the base, or $h/4$ above the base. For the moment of inertia, with density M/V,

$$I_z = \frac{M}{V}\iiint \rho^2\,dV = \frac{M}{V}\cdot 2\pi \int_0^h \frac{a^4 z^4}{4h^4}\,dz = \frac{M}{V}\cdot\frac{\pi a^4 h}{10} = \frac{3}{10}Ma^2.$$

Compare a solid cylinder, $\frac{1}{2}Ma^2$, and a solid sphere, $\frac{2}{5}Ma^2$. All three have the form kMa^2 with k smaller when more of the mass lies near the axis. In the cone only the base reaches the full radius a, and every other slice is a smaller disk, so it has the smallest k of the three.

Problem 5.4 A linear change of variables

Evaluate $\iint_R (x + y)^2 e^{x-y}\,dA$, where R is the square bounded by $x + y = 1$, $x + y = 2$, $x - y = 0$, $x - y = 1$.

Solution. The boundaries suggest $u = x + y$ and $v = x - y$, which map R to the rectangle $1 \le u \le 2$, $0 \le v \le 1$. Then $\dfrac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$, so by Theorem 5.25, $dA = \left|\dfrac{\partial(x,y)}{\partial(u,v)}\right| du\,dv = \frac{1}{2}\,du\,dv$. The integral separates:

$$\iint_R (x + y)^2 e^{x-y}\,dA = \frac{1}{2}\int_1^2 u^2\,du \int_0^1 e^v\,dv = \frac{1}{2}\cdot\frac{7}{3}\cdot(e - 1) = \frac{7(e-1)}{6} \approx 2.005.$$

The Jacobian of the *inverse* map is the reciprocal of the forward Jacobian. Here $|-2|$ inverts to $\frac{1}{2}$. Forgetting to invert is the most common error with this theorem, and the area check catches it: R is a square with side $\frac{1}{\sqrt{2}}$, of area $\frac{1}{2}$, which matches $\frac{1}{2}\cdot$ (area of the unit square in uv).

Problem 5.5 Volume and moment of inertia of an ellipsoid

Let E be the solid ellipsoid $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}\le 1$ with uniform density and mass M. Find its volume and its moment of inertia about the z-axis.

Solution. Stretch the unit ball B: $x=au$, $y=bv$, $z=cw$ maps B onto E with Jacobian abc. So $\operatorname{vol}(E)=abc\cdot\frac{4}{3}\pi$. For the moment of inertia, with density $M/\operatorname{vol}(E)$,

$$I_z=\frac{M}{\frac{4}{3}\pi abc}\iiint_E (x^2+y^2)\,dV=\frac{M}{\frac{4}{3}\pi}\iiint_B (a^2u^2+b^2v^2)\,dV.$$

By symmetry of the ball, $\iiint_B u^2\,dV=\iiint_B v^2\,dV=\frac{1}{3}\iiint_B r^2\,dV=\frac{1}{3}\cdot 4\pi\int_0^1 r^4\,dr=\frac{4\pi}{15}$. Therefore

$$I_z=\frac{M}{\frac{4}{3}\pi}\cdot\frac{4\pi}{15}(a^2+b^2)=\frac{M(a^2+b^2)}{5}.$$

For $a=b=c$ this is $\frac{2}{5}Ma^2$, the result of Example 5.21. The symmetry argument, replacing u^2 by one third of r^2, avoided a trigonometric integral entirely.

Problem 5.6 Convergence of radial integrals

For which $p>0$ does $\displaystyle\iint_{\mathbb{R}^2}\frac{dA}{(1+x^2+y^2)^p}$ converge? What is its value? Generalize to $\mathbb{R}^n$.

Solution. The integrand is positive, so the improper integral is the limit over disks $r\le R$. In polar coordinates,

$$\iint_{r\le R}\frac{dA}{(1+r^2)^p}=2\pi\int_0^R\frac{r\,dr}{(1+r^2)^p}=\frac{\pi}{1-p}\Big[(1+R^2)^{1-p}-1\Big]\quad(p\neq 1).$$

As $R\to\infty$ this converges if and only if $p>1$, to $\dfrac{\pi}{p-1}$. For $p=1$ the integral is $\pi\ln(1+R^2)\to\infty$.

In $\mathbb{R}^n$ the volume element is $S_{n-1}r^{n-1}dr$ (Example 5.35), and the integrand behaves like r^{-2p} for large r. So the integral converges if and only if $\int^\infty r^{n-1-2p}dr<$

∞, that is, $2p > n$. The critical exponent grows with dimension, because there is more room at large r. This counting governs the normalizability of wavefunctions: in three dimensions, a bound state that decays like r^{-k} can be normalized only if $|\psi|^2 \sim r^{-2k}$ is integrable, which needs $2k > 3$.

Problem 5.7 An ice-cream cone

Find the volume of the solid inside the sphere $r = 2$ and inside the cone $\theta = \pi/6$.

Solution. In spherical coordinates the solid is the box $0 \le r \le 2$, $0 \le \theta \le \pi/6$, $0 \le \phi < 2\pi$:

$$V = \int_0^{2\pi} d\phi \int_0^{\pi/6} \sin\theta \, d\theta \int_0^{2} r^2 \, dr = 2\pi\Big(1 - \frac{\sqrt{3}}{2}\Big)\frac{8}{3} = \frac{16\pi}{3}\Big(1 - \frac{\sqrt{3}}{2}\Big) \approx 2.245.$$

As a check on the geometry, the formula $V = \frac{2\pi}{3}R^3(1-\cos\alpha)$ for a cone of half-angle α capped by a sphere of radius R gives the full ball when $\alpha = \pi$ and the hemisphere when $\alpha = \pi/2$. The factor $2\pi(1-\cos\alpha)$ is the solid angle of the cone, and $V = \frac{1}{3}R^3 \times (\text{solid angle})$, just as the area of a circular sector is $\frac{1}{2}R^2 \times (\text{angle})$.

Problem 5.8 The Dirichlet integral over a simplex

Let T be the tetrahedron $x, y, z \ge 0$, $x + y + z \le 1$. Evaluate $\iiint_T xyz \, dV$, and verify the general formula

$$\iiint_T x^a y^b z^c \, dV = \frac{a!\, b!\, c!}{(a+b+c+3)!}$$

in this case and for $a = b = c = 0$ and $a = b = 0, c = 1$.

Solution. Integrate in z, then y, then x. The innermost integral is $\int_0^{1-x-y} z \, dz = \frac{1}{2}(1-x-y)^2$. Then with $s = 1 - x$,

$$\int_0^{s} y\,\frac{(s-y)^2}{2}\,dy = \frac{1}{2}\int_0^{s}(s^2 y - 2sy^2 + y^3)\,dy = \frac{1}{2}\Big(\frac{s^4}{2} - \frac{2s^4}{3} + \frac{s^4}{4}\Big) = \frac{s^4}{24}.$$

Finally $\displaystyle\int_0^1 x\frac{(1-x)^4}{24}\,dx = \frac{1}{24}B(2,5) = \frac{1}{24}\cdot\frac{1!\,4!}{6!} = \frac{1}{720}$, using the Beta integral $\int_0^1 x^m(1-x)^n dx = \frac{m!\,n!}{(m+n+1)!}$.

The general formula gives $\frac{1!\,1!\,1!}{6!} = \frac{1}{720}$, in agreement. With $a = b = c = 0$ it gives $\frac{1}{3!} = \frac{1}{6}$, the volume of T. With $c = 1$ it gives $\frac{1}{4!} = \frac{1}{24}$, matching Example 5.17. The general case follows by the same steps: each integration is a Beta integral, and the factorials telescope. This formula is the basis of Feynman-parameter integrals in quantum field theory and of the Dirichlet distribution in statistics.

EXERCISES

A. COMPUTATION

5.1 Evaluate $\iint_R xy^2\,dA$ for $R = [0,2] \times [1,3]$.

5.2 Evaluate $\iint_R \dfrac{x}{1+xy}\,dA$ for $R = [0,1] \times [0,1]$.

5.3 Evaluate $\iint_R ye^{-xy}\,dA$ for $R = [0,2] \times [0,3]$.

5.4 Evaluate $\iint_D x\,dA$ where D is bounded by $y = x^2$ and $y = 2x$.

5.5 Evaluate $\iint_D (x^2 + 2y)\,dA$ where D is bounded by $y = x$, $y = x^3$, $x \geq 0$.

5.6 Find the volume of the solid bounded by $z = x^2 + y^2$, $y = x^2$, $y = 1$, and $z = 0$.

5.7 Reverse the order and evaluate $\int_0^1 \int_{\sqrt{y}}^1 \sqrt{x^3+1}\,dx\,dy$.

5.8 Reverse the order and evaluate $\int_0^4 \int_{\sqrt{x}}^2 \frac{1}{y^3+1}\,dy\,dx$.

5.9 Reverse the order and evaluate $\int_0^1 \int_y^1 e^{x^2}\,dx\,dy$.

5.10 Use polar coordinates to evaluate $\iint_D e^{-x^2-y^2}\,dA$, where D is the half-disk $x^2 + y^2 \leq 4$, $x \geq 0$.

5.11 Find the area enclosed by one loop of the rose $r = \cos 3\theta$.

5.12 Find the volume inside the sphere $x^2 + y^2 + z^2 = 16$ and outside the cylinder $x^2 + y^2 = 4$.

5.13 Evaluate $\int_0^1 \int_0^{\sqrt{1-x^2}} (x^2+y^2)^{3/2}\,dy\,dx$ by converting to polar coordinates.

5.14 Find the mass and center of mass of the triangular lamina with vertices $(0,0)$, $(1,0)$, $(0,2)$ and density $\rho = 1 + 3x + y$.

5.15 Find the moment of inertia I_0 of a uniform disk of mass M and radius a about its center.

5.16 The joint density of (X, Y) is $f(x,y) = Cx(1+y)$ on $[0,1] \times [0,2]$. Find C, $P(X \leq \frac{1}{2})$, and $E[X]$.

5.17 Evaluate $\iiint_E 2x\,dV$, where $E = \{0 \leq y \leq 2,\ 0 \leq x \leq \sqrt{4-y^2},\ 0 \leq z \leq y\}$.

5.18 Find the volume of the tetrahedron bounded by $x + 2y + z = 2$, $x = 2y$, $x = 0$, $z = 0$.

5.19 Find the centroid of the solid hemisphere $x^2+y^2+z^2 \le a^2$, $z \ge 0$.

5.20 Evaluate $\iiint_E \sqrt{x^2+y^2}\,dV$, where E is inside $x^2+y^2=16$ and between $z=-5$ and $z=4$.

5.21 Find the volume of the region bounded by the paraboloids $z = x^2+y^2$ and $z = 36-3x^2-3y^2$.

5.22 Evaluate $\iiint_B (x^2+y^2+z^2)^2\,dV$ for the unit ball B.

5.23 Evaluate $\iiint_E z\,dV$ where E lies between the spheres $r=1$ and $r=2$ in the first octant.

5.24 Find the mass of a ball of radius a whose density is proportional to the distance from the center, $\rho = kr$.

5.25 Find the moment of inertia of a uniform thin spherical shell of mass M and radius a about a diameter.

5.26 Use $x=2u+v$, $y=u+2v$ to evaluate $\iint_R (x-3y)\,dA$, where R is the triangle with vertices $(0,0)$, $(2,1)$, $(1,2)$.

5.27 Evaluate $\iint_R (x+y)e^{x^2-y^2}\,dA$, where R is the rectangle enclosed by $x-y=0$, $x-y=2$, $x+y=0$, $x+y=3$.

5.28 Evaluate $\iint_R e^{(x+y)/(x-y)}\,dA$ over the trapezoid with vertices $(1,0)$, $(2,0)$, $(0,-2)$, $(0,-1)$.

5.29 Compute $\Gamma(\tfrac{7}{2})$ and $\int_0^\infty x^5 e^{-x^2}\,dx$.

5.30 Compute $\int_{-\infty}^{\infty} x^2 e^{-3x^2}\,dx$ and $\int_{-\infty}^{\infty} x^6 e^{-x^2}\,dx$.

5.31 Compute $\int_0^\infty r^4 e^{-r/a}\,dr$ and $\int_{-\infty}^{\infty} e^{-x^2+2x}\,dx$ (complete the square).

5.32 For the hydrogen ground state of Example 5.23, compute $\langle r^2\rangle$ and $\langle 1/r\rangle$.

5.33 Find the probability that the hydrogen ground-state electron is found farther than a from the nucleus.

B. THEORY AND PROOF

5.34 Verify the two iterated integrals in the Fubini pitfall. (Hint: $\frac{\partial}{\partial y}\frac{y}{x^2+y^2} = \frac{x^2-y^2}{(x^2+y^2)^2}$.)

5.35 Prove that if f is continuous on $[a,b]$, then $\int_a^b \int_a^x f(t)\,dt\,dx = \int_a^b (b-t)f(t)\,dt$. Generalize to n-fold iterated integrals (Cauchy's formula for repeated integration).

5.36 Prove Pappus's theorem: the volume of a solid of revolution obtained by rotating a plane region D about an external axis equals the area of D times the distance travelled by its centroid. Use it to find the volume of a torus.

5.37 Prove the parallel-axis theorem $I = I_{\text{cm}} + Md^2$.

5.38 Derive the cylindrical and spherical volume elements from Theorem 5.25.

5.39 Show that $\Gamma(s)\Gamma(t) = B(s,t)\,\Gamma(s+t)$, where $B(s,t) = \int_0^1 u^{s-1}(1-u)^{t-1}du$, by writing $\Gamma(s)\Gamma(t)$ as a double integral and changing variables.

5.40 Use differentiation under the integral sign to evaluate $\displaystyle\int_0^\infty \frac{e^{-ax} - e^{-bx}}{x}\,dx$ for $a, b > 0$.

5.41 Show that $\int_{-\infty}^{\infty} e^{-ax^2+bx}\,dx = \sqrt{\pi/a}\; e^{b^2/(4a)}$ for $a > 0$ and real b.

C. CHALLENGE

5.42 Compute $\int_{-\infty}^{\infty} e^{-x^2}\cos(kx)\,dx$ by differentiating with respect to k and solving the resulting differential equation. This is the Fourier transform of a Gaussian, computed without complex analysis.

5.43 Find the fraction of the volume of the unit cube $[-1,1]^n$ occupied by the unit ball, for $n = 2, 3, 10$. What happens as $n \to \infty$?

5.44 (Stirling's approximation.) Write $n! = \int_0^\infty e^{n\ln t - t}dt$, substitute $t = n + s\sqrt{n}$, and expand the exponent to second order in s to show that $n! \sim \sqrt{2\pi n}\,(n/e)^n$.

5.45 The density $|\psi|^2$ of a $2p_z$ electron in hydrogen is proportional to $r^2 e^{-r/a}\cos^2\theta$. Normalize it and find the probability that the electron is in the upper half-space $z > 0$, and the expected value of r.

CHAPTER 6

Vector Fields and Line Integrals

LEARNING OBJECTIVES

- Describe vector fields and their flow lines, including gravitational and electric fields.
- Define and evaluate line integrals of scalar functions and of vector fields, and interpret them as mass, work, circulation, and flux.
- Prove the fundamental theorem for line integrals and characterize conservative vector fields.
- Find potential functions, and recognize when a curl-free field fails to be conservative.
- Define divergence and curl, interpret them physically, and use the standard vector identities.
- State and prove Green's theorem, and apply it to circulation, flux, and area.

A vector field assigns a vector to every point of a region: the velocity of a fluid, the force of gravity, the electric field of a charge. Integrating a vector field along a curve measures the work it does on a particle moving along the curve, or how much fluid circulates around the curve. This chapter develops these *line integrals* and proves the first of the great integral theorems of vector calculus, Green's theorem. Its higher-dimensional versions, the theorems of Stokes and Gauss, complete Part I in Chapter 7.

6.1 VECTOR FIELDS

Definition 6.1 (Vector field). A *vector field* on a region $D \subseteq \mathbb{R}^n$ is a function $\mathbf{F} : D \to \mathbb{R}^n$. In three dimensions, $\mathbf{F}(x,y,z) = \langle P(x,y,z),\, Q(x,y,z),\, R(x,y,z)\rangle$, and $\mathbf{F}$ is *continuous* or C^1 when its component functions are.

We picture a vector field by drawing the arrow $\mathbf{F}(\mathbf{x})$ with its tail at $\mathbf{x}$, at a representative sample of points.

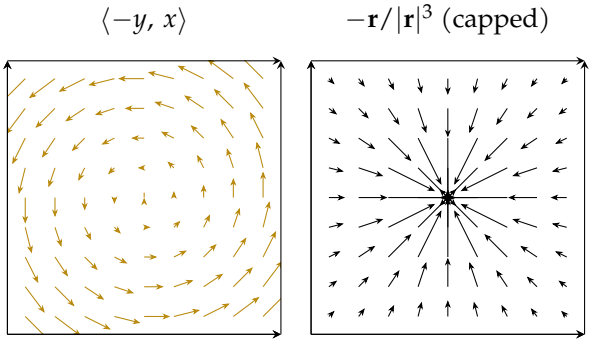


Figure 6.1: A rotational field (left) and an inverse-square attractive field (right). On the right, arrows near the origin are shortened for legibility.

Example 6.2 (Fields from physics). (a) *Rotation.* A rigid body spinning with angular velocity $\boldsymbol{\omega}$ about an axis through the origin has velocity field $\mathbf{v}(\mathbf{r}) = \boldsymbol{\omega} \times \mathbf{r}$. For $\boldsymbol{\omega} = \hat{\mathbf{k}}$ this is $\langle -y, x, 0\rangle$.

(b) *Gravity.* A mass M at the origin exerts on a mass m at $\mathbf{r}$ the force $\mathbf{F} = -\dfrac{GMm}{|\mathbf{r}|^3}\,\mathbf{r} = -\dfrac{GMm}{r^2}\,\hat{\mathbf{r}}$.

(c) *Electrostatics.* A point charge q at the origin produces the field $\mathbf{E} = \dfrac{q}{4\pi\varepsilon_0}\dfrac{\hat{\mathbf{r}}}{r^2}$, also an inverse-square field.

Definition 6.3 (Gradient field, potential). A vector field $\mathbf{F}$ is a *gradient field*, or *conservative*, on D if $\mathbf{F} = \nabla f$ for some scalar function f, called a *potential function* for $\mathbf{F}$.

In physics the convention is $\mathbf{F} = -\nabla U$, where U is the *potential energy*; the sign makes forces push toward lower energy. The gravitational field of Exam-

ple 6.2(b) is conservative with $U = -GMm/r$, since $\nabla(1/r) = -\mathbf{r}/r^3$ (Example 6.4). By Theorem 3.25, a gradient field is everywhere perpendicular to the level sets of its potential, the *equipotential surfaces*.

Example 6.4 (Gradient of a power of r). With $r = |\mathbf{r}| = \sqrt{x^2+y^2+z^2}$, $\partial r/\partial x = x/r$, so $\nabla r = \mathbf{r}/r = \hat{\mathbf{r}}$. By the chain rule, $\nabla(r^n) = nr^{n-1}\hat{\mathbf{r}} = nr^{n-2}\mathbf{r}$. In particular $\nabla(1/r) = -\mathbf{r}/r^3$.

Definition 6.5 (Flow lines). A *flow line* (or *field line*, or *integral curve*) of $\mathbf{F}$ is a curve $\mathbf{r}(t)$ with $\mathbf{r}'(t) = \mathbf{F}(\mathbf{r}(t))$ for all t.

Flow lines are the paths of particles carried by a fluid with velocity field $\mathbf{F}$. Finding them means solving a system of ordinary differential equations. For $\mathbf{F} = \langle -y, x\rangle$ the system $x' = -y$, $y' = x$ has solutions $\mathbf{r}(t) = \langle a\cos t - b\sin t,\ a\sin t + b\cos t\rangle$: concentric circles traversed counterclockwise. Systems of this kind are the subject of Chapter 10.

6.2 Line Integrals of Scalar Functions

Suppose a thin wire lies along a smooth curve C and has linear density $f(x,y,z)$, in mass per unit length. Chop C into small arcs of lengths Δs_i, choose a point P_i^* in each, and add up $f(P_i^*)\Delta s_i$. The limit of these sums is the mass.

Definition 6.6 (Line integral with respect to arc length). Let C be a smooth curve parametrized by $\mathbf{r}(t)$, $a \le t \le b$, traversed once, and let f be continuous on C. Then

$$\int_C f\,ds = \int_a^b f(\mathbf{r}(t))\,|\mathbf{r}'(t)|\,dt.$$

The definition is independent of the parametrization, by the same substitution argument that proved invariance of arc length. It is also independent of the direction of traversal, since $ds = |\mathbf{r}'(t)|\,dt$ is always positive. If C is *piecewise smooth*, a union of smooth curves $C_1, \dots, C_k$ joined end to end, we set $\int_C f\,ds = \sum_i \int_{C_i} f\,ds$. With $f = 1$, $\int_C ds$ is the length of C.

Example 6.7 (Mass of a semicircular wire). A wire bent into the semicircle $x^2 + y^2 = 1$, $y \ge 0$, is denser near its base, with density $\rho = k(2 - y)$. Find

its mass and center of mass.

Solution. With $\mathbf{r}(t) = \langle \cos t, \sin t \rangle$, $0 \le t \le \pi$, we have $ds = dt$, and $m = \int_0^\pi k(2 - \sin t)\,dt = k(2\pi - 2)$. By symmetry $\bar{x} = 0$. Also $\int_C y\rho\,ds = k\int_0^\pi (2\sin t - \sin^2 t)\,dt = k\left(4 - \frac{\pi}{2}\right)$, so $\bar{y} = \dfrac{4 - \pi/2}{2\pi - 2} = \dfrac{8 - \pi}{4(\pi - 1)} \approx 0.567$. A uniform wire would have $\bar{y} = 2/\pi \approx 0.637$.

Example 6.8 (A helical wire). Evaluate $\int_C y \sin z\,ds$ where C is the helix $\mathbf{r}(t) = \langle \cos t, \sin t, t \rangle$, $0 \le t \le 2\pi$.

Solution. $|\mathbf{r}'(t)| = \sqrt{2}$, so $\int_0^{2\pi} \sin t \sin t\,\sqrt{2}\,dt = \sqrt{2}\,\pi$.

6.3 Line Integrals of Vector Fields

A particle moves along a curve C through a force field $\mathbf{F}$. Over a short arc of length Δs, its displacement is approximately $\hat{\mathbf{T}}\Delta s$, and the work done is $\mathbf{F}\cdot\hat{\mathbf{T}}\,\Delta s$. Summing gives the total work.

Definition 6.9 (Line integral of a vector field). Let C be a smooth *oriented* curve given by $\mathbf{r}(t)$, $a \le t \le b$, and let $\mathbf{F}$ be continuous on C. Then

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_C \mathbf{F}\cdot\hat{\mathbf{T}}\,ds = \int_a^b \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt.$$

In components, with $\mathbf{F} = \langle P, Q, R\rangle$ and $d\mathbf{r} = \langle dx, dy, dz\rangle$, this is $\int_C P\,dx + Q\,dy + R\,dz$. When C is closed, we write $\oint_C$ and call the integral the *circulation* of $\mathbf{F}$ around C.

Unlike $\int_C f\,ds$, this integral depends on the orientation of C: reversing it changes the sign of $\hat{\mathbf{T}}$. If $-C$ denotes the curve with the opposite orientation, then $\int_{-C}\mathbf{F}\cdot d\mathbf{r} = -\int_C \mathbf{F}\cdot d\mathbf{r}$. It is, however, independent of orientation-preserving reparametrizations.

Example 6.10 (Work along two paths). Let $\mathbf{F} = \langle y^2,\, x\rangle$. Compute $\int_C \mathbf{F}\cdot d\mathbf{r}$ from $(-5,-3)$ to $(0,2)$ along (a) the straight segment and (b) the parabola $x = 4 - y^2$.

Solution. (a) $\mathbf{r}(t) = \langle -5 + 5t,\, -3 + 5t\rangle$, $0 \le t \le 1$: $\int_0^1 \big[(5t-3)^2\cdot 5 + (5t - 5)\cdot 5\big]\,dt = -\frac{5}{6}$. (b) With y as parameter, $x = 4 - y^2$ and $dx = -2y\,dy$ for

$-3 \le y \le 2$: $\int_{-3}^{2} \left[y^2(-2y) + (4 - y^2)\right] dy = \frac{245}{6} \approx 40.83$. The two answers differ. The work depends on the path, so this field is not conservative.

Example 6.11 (Work in a rotational field). Compute the circulation of $\mathbf{F} = \langle -y, x \rangle$ counterclockwise around the circle of radius a.

Solution. $\mathbf{r} = \langle a\cos t, a\sin t \rangle$: $\oint \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} (a^2 \sin^2 t + a^2 \cos^2 t)\, dt = 2\pi a^2$. The field points along the direction of motion everywhere, so the circulation is positive and grows with the circle. It equals twice the enclosed area, a fact Green's theorem will explain.

6.3.1 FLUX ACROSS A PLANE CURVE

For a plane curve, a second integral is natural: how much of the field crosses the curve rather than flowing along it. If C is traversed with the region on the left, the outward unit normal is $\hat{\mathbf{n}} = \hat{\mathbf{T}} \times \hat{\mathbf{k}}$, which is $\hat{\mathbf{T}}$ rotated clockwise by $90°$. With $\mathbf{r}' = \langle x', y' \rangle$ we have $\hat{\mathbf{n}}\, ds = \langle y', -x' \rangle\, dt$.

Definition 6.12 (Flux across a plane curve). The *flux* of $\mathbf{F} = \langle P, Q \rangle$ across C is $\displaystyle\int_C \mathbf{F} \cdot \hat{\mathbf{n}}\, ds = \int_C P\, dy - Q\, dx$.

Example 6.13 (Flux of a radial field). The flux of $\mathbf{F} = \langle x, y \rangle$ outward across the circle of radius a is $\int_0^{2\pi} (a\cos t \cdot a\cos t + a\sin t \cdot a\sin t)\, dt = 2\pi a^2$. The flux of $\mathbf{G} = \mathbf{r}/|\mathbf{r}|^2$ is $\int_0^{2\pi} 1\, dt = 2\pi$ for *every* radius. In two dimensions the inverse-first-power field is the analog of the inverse-square field: its flux through any circle about the source is the same. This is Gauss's law in the plane.

6.4 THE FUNDAMENTAL THEOREM FOR LINE INTEGRALS

Theorem 6.14 (Fundamental theorem for line integrals). *Let C be a piecewise*

smooth curve from A to B, and let f be C^1 on an open set containing C. Then

$$\int_C \nabla f \cdot d\mathbf{r} = f(B) - f(A).$$

Proof. For a smooth piece $\mathbf{r}(t)$, $a \le t \le b$, the chain rule gives $\nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = \mathrm{d}/\mathrm{d}t\, f(\mathbf{r}(t))$. The ordinary fundamental theorem of calculus then gives $f(\mathbf{r}(b)) - f(\mathbf{r}(a))$. For a piecewise smooth curve, the intermediate values telescope. ∎

The integral of a gradient depends only on the endpoints, just as $\int_a^b F'(x)\,dx$ depends only on $F(a)$ and $F(b)$.

Definition 6.15 (Path independence). $\int_C \mathbf{F} \cdot d\mathbf{r}$ is *path independent* in D if, for any two points $A, B \in D$, it has the same value for every piecewise smooth path in D from A to B. A region D is *open* if every point of D is the center of some ball contained in D, and *connected* if any two of its points are joined by a path in D.

Theorem 6.16 (Characterization of conservative fields). *Let $\mathbf{F}$ be continuous on an open connected region D. The following are equivalent:*

(*i*) $\mathbf{F} = \nabla f$ *for some C^1 function f on D;*

(*ii*) $\int_C \mathbf{F} \cdot d\mathbf{r}$ *is path independent in D;*

(*iii*) $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ *for every closed piecewise smooth curve C in D.*

Proof. (i)⇒(ii) is Theorem 6.14. (ii)⇔(iii): a closed curve through A and B splits into two paths C_1 and $-C_2$ from A to B, so $\oint = \int_{C_1} - \int_{C_2}$. That is zero exactly when the two path integrals agree.

(ii)⇒(i): Fix $A \in D$ and define $f(\mathbf{x}) = \int_A^{\mathbf{x}} \mathbf{F} \cdot d\mathbf{r}$ along any path, which is well defined by (ii). To compute $\partial f/\partial x$ at $\mathbf{x}$, choose a path that arrives at $\mathbf{x}$ along a short horizontal segment from $\mathbf{x} - h\hat{\mathbf{i}}$. This is possible because D is open. Along that segment only the P-component contributes:

$$f(\mathbf{x}) = \text{const} + \int_{-h}^{0} P(\mathbf{x} + t\hat{\mathbf{i}})\,dt, \qquad \text{so} \qquad \frac{\partial f}{\partial x} = P(\mathbf{x})$$

by the fundamental theorem of calculus. The other partials are found similarly, so $\nabla f = \mathbf{F}$. ∎

Theorem 6.17 (Conservation of energy). *If a particle of mass m moves under a conservative force $\mathbf{F} = -\nabla U$, its total energy $E = \frac{1}{2}m|\mathbf{v}|^2 + U(\mathbf{r})$ is constant.*

Proof. By Newton's law, the work done from time t_0 to t_1 is $\int \mathbf{F}\cdot d\mathbf{r} = \int_{t_0}^{t_1} m\mathbf{r}''\cdot\mathbf{r}'\,dt = \left[\tfrac{1}{2}m|\mathbf{r}'|^2\right]_{t_0}^{t_1}$. By Theorem 6.14 it also equals $-U(\mathbf{r}(t_1)) + U(\mathbf{r}(t_0))$. Equate the two. ∎

This is the origin of the word *conservative*. In quantum mechanics, energy conservation reappears as the statement that the Hamiltonian commutes with itself and generates time evolution (Chapter 15).

6.4.1 FINDING A POTENTIAL

In two dimensions, if $\mathbf{F} = \langle P, Q\rangle = \nabla f$ with f of class C^2, then $P_y = f_{xy} = f_{yx} = Q_x$ by Clairaut's theorem. So $P_y = Q_x$ is *necessary* for $\mathbf{F}$ to be conservative. In three dimensions the corresponding conditions are $P_y = Q_x$, $P_z = R_x$, and $Q_z = R_y$, which together say $\nabla \times \mathbf{F} = \mathbf{0}$ (Section 6.5).

Example 6.18 (Constructing a potential). Show that $\mathbf{F} = \langle y^2,\ 2xy + e^{3z},\ 3ye^{3z}\rangle$ is conservative, find a potential, and evaluate $\int_C \mathbf{F}\cdot d\mathbf{r}$ along any path from $(0,0,0)$ to $(1,2,0)$.

Solution. Check: $P_y = 2y = Q_x$, $P_z = 0 = R_x$, and $Q_z = 3e^{3z} = R_y$. Integrate the first component in x: $f = xy^2 + g(y,z)$. Then $f_y = 2xy + g_y = 2xy + e^{3z}$ forces $g_y = e^{3z}$, so $g = ye^{3z} + h(z)$. Finally $f_z = 3ye^{3z} + h'(z) = R$ forces $h' = 0$. So $f = xy^2 + ye^{3z}$, and the integral is $f(1,2,0) - f(0,0,0) = 4 + 2 = 6$.

COMMON ERROR (curl-free is not always conservative). The condition $P_y = Q_x$ is necessary but not sufficient. Consider

$$\mathbf{F} = \frac{\langle -y,\ x\rangle}{x^2 + y^2} \qquad \text{on } D = \mathbb{R}^2 \setminus \{(0,0)\}.$$

A direct computation gives $P_y = Q_x = \dfrac{y^2 - x^2}{(x^2+y^2)^2}$. Yet around the unit circle, $\oint \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi}(\sin^2 t + \cos^2 t)\,dt = 2\pi \neq 0$. Locally $\mathbf{F} = \nabla\theta$, the gradient of the polar angle, but θ cannot be defined continuously on any region that encircles the origin. The hole in D is the obstruction.

Definition 6.19 (Simply connected). A connected region D is *simply connected* if every closed curve in D can be continuously shrunk to a point without leaving D. In the plane: D has no holes.

Theorem 6.20 (Test for conservative fields). *If* $\mathbf{F}$ *is* C^1 *on an open, simply connected region* D *and* $\nabla \times \mathbf{F} = \mathbf{0}$ *(in the plane,* $P_y = Q_x$*) on* D*, then* $\mathbf{F}$ *is conservative on* D*.*

In the plane, this follows from Green's theorem (Corollary 6.32); in space, from Stokes' theorem.

LOOKING AHEAD. The field of the pitfall is not merely a curiosity. An infinitely long solenoid confines a magnetic field to its interior, yet outside it the vector potential $\mathbf{A}$ has exactly this form: curl-free but with nonzero circulation. Classically, a charged particle outside the solenoid feels no force. Quantum mechanically, its wavefunction picks up a phase $e^{iq\oint \mathbf{A}\cdot d\mathbf{r}/\hbar}$ around the solenoid, and that phase shifts an interference pattern. This *Aharonov–Bohm effect* has been observed, and it shows that in quantum mechanics potentials are more fundamental than forces.

6.5 Divergence and Curl

Treat the symbol $\nabla = \langle \partial_x, \partial_y, \partial_z \rangle$ as a vector of operators. Then besides the gradient ∇f there are two natural ways for it to act on a vector field: a dot product and a cross product.

Definition 6.21 (Divergence and curl). For a C^1 field $\mathbf{F} = \langle P, Q, R \rangle$,

$$\nabla \cdot \mathbf{F} = \nabla \cdot \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z},$$

$$\nabla \times \mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \partial_x & \partial_y & \partial_z \\ P & Q & R \end{vmatrix} = \langle R_y - Q_z,\ P_z - R_x,\ Q_x - P_y \rangle .$$

In index notation, $\nabla \cdot \mathbf{F} = \partial_i F_i$ and $(\nabla \times \mathbf{F})_i = \varepsilon_{ijk} \partial_j F_k$.

Divergence measures the net outflow per unit volume at a point. It is positive at a source and negative at a sink. *Curl* measures local rotation: if $\mathbf{F}$ is a fluid

velocity field, a tiny paddle wheel at $\mathbf{x}$ spins fastest when its axle points along $\boldsymbol{\nabla}\times\mathbf{F}(\mathbf{x})$, with angular velocity $\tfrac{1}{2}|\boldsymbol{\nabla}\times\mathbf{F}|$. Both interpretations become theorems in Chapter 7, where divergence is shown to be a flux density and curl a circulation density.

Example 6.22 (Rigid rotation and a radial field). (a) For $\mathbf{v}=\boldsymbol{\omega}\times\mathbf{r}$ with constant $\boldsymbol{\omega}$: $\boldsymbol{\nabla}\cdot\mathbf{v}=0$ and $\boldsymbol{\nabla}\times\mathbf{v}=2\boldsymbol{\omega}$. A rotating body neither compresses nor expands, and its curl is twice its angular velocity. (b) For $\mathbf{F}=\mathbf{r}=\langle x,y,z\rangle$: $\boldsymbol{\nabla}\cdot\mathbf{F}=3$ and $\boldsymbol{\nabla}\times\mathbf{F}=\mathbf{0}$, a pure source with no rotation. (c) For $\mathbf{F}=\mathbf{r}/r^3$: by Example 6.4 and the product rule, $\boldsymbol{\nabla}\cdot\mathbf{F}=\dfrac{3}{r^3}+\mathbf{r}\cdot\nabla(r^{-3})=\dfrac{3}{r^3}-\dfrac{3}{r^3}=0$ for $r\neq 0$. The inverse-square field is divergence-free away from the origin. All of its divergence is concentrated at the point charge, a fact we make precise with the Dirac delta function in Chapter 7.

Theorem 6.23 (Second-derivative identities). *For C^2 functions f and fields $\mathbf{F}$:*

$$\boldsymbol{\nabla}\times(\nabla f)=\mathbf{0},\qquad \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{F})=0,\qquad \boldsymbol{\nabla}\cdot(\nabla f)=\nabla^2 f,$$

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{F})=\nabla(\boldsymbol{\nabla}\cdot\mathbf{F})-\nabla^2\mathbf{F},$$

where $\nabla^2\mathbf{F}$ is the Laplacian applied to each component.

Proof. $(\nabla\times\nabla f)_i=\varepsilon_{ijk}\partial_j\partial_k f$. Here ε_{ijk} is antisymmetric in j,k while $\partial_j\partial_k$ is symmetric by Clairaut's theorem, so the sum vanishes. The second identity is the same argument applied to $\partial_i\varepsilon_{ijk}\partial_j F_k$. The last follows from the ε–δ identity exactly as in the proof of BAC–CAB (Theorem 1.26): $\varepsilon_{ijk}\partial_j\varepsilon_{kmn}\partial_m F_n=(\delta_{im}\delta_{jn}-\delta_{in}\delta_{jm})\partial_j\partial_m F_n=\partial_i\partial_n F_n-\partial_j\partial_j F_i$. ∎

Theorem 6.24 (Product rules). *For C^1 scalar functions f,g and fields $\mathbf{F},\mathbf{G}$:*

$$\nabla(fg)=f\nabla g+g\nabla f,\qquad \boldsymbol{\nabla}\cdot(f\mathbf{F})=f\,\boldsymbol{\nabla}\cdot\mathbf{F}+\mathbf{F}\cdot\nabla f,$$
$$\boldsymbol{\nabla}\times(f\mathbf{F})=f\,\boldsymbol{\nabla}\times\mathbf{F}+\nabla f\times\mathbf{F},\qquad \boldsymbol{\nabla}\cdot(\mathbf{F}\times\mathbf{G})=\mathbf{G}\cdot\boldsymbol{\nabla}\times\mathbf{F}-\mathbf{F}\cdot\boldsymbol{\nabla}\times\mathbf{G}.$$

Proof. Each is a one-line computation in index notation. For example, $\partial_i(\varepsilon_{ijk}F_jG_k)=\varepsilon_{ijk}(\partial_i F_j)G_k+\varepsilon_{ijk}F_j\partial_i G_k$. The first term is $G_k(\nabla\times\mathbf{F})_k$, and the second, after the cyclic relabeling $\varepsilon_{ijk}=-\varepsilon_{jik}$, is $-F_j(\nabla\times\mathbf{G})_j$. ∎

Example 6.25 (Maxwell's equations). In vacuum with no charges or currents, Maxwell's equations read

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

Take the curl of the third equation and use Theorem 6.23: $\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\,\partial/\partial t\,(\nabla \times \mathbf{B}) = -\mu_0 \varepsilon_0\, \partial^2 \mathbf{E}/\partial t^2$. Since $\nabla \cdot \mathbf{E} = 0$,

$$\nabla^2 \mathbf{E} = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

This is the wave equation with speed $c = 1/\sqrt{\mu_0 \varepsilon_0} \approx 3.00 \times 10^8\,\mathrm{m/s}$, the measured speed of light. Maxwell's 1865 deduction that light is an electromagnetic wave is one of the most consequential calculations in the history of physics.

6.6 Green's Theorem

Let D be a bounded region in the plane whose boundary ∂D consists of finitely many piecewise smooth simple closed curves. The *positive orientation* of ∂D keeps D on the left as the boundary is traversed: the outer boundary counterclockwise, and any inner boundaries clockwise.

Theorem 6.26 (Green). *If P and Q are C^1 on an open set containing D, then*

$$\oint_{\partial D} P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA.$$

Proof. We prove $\oint_{\partial D} P\,dx = -\iint_D P_y\,dA$ for a type I region $D = \{a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$. By the fundamental theorem of calculus,

$$-\iint_D P_y\,dA = -\int_a^b \big[P(x, g_2(x)) - P(x, g_1(x))\big]\,dx.$$

Now compute the line integral. On the bottom curve, $y = g_1(x)$ traversed left to right, it contributes $\int_a^b P(x, g_1(x))\,dx$. On the top curve, traversed right to left, it contributes $-\int_a^b P(x, g_2(x))\,dx$. On any vertical sides $dx = 0$, so they contribute nothing. The two expressions agree. Symmetrically, $\oint Q\,dy = \iint Q_x\,dA$ for type II regions. A region that is both type I and type II therefore satisfies the theorem. A general region can be cut into such pieces. The line integrals along the cuts

are traversed twice in opposite directions and cancel, while the double integrals add. ∎

Green's theorem is a two-dimensional fundamental theorem of calculus. The integral of a derivative over a region equals a boundary term. In vector form, with $\mathbf{F} = \langle P, Q, 0\rangle$, it has two readings:

$$\oint_{\partial D} \mathbf{F} \cdot \hat{\mathbf{T}}\, ds = \iint_D (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{k}}\, dA \qquad \text{(circulation form)} \tag{6.1}$$

$$\oint_{\partial D} \mathbf{F} \cdot \hat{\mathbf{n}}\, ds = \iint_D \nabla \cdot \mathbf{F}\, dA \qquad \text{(flux form)} \tag{6.2}$$

The flux form follows by applying Theorem 6.26 to the field $\langle -Q, P\rangle$. The first says circulation around the boundary is the total curl inside. The second says flux out through the boundary is the total divergence inside. These are the planar cases of Stokes' theorem and the divergence theorem.

Example 6.27 (Using Green's theorem). Evaluate $\oint_C (3y - e^{\sin x})\, dx + \left(7x + \sqrt{y^4+1}\right) dy$, where C is the circle $x^2 + y^2 = 9$ traversed counterclockwise.

Solution. Direct evaluation is hopeless. But $Q_x - P_y = 7 - 3 = 4$, so the integral is $4 \cdot \text{area} = 4 \cdot 9\pi = 36\pi$.

Corollary 6.28 (Area from a boundary integral). *The area of D is* $\displaystyle\oint_{\partial D} x\, dy = -\oint_{\partial D} y\, dx = \frac{1}{2}\oint_{\partial D} x\, dy - y\, dx.$

Proof. Each of the fields $\langle 0, x\rangle$, $\langle -y, 0\rangle$, and $\frac{1}{2}\langle -y, x\rangle$ has $Q_x - P_y = 1$. ∎

This is the principle of the *planimeter*, a mechanical instrument that measures the area of a region by tracing its boundary. It also explains the earlier observation that the circulation of $\langle -y, x\rangle$ is twice the enclosed area.

Example 6.29 (Area of an ellipse). With $x = a\cos t$, $y = b\sin t$: $A = \frac{1}{2}\int_0^{2\pi} (ab\cos^2 t + ab\sin^2 t)\, dt = \pi ab$.

Example 6.30 (Area of a polygon: the shoelace formula). For a polygon with vertices $(x_1, y_1), \ldots, (x_n, y_n)$ listed counterclockwise, apply Corollary 6.28 edge by edge. On the segment from (x_i, y_i) to (x_{i+1}, y_{i+1}), $\int x\, dy -$

$y\,dx = x_i y_{i+1} - x_{i+1} y_i$. So $A = \frac{1}{2}\sum_{i=1}^{n}(x_i y_{i+1} - x_{i+1} y_i)$, with indices taken modulo n. For the quadrilateral $(0,0)$, $(4,0)$, $(3,2)$, $(1,3)$: $A = \frac{1}{2}(0+8+7+0) = \frac{15}{2}$.

6.6.1 REGIONS WITH HOLES

Example 6.31 (The winding field revisited). Let $\mathbf{F} = \langle -y, x\rangle / (x^2+y^2)$, and let C be any positively oriented simple closed curve enclosing the origin. Show that $\oint_C \mathbf{F}\cdot d\mathbf{r} = 2\pi$.

Solution. Choose a small circle C_a of radius a about the origin, inside C, and let D be the region between them. Its positive boundary is C together with C_a traversed clockwise. Green's theorem applies on D, which excludes the origin, and $Q_x - P_y = 0$ there. So $\oint_C \mathbf{F}\cdot d\mathbf{r} - \oint_{C_a} \mathbf{F}\cdot d\mathbf{r} = 0$. The circle integral is 2π by the computation in the pitfall of the previous section. A curve that does not enclose the origin gives 0, and a curve that winds n times around it gives $2\pi n$. The integral counts windings: it is 2π times the *winding number*.

Corollary 6.32 (Simply connected regions). *If P, Q are C^1 on an open simply connected region D and $P_y = Q_x$ there, then $\langle P, Q\rangle$ is conservative on D.*

Proof. For a simple closed curve C in D, the region it encloses lies in D because D has no holes, so Green's theorem gives $\oint_C = \iint (Q_x - P_y)\,dA = 0$. A general closed curve can be decomposed into simple loops, and Theorem 6.16 applies. ∎

LOOKING AHEAD. The winding-number integral of Example 6.31 is the prototype of Cauchy's integral theorem in complex analysis (Chapter 8). There, Green's theorem shows that integrals of analytic functions around closed curves vanish, except for contributions from singularities: the *residues*. The residue theorem then evaluates real integrals, such as $\int_0^\infty dx/(1+x^4)$, that resist every other method.

6.7 WORKED PROBLEMS

The problems below connect the three faces of this chapter: line integrals computed directly, the test for conservative fields, and Green's theorem as a bridge between line and double integrals. Two problems develop vector identities used constantly in electromagnetism and quantum mechanics.

Problem 6.1 Making a field conservative

Find constants a and b for which $\mathbf{F} = \langle 2xy + z,\ ax^2 + ze^{yz},\ ye^{yz} + bx\rangle$ is conservative on $\mathbb{R}^3$. For those values, find a potential and the work done by $\mathbf{F}$ along any path from $(0,0,0)$ to $(1,2,1)$.

Solution. $\mathbb{R}^3$ is simply connected, so by Theorem 6.20 $\mathbf{F}$ is conservative if and only if $\nabla \times \mathbf{F} = \mathbf{0}$. The three conditions are

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}: \quad 2x = 2ax, \qquad \frac{\partial F_1}{\partial z} = \frac{\partial F_3}{\partial x}: \quad 1 = b,$$
$$\frac{\partial F_2}{\partial z} = \frac{\partial F_3}{\partial y}: \quad e^{yz} + yze^{yz} = e^{yz} + yze^{yz}.$$

The last holds identically, and the others force $a = 1$ and $b = 1$.

To find f with $\nabla f = \mathbf{F}$, integrate the first component in x: $f = x^2y + xz + g(y,z)$. Then $f_y = x^2 + g_y = x^2 + ze^{yz}$ gives $g_y = ze^{yz}$, so $g = e^{yz} + h(z)$. Finally $f_z = x + ye^{yz} + h'(z) = ye^{yz} + x$ gives $h' = 0$. So

$$f(x,y,z) = x^2y + xz + e^{yz}.$$

By the fundamental theorem for line integrals (Theorem 6.14), the work is $f(1,2,1) - f(0,0,0) = (2 + 1 + e^2) - 1 = 2 + e^2 \approx 9.39$, for every path.

Problem 6.2 Path dependence measured by Green's theorem

Let $\mathbf{F} = \langle x^2,\ xy\rangle$. Compute $\int \mathbf{F}\cdot d\mathbf{r}$ from $(0,0)$ to $(1,1)$ along the line $y = x$ and along the parabola $y = x^2$. Explain the difference using Green's theorem.

Solution. Along the line, $\mathbf{r} = \langle t, t\rangle$: $\int_0^1 (t^2 + t^2)\,dt = \frac{2}{3}$. Along the parabola, $\mathbf{r} = \langle t, t^2\rangle$ and $d\mathbf{r} = \langle 1, 2t\rangle\ dt$: $\int_0^1 (t^2 + t\cdot t^2\cdot 2t)\,dt = \frac{1}{3} + \frac{2}{5} = \frac{11}{15}$.

The difference is $\frac{11}{15} - \frac{2}{3} = \frac{1}{15}$. Out along the parabola and back along the line is a closed curve C that encloses the region D between them, $x^2 \le y \le x$, counterclockwise (the parabola lies below the line). Green's theorem gives

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \iint_D \Big(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\Big)dA = \int_0^1\int_{x^2}^{x} y\,dy\,dx = \frac{1}{2}\int_0^1 (x^2 - x^4)\,dx = \frac{1}{15}.$$

So the path dependence is exactly the flux of the curl through the region between the paths. Two paths give the same work precisely when the curl integrates to zero between them.

Problem 6.3 The area of the loop of the folium

Find the area enclosed by the loop of the folium of Descartes, $x^3 + y^3 = 3xy$.

Solution. Lines through the origin, $y = tx$, meet the curve where $x^3(1 + t^3) = 3tx^2$, giving the rational parametrization

$$x = \frac{3t}{1+t^3}, \qquad y = \frac{3t^2}{1+t^3}.$$

As t runs from 0 to ∞, the point leaves the origin along the x-axis ($y/x = t$ small), goes around the loop counterclockwise, and returns to the origin tangent to the y-axis. By Corollary 6.28, $A = \frac{1}{2}\oint (x\,dy - y\,dx)$. Writing $y = tx$,

$$x\,dy - y\,dx = x(t\,dx + x\,dt) - tx\,dx = x^2\,dt.$$

This identity, valid for any curve written as $y = tx$, avoids differentiating the quotients. Then

$$A = \frac{1}{2}\int_0^\infty \frac{9t^2}{(1+t^3)^2}\,dt = \frac{3}{2}\Big[-\frac{1}{1+t^3}\Big]_0^\infty = \frac{3}{2}.$$

Notice what the formula is computing. With $\theta = \arctan t$ the polar angle, $r^2 = x^2(1+t^2)$ and $d\theta = dt/(1+t^2)$, so $x^2\,dt = r^2\,d\theta$. Green's area formula has reproduced the polar area element $\frac{1}{2}r^2\,d\theta$, and the parameter $t = \tan\theta$ is the slope of the ray from the origin.

Problem 6.4 Circulation around a half-annulus

Evaluate $\oint_C y^2\,dx + 3xy\,dy$, where C is the positively oriented boundary of the region D in the upper half-plane between the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$.

Solution. The boundary has four pieces: two semicircles and two segments of the x-axis. Computing four line integrals is possible, but Green's theorem replaces them with one double integral. With $P = y^2$ and $Q = 3xy$, $Q_x - P_y = 3y - 2y = y$, so

$$\begin{aligned}\oint_C y^2\,dx + 3xy\,dy &= \iint_D y\,dA = \int_0^\pi \int_1^2 r\sin\theta\; r\,dr\,d\theta \\ &= \Big[\frac{r^3}{3}\Big]_1^2 \cdot \Big[-\cos\theta\Big]_0^\pi = \frac{7}{3}\cdot 2 = \frac{14}{3}.\end{aligned}$$

As a partial check, the two segments on the x-axis contribute nothing, since $y = 0$ and $dy = 0$ there. Then the outer semicircle, traversed counterclockwise, and the inner semicircle, traversed clockwise, must together give $\frac{14}{3}$. On the outer one, $\mathbf{r} = \langle 2\cos t, 2\sin t\rangle$ gives $\int_0^{\pi} \big(4\sin^2 t(-2\sin t) + 12\cos t \sin t \cdot 2\cos t\big)dt = -\frac{32}{3} + 16 = \frac{16}{3}$. On a circle of radius a the same computation gives $\frac{2a^3}{3}$, so the inner semicircle, traversed backward, gives $-\frac{2}{3}$. The total is $\frac{16}{3} - \frac{2}{3} = \frac{14}{3}$.

Problem 6.5 Mass and center of mass of a cardioid wire

A uniform wire has the shape of the cardioid $r = 1 + \cos\theta$. Find its length and its center of mass.

Solution. With $x = (1+\cos\theta)\cos\theta$ and $y = (1+\cos\theta)\sin\theta$, the polar arc-length formula $ds^2 = (r^2 + r'^2)\,d\theta^2$ gives

$$r^2 + r'^2 = (1+\cos\theta)^2 + \sin^2\theta = 2 + 2\cos\theta = 4\cos^2\tfrac{\theta}{2}, \qquad ds = 2\big|\cos\tfrac{\theta}{2}\big|\,d\theta.$$

The absolute value matters: $\cos\frac{\theta}{2} < 0$ for $\pi < \theta < 2\pi$. Use symmetry and double the integral over $[0,\pi]$:

$$L = 2\int_0^{\pi} 2\cos\tfrac{\theta}{2}\,d\theta = 8.$$

By symmetry $\bar{y} = 0$. For $\bar{x}$, using $1+\cos\theta = 2\cos^2\frac{\theta}{2}$ and $\cos\theta = 2\cos^2\frac{\theta}{2} - 1$, and then the substitution $u = \sin\frac{\theta}{2}$,

$$\begin{aligned}\int_C x\,ds &= 2\int_0^{\pi} 2\cos^2\tfrac{\theta}{2}\big(2\cos^2\tfrac{\theta}{2} - 1\big)2\cos\tfrac{\theta}{2}\,d\theta \\ &= 16\int_0^1 \big(2(1-u^2)^2 - (1-u^2)\big)du = 16\Big(\frac{16}{15} - \frac{2}{3}\Big) = \frac{32}{5}.\end{aligned}$$

So $\bar{x} = \frac{32/5}{8} = \frac{4}{5}$, and the center of mass is $(\frac{4}{5}, 0)$. For comparison, the centroid of the region the cardioid encloses is at $(\frac{5}{6}, 0)$, since $\iint x\,dA = \frac{1}{3}\int_0^{2\pi}(1+\cos\theta)^3\cos\theta\,d\theta = \frac{5\pi}{4}$ and the area is $\frac{3\pi}{2}$. A wire and a plate of the same outline do not balance at the same point.

Problem 6.6 Divergence and Laplacian of powers of r

Let $\mathbf{r} = \langle x, y, z\rangle$ and $r = |\mathbf{r}|$. For $r \neq 0$, prove

$$\nabla\cdot\big(r^n\mathbf{r}\big) = (n+3)r^n, \qquad \nabla^2\big(r^n\big) = n(n+1)r^{n-2}.$$

Which powers of r are harmonic away from the origin? Which radial fields $r^n\mathbf{r}$ are divergence-free?

Solution. By the product rule (Theorem 6.24), $\nabla \cdot (f\mathbf{G}) = \nabla f \cdot \mathbf{G} + f\nabla \cdot \mathbf{G}$. Here $\nabla r^n = nr^{n-2}\mathbf{r}$ (Example 6.4) and $\nabla \cdot \mathbf{r} = 3$, so

$$\nabla \cdot \big(r^n \mathbf{r}\big) = nr^{n-2}\mathbf{r} \cdot \mathbf{r} + 3r^n = (n+3)r^n.$$

For the Laplacian, $\nabla^2(r^n) = \nabla \cdot \nabla(r^n) = \nabla \cdot \big(nr^{n-2}\mathbf{r}\big) = n\big((n-2)+3\big)r^{n-2} = n(n+1)r^{n-2}$.

So r^n is harmonic away from the origin exactly for $n = 0$ (constants) and $n = -1$: the potential $1/r$ of a point charge. The field $r^n\mathbf{r}$ is divergence-free exactly for $n = -3$, that is, $\mathbf{r}/r^3 = \hat{\mathbf{r}}/r^2$, the inverse-square field. These are the same fact, since $\nabla(1/r) = -\mathbf{r}/r^3$.

In $\mathbb{R}^d$, $\nabla \cdot \mathbf{r} = d$, and the same computation gives $\nabla^2(r^n) = n(n+d-2)r^{n-2}$. The fundamental harmonic function is r^{2-d} for $d \geq 3$, and $\ln r$ in the plane. The inverse-square law is tied to the three dimensions of space.

Problem 6.7 The divergence of a cross product

Prove that $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot (\nabla \times \mathbf{F}) - \mathbf{F} \cdot (\nabla \times \mathbf{G})$. Use it to show that, in a region of space with no charges or currents, the energy flux of the electromagnetic field, the *Poynting vector* $\mathbf{S} = \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}$, satisfies

$$\nabla \cdot \mathbf{S} = -\frac{\partial}{\partial t}\Big(\frac{\varepsilon_0}{2}|\mathbf{E}|^2 + \frac{1}{2\mu_0}|\mathbf{B}|^2\Big).$$

Solution. In index notation, with summation over repeated indices,

$$\partial_i(\varepsilon_{ijk}F_jG_k) = \varepsilon_{ijk}(\partial_i F_j)G_k + \varepsilon_{ijk}F_j(\partial_i G_k).$$

In the first term, $\varepsilon_{ijk}\partial_i F_j = \varepsilon_{kij}\partial_i F_j = (\nabla \times \mathbf{F})_k$, giving $\mathbf{G} \cdot \nabla \times \mathbf{F}$. In the second, $\varepsilon_{ijk}\partial_i G_k = -\varepsilon_{jik}\partial_i G_k = -(\nabla \times \mathbf{G})_j$, giving $-\mathbf{F} \cdot \nabla \times \mathbf{G}$.

Now take $\mathbf{F} = \mathbf{E}$, $\mathbf{G} = \mathbf{B}$, and use the Maxwell equations of Example 6.25:

$$\mu_0 \nabla \cdot \mathbf{S} = \mathbf{B} \cdot \Big(-\frac{\partial \mathbf{B}}{\partial t}\Big) - \mathbf{E} \cdot \Big(\mu_0\varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}\Big) = -\frac{\partial}{\partial t}\Big(\frac{1}{2}|\mathbf{B}|^2 + \frac{\mu_0\varepsilon_0}{2}|\mathbf{E}|^2\Big).$$

Dividing by μ_0 gives the result. This is a *continuity equation*: by the divergence theorem of Chapter 7, the rate at which field energy leaves a region equals the flux of $\mathbf{S}$ out of its boundary. Energy is conserved locally. The same structure appears for quantum probability in Chapter 13.

EXERCISES

A. COMPUTATION

6.1 Sketch the vector field $\mathbf{F} = \langle y,\, x\rangle$ and find its flow lines.

6.2 Find the flow lines of $\mathbf{F} = \langle x,\, -y\rangle$ and describe them.

6.3 Find the gradient field of $f(x,y,z) = \sqrt{x^2+y^2+z^2}$ and describe it.

6.4 Evaluate $\int_C xy^4\,ds$, where C is the right half of the circle $x^2+y^2=16$.

6.5 Evaluate $\int_C (x^2+y^2+z^2)\,ds$ along the helix $\langle \cos t, \sin t, t\rangle$, $0 \le t \le 2\pi$.

6.6 Find the mass of a wire along $y = x^2$ from $(0,0)$ to $(1,1)$ with density $\rho = x$.

6.7 Evaluate $\int_C \mathbf{F}\cdot d\mathbf{r}$ for $\mathbf{F} = \langle xy,\, yz,\, zx\rangle$ along the twisted cubic $\langle t, t^2, t^3\rangle$, $0 \le t \le 1$.

6.8 Evaluate $\int_C \mathbf{F}\cdot d\mathbf{r}$ for $\mathbf{F} = \langle \sin x,\, \cos y,\, xz\rangle$ along $\langle t^3, -t^2, t\rangle$, $0 \le t \le 1$.

6.9 Find the work done by $\mathbf{F} = \langle x^2,\, xy\rangle$ moving a particle once counterclockwise around the circle $x^2+y^2=4$.

6.10 Evaluate $\int_C y\,dx + z\,dy + x\,dz$ along the line segment from $(2,0,0)$ to $(3,4,5)$ followed by the segment to $(3,4,0)$.

6.11 Find the flux of $\mathbf{F} = \langle x-y,\, x\rangle$ outward across the unit circle, and its circulation around it.

6.12 Determine whether $\mathbf{F} = \langle 2xy+y^2,\; x^2+2xy\rangle$ is conservative; if so, find a potential.

6.13 Determine whether $\mathbf{F} = \langle ye^x + \sin y,\; e^x + x\cos y\rangle$ is conservative; if so, find a potential.

6.14 Find a potential for $\mathbf{F} = \langle yz,\, xz,\, xy+2z\rangle$ and evaluate $\int_C \mathbf{F}\cdot d\mathbf{r}$ from $(1,0,-2)$ to $(4,6,3)$.

6.15 Find a potential for $\mathbf{F} = \langle e^y,\; xe^y + e^z,\; ye^z\rangle$ and evaluate the integral from $(0,2,0)$ to $(4,0,3)$.

6.16 Find the work done by the gravitational field $\mathbf{F} = -GMm\,\mathbf{r}/r^3$ moving a particle from $(3,4,12)$ to $(2,2,1)$.

6.17 Compute $\nabla\cdot\mathbf{F}$ and $\nabla\times\mathbf{F}$ for $\mathbf{F} = \langle xyz,\; 0,\; -x^2y\rangle$.

6.18 Compute $\nabla\cdot\mathbf{F}$ and $\nabla\times\mathbf{F}$ for $\mathbf{F} = \langle e^x\sin y,\; e^y\sin z,\; e^z\sin x\rangle$.

6.19 Is there a vector field $\mathbf{G}$ with $\nabla\times\mathbf{G} = \langle x\sin y,\; \cos y,\; z - xy\rangle$? Explain.

6.20 Show that $\mathbf{F} = \langle y^2z^3,\ 2xyz^3,\ 3xy^2z^2\rangle$ is curl-free on $\mathbb{R}^3$ and find a potential.

6.21 Compute $\nabla^2(r^n)$ in $\mathbb{R}^3$, where $r = |\mathbf{r}|$, and find all n for which r^n is harmonic away from the origin.

6.22 Use Green's theorem to evaluate $\oint_C (y + e^{\sqrt{x}})\,dx + (2x + \cos y^2)\,dy$, where C bounds the region between $y = x^2$ and $x = y^2$, positively oriented.

6.23 Use Green's theorem to evaluate $\oint_C x^2y\,dx - xy^2\,dy$ around the circle $x^2 + y^2 = 4$, counterclockwise.

6.24 Use Green's theorem to find the work done by $\mathbf{F} = \langle x(x+y),\ xy^2\rangle$ around the triangle $(0,0) \to (1,0) \to (0,1) \to (0,0)$.

6.25 Use Corollary 6.28 to find the area enclosed by the astroid $\langle \cos^3 t, \sin^3 t\rangle$.

6.26 Use the shoelace formula to find the area of the pentagon with vertices $(0,0)$, $(2,1)$, $(1,3)$, $(0,2)$, $(-1,1)$.

6.27 Find the flux of $\mathbf{F} = \langle x^3,\ y^3\rangle$ outward across the circle $x^2 + y^2 = 1$ using the flux form of Green's theorem.

6.28 Evaluate $\oint_C \dfrac{x\,dy - y\,dx}{x^2 + y^2}$ around the ellipse $x^2/4 + y^2/9 = 1$, counterclockwise, and around the circle $(x-3)^2 + y^2 = 1$.

B. THEORY AND PROOF

6.29 Prove that $\int_C f\,ds$ is independent of parametrization and of orientation.

6.30 Prove that $\boldsymbol{\nabla} \cdot (\nabla f \times \nabla g) = 0$ for all C^2 functions f and g.

6.31 Prove $\nabla(\mathbf{F}\cdot\mathbf{G}) = (\mathbf{F}\cdot\nabla)\mathbf{G} + (\mathbf{G}\cdot\nabla)\mathbf{F} + \mathbf{F}\times(\boldsymbol{\nabla}\times\mathbf{G}) + \mathbf{G}\times(\boldsymbol{\nabla}\times\mathbf{F})$.

6.32 Show that the curl of $\mathbf{v} = \boldsymbol{\omega}\times\mathbf{r}$ is $2\boldsymbol{\omega}$ for constant $\boldsymbol{\omega}$, using index notation.

6.33 Prove *Green's first identity*: $\iint_D f\nabla^2 g\,dA = \oint_{\partial D} f\,(\nabla g\cdot\hat{\mathbf{n}})\,ds - \iint_D \nabla f\cdot\nabla g\,dA$. Deduce *Green's second identity* by antisymmetrizing in f and g.

6.34 Show that a harmonic function on D that vanishes on ∂D is identically zero. (Use Green's first identity with $f = g$.) Deduce that the Dirichlet problem for Laplace's equation has at most one solution.

6.35 Prove that if $\mathbf{F}$ is a central field $\mathbf{F} = g(r)\,\hat{\mathbf{r}}$ with g continuous, then $\mathbf{F}$ is conservative on $\mathbb{R}^3\setminus\{\mathbf{0}\}$ and find a potential.

6.36 Show that the flux form (6.2) follows from Theorem 6.26.

6.37 Let $\mathbf{F}$ be the winding field of Example 6.31. Show that $\mathbf{F} = \nabla\theta$ on the plane slit along the negative x-axis, and explain why this is consistent with Theorem 6.20.

C. CHALLENGE

6.38 (Cardioid gear.) A circle of radius 1 rolls without slipping around the outside of a fixed circle of radius 1. Show that a point on the rolling circle traces the cardioid $\langle 2\cos t - \cos 2t,\ 2\sin t - \sin 2t\rangle$, and use Corollary 6.28 to find the area it encloses.

6.39 Prove that the centroid of a region D satisfies $\bar{x} = \dfrac{1}{2A}\oint_{\partial D} x^2\,dy$ and $\bar{y} = -\dfrac{1}{2A}\oint_{\partial D} y^2\,dx$. Use this to find the centroid of the half-disk $x^2 + y^2 \le 1$, $y \ge 0$.

6.40 The vector potential of an infinite solenoid of radius R carrying a uniform field $B\hat{\mathbf{k}}$ inside is $\mathbf{A} = \frac{B}{2}\langle -y, x, 0\rangle$ for $\rho < R$ and $\mathbf{A} = \frac{BR^2}{2}\langle -y, x, 0\rangle / \rho^2$ for $\rho > R$, where $\rho^2 = x^2 + y^2$. Verify that $\boldsymbol{\nabla}\times\mathbf{A} = B\hat{\mathbf{k}}$ inside and $\mathbf{0}$ outside, and compute $\oint \mathbf{A}\cdot d\mathbf{r}$ around any circle $\rho = a > R$. Compare with the enclosed magnetic flux.

6.41 Show that the Laplacian in the plane can be written $\nabla^2 f = \lim_{a\to 0}\dfrac{4}{a^2}\big(\bar{f}_a - f(\mathbf{x})\big)$, where $\bar{f}_a$ is the average of f over the circle of radius a about $\mathbf{x}$. (Use Taylor's theorem.) Deduce the mean value property of harmonic functions.

CHAPTER 7

Surface Integrals and the Integral Theorems

LEARNING OBJECTIVES

- Parametrize surfaces, compute tangent planes and normals, and find surface area.
- Evaluate surface integrals of scalar functions and flux integrals of vector fields across oriented surfaces.
- State, prove (for simple regions), and apply Stokes' theorem and the divergence theorem.
- Interpret curl as circulation density and divergence as flux density, and derive Gauss's law and the continuity equation.
- Use the Dirac delta function to describe point sources, including $\nabla^2(1/r) = -4\pi\delta^3(\mathbf{r})$.
- Derive the gradient, divergence, curl, and Laplacian in orthogonal curvilinear coordinates, especially spherical coordinates.

This chapter completes the calculus of Part I. We learn to integrate over surfaces, and then prove the two theorems that crown vector calculus. *Stokes' theorem* equates the circulation of a field around a closed curve with the flux of its curl through any surface the curve bounds. The *divergence theorem* equates the flux of a field out of a closed surface with the integral of its divergence over the enclosed volume. Together with the fundamental theorem of calculus, the fundamental theorem for line integrals, and Green's theorem, they form a single pattern. The integral of a derivative over a region equals an integral over its boundary. The chapter ends by expressing ∇^2 in spherical coordinates, the form in which it appears in the hydrogen atom.

7.1 PARAMETRIC SURFACES

Definition 7.1 (Parametric surface). A *parametric surface* is the image S of a continuous map $\mathbf{r} : D \to \mathbb{R}^3$ from a region D of the uv-plane,

$$\mathbf{r}(u,v) = \langle x(u,v),\, y(u,v),\, z(u,v)\rangle\,.$$

The curves $u = \text{const}$ and $v = \text{const}$ on S are the *grid curves*. If $\mathbf{r}$ is C^1, the vectors $\mathbf{r}_u = \partial\mathbf{r}/\partial u$ and $\mathbf{r}_v = \partial\mathbf{r}/\partial v$ are tangent to the grid curves. The surface is *smooth* where $\mathbf{r}_u \times \mathbf{r}_v \neq \mathbf{0}$, and there

$$\mathbf{N} = \mathbf{r}_u \times \mathbf{r}_v$$

is a normal vector and the tangent plane is spanned by $\mathbf{r}_u$ and $\mathbf{r}_v$.

Example 7.2 (Standard parametrizations). (a) *Graph* $z = g(x,y)$: $\mathbf{r}(x,y) = \langle x, y, g(x,y)\rangle$, with $\mathbf{N} = \langle -g_x, -g_y, 1\rangle$.

(b) *Sphere* of radius a: $\mathbf{r}(\theta,\phi) = a\langle \sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta\rangle$ on $[0,\pi]\times[0,2\pi]$. Then $\mathbf{N} = \mathbf{r}_\theta \times \mathbf{r}_\phi = a^2\sin\theta\ \hat{\mathbf{r}}$, which points outward, with $|\mathbf{N}| = a^2\sin\theta$.

(c) *Cylinder* $x^2 + y^2 = a^2$: $\mathbf{r}(\phi,z) = \langle a\cos\phi, a\sin\phi, z\rangle$, with $\mathbf{N} = a\langle\cos\phi,\sin\phi,0\rangle$.

(d) *Surface of revolution* obtained by rotating $y = f(x) \geq 0$, $a \leq x \leq b$, about the x-axis: $\mathbf{r}(x,\phi) = \langle x,\ f(x)\cos\phi,\ f(x)\sin\phi\rangle$.

(e) *Torus* with tube radius b and center-circle radius $a > b$: $\mathbf{r}(\alpha,\phi) = \langle (a + b\cos\alpha)\cos\phi,\ (a+b\cos\alpha)\sin\phi,\ b\sin\alpha\rangle$.

Example 7.3 (Tangent plane to a parametric surface). Find the tangent plane to $\mathbf{r}(u,v) = \langle u^2,\ v^2,\ u+2v\rangle$ at $(1,1,3)$.

Solution. The point corresponds to $u = v = 1$. There $\mathbf{r}_u = \langle 2,0,1\rangle$ and $\mathbf{r}_v = \langle 0,2,2\rangle$, so $\mathbf{N} = \langle -2,-4,4\rangle$. The plane is $-2(x-1) - 4(y-1) + 4(z-3) = 0$, or $x + 2y - 2z + 3 = 0$.

7.2 Surface Area

A small rectangle $[u, u+\Delta u] \times [v, v+\Delta v]$ in D maps to a patch of S that is approximately the parallelogram spanned by $\mathbf{r}_u \Delta u$ and $\mathbf{r}_v \Delta v$. Its area is $|\mathbf{r}_u \times \mathbf{r}_v|\, \Delta u\, \Delta v$.

Definition 7.4 (Surface area). For a smooth surface S given by $\mathbf{r} : D \to \mathbb{R}^3$, one-to-one on the interior of D,

$$A(S) = \iint_D |\mathbf{r}_u \times \mathbf{r}_v|\, dA, \qquad dS = |\mathbf{r}_u \times \mathbf{r}_v|\, du\, dv.$$

For a graph $z = g(x,y)$ over D, $dS = \sqrt{1 + g_x^2 + g_y^2}\, dx\, dy$.

Example 7.5 (Area of a sphere). With $|\mathbf{N}| = a^2 \sin\theta$: $A = \int_0^{2\pi} \int_0^{\pi} a^2 \sin\theta\, d\theta\, d\phi = 4\pi a^2$. The area element $dS = a^2\, d\Omega$, with $d\Omega = \sin\theta\, d\theta\, d\phi$ the element of *solid angle*, appears in every scattering calculation and in every normalization of an angular wavefunction.

Example 7.6 (Archimedes' hat-box theorem). The area of the zone of a sphere of radius a between two parallel planes a distance h apart is $2\pi a h$, independent of where the planes cut the sphere.

Solution. For $z_1 \le z \le z_2$, i.e. $\cos\theta_2 \le \cos\theta \le \cos\theta_1$: $A = 2\pi a^2 \int_{\theta_1}^{\theta_2} \sin\theta\, d\theta = 2\pi a^2(\cos\theta_1 - \cos\theta_2) = 2\pi a(z_2 - z_1)$. This is also the lateral area of the circumscribed cylinder between the same planes, which is the content of Archimedes' theorem.

Example 7.7 (Area of a torus). For the torus of Example 7.2(e), a computation gives $|\mathbf{r}_\alpha \times \mathbf{r}_\phi| = b(a + b\cos\alpha)$, so $A = \int_0^{2\pi} \int_0^{2\pi} b(a + b\cos\alpha)\, d\alpha\, d\phi = 4\pi^2 a b$. This is the circumference $2\pi b$ of the tube times the distance $2\pi a$ travelled by its center, as Pappus's theorem predicts.

7.3 Surface Integrals

Definition 7.8 (Surface integral of a scalar function). For f continuous on

a smooth surface S parametrized by $\mathbf{r} : D \to \mathbb{R}^3$,

$$\iint_S f\, dS = \iint_D f(\mathbf{r}(u,v))\, |\mathbf{r}_u \times \mathbf{r}_v|\, du\, dv.$$

For a piecewise smooth surface we sum over the smooth pieces.

As with line integrals, the value is independent of the parametrization. If f is a surface density, $\iint_S f\, dS$ is the total mass or charge.

Example 7.9 (Average of z^2 over a sphere). $\iint_{x^2+y^2+z^2=a^2} z^2\, dS = \int_0^{2\pi}\int_0^{\pi} a^2\cos^2\theta\ a^2\sin\theta\, d\theta\, d\phi = \dfrac{4\pi a^4}{3}$. Dividing by the area, the average of z^2 is $a^2/3$. By symmetry the averages of x^2 and y^2 are the same, and they sum to a^2, as they must.

7.3.1 ORIENTED SURFACES AND FLUX

Definition 7.10 (Orientation). A smooth surface S is *orientable* if there is a continuous choice of unit normal $\hat{\mathbf{n}}$ on all of S. Such a choice is an *orientation*. A closed surface, one that bounds a solid, is *positively oriented* when $\hat{\mathbf{n}}$ points outward.

Every surface has two sides locally, but not globally. The *Möbius strip*, made by giving a strip of paper a half twist before joining its ends, has only one side. A normal vector carried once around it returns pointing the opposite way. Flux integrals are defined only for orientable surfaces.

Definition 7.11 (Flux integral). The *flux* of a continuous field $\mathbf{F}$ across an oriented surface S with unit normal $\hat{\mathbf{n}}$ is

$$\iint_S \mathbf{F}\cdot d\mathbf{S} = \iint_S \mathbf{F}\cdot\hat{\mathbf{n}}\, dS = \pm\iint_D \mathbf{F}(\mathbf{r}(u,v))\cdot(\mathbf{r}_u\times\mathbf{r}_v)\, du\, dv,$$

with the sign $+$ if $\mathbf{r}_u \times \mathbf{r}_v$ points along the chosen orientation and $-$ otherwise.

If $\mathbf{F}$ is the velocity field of a fluid with density ρ, then $\iint_S \rho\mathbf{F}\cdot d\mathbf{S}$ is the mass per unit time crossing S in the direction of $\hat{\mathbf{n}}$. For a graph $z = g(x,y)$ oriented

upward and $\mathbf{F} = \langle P, Q, R \rangle$,

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_D \left(-Pg_x - Qg_y + R \right) dA. \tag{7.1}$$

Example 7.12 (Flux of an inverse-square field). Find the outward flux of $\mathbf{F} = \mathbf{r}/r^3$ through the sphere of radius a.

Solution. On the sphere $\mathbf{F} = \hat{\mathbf{r}}/a^2$ and $\hat{\mathbf{n}} = \hat{\mathbf{r}}$, so $\mathbf{F} \cdot \hat{\mathbf{n}} = 1/a^2$ and the flux is $(1/a^2)(4\pi a^2) = 4\pi$, independent of a. Compare Example 6.13. The flux of an inverse-square field through a sphere centered at the source does not depend on the radius: the field weakens as $1/a^2$ while the area grows as a^2.

Example 7.13 (Flux through a paraboloid). Find the flux of $\mathbf{F} = \langle y, x, z \rangle$ upward through the part of $z = 1 - x^2 - y^2$ above the xy-plane.

Solution. By (7.1) with $g_x = -2x$ and $g_y = -2y$, the integrand is $2xy + 2xy + 1 - x^2 - y^2$. In polar coordinates, $\int_0^{2\pi} \int_0^1 \left(4r^2 \cos\theta \sin\theta + 1 - r^2\right) r \, dr \, d\theta = 0 + 2\pi \cdot \frac{1}{4} = \frac{\pi}{2}$.

7.4 Stokes' Theorem

Let S be an oriented, piecewise smooth surface bounded by a simple closed, piecewise smooth curve ∂S. The *induced orientation* on ∂S is the one for which, walking along the boundary with your head pointing along $\hat{\mathbf{n}}$, the surface lies on your left.

Theorem 7.14 (Stokes). *If $\mathbf{F}$ is C^1 on an open region containing S, then*

$$\oint_{\partial S} \mathbf{F} \cdot d\mathbf{r} = \iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S}.$$

Proof for graphs. Let S be the graph $z = g(x, y)$ over a region D to which Green's theorem applies, oriented upward, with g of class C^2. Write $\mathbf{F} = \langle P, Q, R \rangle$. If ∂D is parametrized by $(x(t), y(t))$, then ∂S is $(x, y, g(x, y))$, and $dz = g_x dx + g_y dy$. Hence

$$\oint_{\partial S} P\,dx + Q\,dy + R\,dz = \oint_{\partial D} (P + Rg_x)\,dx + (Q + Rg_y)\,dy,$$

where P, Q, R are evaluated at $(x, y, g(x, y))$. By Green's theorem this equals $\iint_D \left[\partial_x(Q + Rg_y) - \partial_y(P + Rg_x)\right] dA$. Expanding with the chain rule, and noting

that the terms Rg_{xy} and Rg_{yx} cancel by Clairaut's theorem, the integrand is

$$-(R_y - Q_z)g_x - (P_z - R_x)g_y + (Q_x - P_y).$$

By (7.1), this is precisely the integrand of $\iint_S \nabla \times \mathbf{F} \cdot d\mathbf{S}$. ∎

A general surface is divided into pieces that are graphs over one of the coordinate planes. Integrals along interior cuts cancel in pairs, exactly as in the proof of Green's theorem.

Example 7.15 (Using Stokes to compute a line integral). Evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$ for $\mathbf{F} = \langle -y^2,\, x,\, z^2 \rangle$, where C is the intersection of the plane $y + z = 2$ with the cylinder $x^2 + y^2 = 1$, oriented counterclockwise as seen from above.

Solution. $\nabla \times \mathbf{F} = \langle 0, 0, 1 + 2y \rangle$. Take S to be the elliptical disk in the plane $z = 2 - y$ inside the cylinder, oriented upward. By (7.1) only the R-component matters:

$$\iint_{x^2+y^2 \le 1} (1 + 2y)\, dA = \pi + 0 = \pi.$$

Example 7.16 (Surface independence). Stokes' theorem implies that the flux of $\nabla \times \mathbf{F}$ is the same through *every* surface with the same boundary. For instance, the flux of $\nabla \times \mathbf{F}$ through the upper hemisphere $x^2 + y^2 + z^2 = 1$, $z \ge 0$, equals its flux through the flat disk $x^2 + y^2 \le 1$, $z = 0$, which is usually far easier to compute.

Theorem 7.17 (Curl as circulation density). *Let $\hat{\mathbf{n}}$ be a unit vector and D_a the disk of radius a centered at $\mathbf{x}$ and perpendicular to $\hat{\mathbf{n}}$, with boundary circle C_a oriented by $\hat{\mathbf{n}}$. Then*

$$\nabla \times \mathbf{F}(\mathbf{x}) \cdot \hat{\mathbf{n}} = \lim_{a \to 0} \frac{1}{\pi a^2} \oint_{C_a} \mathbf{F} \cdot d\mathbf{r}.$$

Proof. By Stokes' theorem, the circulation equals $\iint_{D_a} \nabla \times \mathbf{F} \cdot \hat{\mathbf{n}}\, dS$. By the mean value theorem for integrals this is πa^2 times the value of $\nabla \times \mathbf{F} \cdot \hat{\mathbf{n}}$ at some point of D_a. Divide by πa^2 and let $a \to 0$, using continuity. ∎

This justifies the paddle-wheel picture of Section 6.5: the component of curl along an axis is the circulation per unit area around that axis. It also gives a coordinate-

free definition of curl. That definition is how we compute curl in curvilinear coordinates in Section 7.7.

Example 7.18 (Faraday's law). The differential law $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$ and Stokes' theorem give, for a fixed loop C bounding S,

$$\oint_C \mathbf{E} \cdot d\mathbf{r} = -\frac{\mathrm{d}}{\mathrm{d}t} \iint_S \mathbf{B} \cdot d\mathbf{S}.$$

The voltage induced around a loop equals minus the rate of change of magnetic flux through it. Every electric generator and transformer runs on this identity.

7.5 The Divergence Theorem

Theorem 7.19 (Divergence theorem (Gauss)). *Let E be a bounded solid region whose boundary ∂E is a closed, piecewise smooth surface oriented outward, and let $\mathbf{F}$ be C^1 on an open set containing E. Then*

$$\oiint_{\partial E} \mathbf{F} \cdot d\mathbf{S} = \iiint_E \nabla \cdot \mathbf{F}\, dV.$$

Proof for simple solids. Write $\mathbf{F} = P\hat{\mathbf{i}} + Q\hat{\mathbf{j}} + R\hat{\mathbf{k}}$. It suffices to prove $\oiint_{\partial E} R\,\hat{\mathbf{k}} \cdot \hat{\mathbf{n}}\, dS = \iiint_E R_z\, dV$ together with the analogous statements for P and Q. Suppose $E = \{(x,y) \in D,\ u_1(x,y) \le z \le u_2(x,y)\}$. By the fundamental theorem of calculus,

$$\iiint_E R_z\, dV = \iint_D \big[R(x,y,u_2) - R(x,y,u_1)\big]\, dA.$$

On the boundary, ∂E consists of a top S_2 (the graph of u_2, outward normal pointing up), a bottom S_1 (the graph of u_1, outward normal pointing down), and possibly vertical sides, where $\hat{\mathbf{k}} \cdot \hat{\mathbf{n}} = 0$. By (7.1), $\iint_{S_2} R\hat{\mathbf{k}} \cdot d\mathbf{S} = \iint_D R(x,y,u_2)\, dA$, and the bottom contributes $-\iint_D R(x,y,u_1)\, dA$. The two sides agree. A solid that is simple in all three coordinate directions satisfies all three identities, and general solids are handled by subdivision. ∎

Example 7.20 (Flux through a cube). Find the flux of $\mathbf{F} = \langle x^3, y^3, z^3 \rangle$ outward through the surface of the cube $[0,1]^3$.

Solution. $\nabla \cdot \mathbf{F} = 3(x^2+y^2+z^2)$, so the flux is $\iiint_{[0,1]^3} 3(x^2+y^2+z^2)\, dV =$

$3 \cdot 3 \cdot \frac{1}{3} = 3$. Computing the same answer directly requires six surface integrals.

Example 7.21 (A flux through an open surface). Find the flux of $\mathbf{F} = \left\langle xy,\ y^2 + e^{xz^2},\ \sin(xy) \right\rangle$ out of the region E bounded by $z = 1 - x^2$, $z = 0$, $y = 0$, and $y + z = 2$.

Solution. Direct computation is hopeless, but $\nabla \cdot \mathbf{F} = y + 2y = 3y$. So the flux is

$$\int_{-1}^{1} \int_{0}^{1-x^2} \int_{0}^{2-z} 3y \, dy \, dz \, dx = \int_{-1}^{1} \int_{0}^{1-x^2} \tfrac{3}{2}(2-z)^2 \, dz \, dx = \frac{184}{35}.$$

Theorem 7.22 (Divergence as flux density). *If B_a is the ball of radius a about $\mathbf{x}$, then* $\nabla \cdot \mathbf{F}(\mathbf{x}) = \lim_{a \to 0} \dfrac{1}{\frac{4}{3}\pi a^3} \oiint_{\partial B_a} \mathbf{F} \cdot d\mathbf{S}$.

The proof is identical to that of Theorem 7.17. Divergence is outward flux per unit volume, independent of any coordinate system.

7.5.1 GAUSS'S LAW

Theorem 7.23 (Gauss's law for a point charge). *Let $\mathbf{F} = \mathbf{r}/r^3$. For any closed, piecewise smooth surface S bounding a solid region E, with $\mathbf{0} \notin S$,*

$$\oiint_S \mathbf{F} \cdot d\mathbf{S} = \begin{cases} 4\pi, & \mathbf{0} \in E, \\ 0, & \mathbf{0} \notin E. \end{cases}$$

Proof. If $\mathbf{0} \notin E$, $\mathbf{F}$ is C^1 on E and $\nabla \cdot \mathbf{F} = 0$ (Example 6.22), so the divergence theorem gives 0. If $\mathbf{0} \in E$, remove a small ball B_a about the origin. On the region $E \setminus B_a$, whose boundary is S together with the sphere ∂B_a oriented inward, the divergence theorem gives $\oiint_S - \oiint_{\partial B_a} = 0$, and Example 7.12 gives 4π for the sphere. ∎

The electric field of a point charge q is $\mathbf{E} = \frac{q}{4\pi\varepsilon_0}\mathbf{r}/r^3$, so its flux through any closed surface enclosing the charge is q/ε_0. By superposition, the flux of the field of any collection of charges through a closed surface is $Q_{\text{enc}}/\varepsilon_0$. Applying the divergence theorem to a charge distribution of density ρ and shrinking the

region gives the differential form,

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0},$$

which is the first of Maxwell's equations.

7.5.2 THE CONTINUITY EQUATION

Let $\rho(\mathbf{r}, t)$ be the density of a conserved quantity (mass, charge, or probability) and $\mathbf{J}$ its current density, so that $\mathbf{J} \cdot \hat{\mathbf{n}}$ is the amount crossing a surface per unit area per unit time. For any fixed region E, conservation says that the amount inside decreases exactly by what flows out:

$$\frac{\mathrm{d}}{\mathrm{d}t} \iiint_E \rho \, dV = -\oiint_{\partial E} \mathbf{J} \cdot d\mathbf{S} = -\iiint_E \nabla \cdot \mathbf{J} \, dV.$$

Since E is arbitrary, the integrands must agree.

Theorem 7.24 (Continuity equation). *A conserved density and its current satisfy* $\dfrac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0$.

LOOKING AHEAD. In quantum mechanics the conserved quantity is probability, with density $\rho = |\psi|^2$. In Chapter 13 we derive from the Schrödinger equation that $\partial_t |\psi|^2 + \nabla \cdot \mathbf{J} = 0$ with the *probability current* $\mathbf{J} = \frac{\hbar}{2mi}\left(\psi^* \nabla \psi - \psi \nabla \psi^*\right)$. Integrating over all space, this shows that the total probability $\iiint |\psi|^2 dV$ stays equal to 1 forever. The normalization imposed at one moment is preserved by the dynamics.

7.6 THE DIRAC DELTA FUNCTION

Theorem 7.23 contains a puzzle. The field $\mathbf{r}/r^3$ has zero divergence everywhere except at the origin, where it is undefined, yet its flux out of any region containing the origin is 4π. The divergence theorem can be rescued only if all of the divergence is concentrated at one point. Physicists describe such a concentration with the *Dirac delta*.

Definition 7.25 (Dirac delta, informal). The *delta function* $\delta(x)$ is character-

ized by the *sifting property*

$$\int_{-\infty}^{\infty} f(x)\,\delta(x-a)\,dx = f(a) \qquad \text{for every continuous } f.$$

In three dimensions $\delta^3(\mathbf{r}) = \delta(x)\delta(y)\delta(z)$, so $\iiint f(\mathbf{r})\delta^3(\mathbf{r}-\mathbf{a})\,dV = f(\mathbf{a})$.

No ordinary function has this property. It would have to vanish for $x \neq a$ and yet integrate to 1. Rigorously, δ is a *distribution*: a rule that assigns to each smooth test function f the number $f(a)$. It is also the limit of a sequence of ordinary functions, such as normalized Gaussians of shrinking width,

$$\delta(x) = \lim_{\sigma \to 0^+} \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}, \tag{7.2}$$

in the sense that $\int f(x)\,g_\sigma(x)\,dx \to f(0)$ for every continuous bounded f. All manipulations of δ in this book can be justified this way.

Theorem 7.26 (Properties of δ). *(a)* $\delta(cx) = \delta(x)/|c|$ *for* $c \neq 0$*; in particular* $\delta(-x) = \delta(x)$.

(b) $f(x)\delta(x-a) = f(a)\delta(x-a)$.

(c) $\delta(g(x)) = \sum_i \delta(x-x_i)/|g'(x_i)|$, *summed over the simple zeros* x_i *of* g.

(d) $\int_{-\infty}^{\infty} f(x)\delta'(x)\,dx = -f'(0)$, *by formal integration by parts.*

(e) $\delta(x) = \mathrm{d}/\mathrm{d}x\ H(x)$, *where* H *is the Heaviside step function* (0 *for* $x<0$, 1 *for* $x>0$).

Proof. (a) Substitute $u = cx$ in $\int f(x)\delta(cx)\,dx$; a negative c reverses the limits, which produces the absolute value. (c) Near each simple zero, $g(x) \approx g'(x_i)(x-x_i)$; apply (a). (e) $\int f(x)H'(x)\,dx = -\int f'(x)H(x)\,dx = -\int_0^\infty f' = f(0)$ for f vanishing at infinity. ∎

Theorem 7.27 (The delta function of a point charge). *In the sense of distributions,*

$$\boldsymbol{\nabla}\cdot\left(\frac{\hat{\mathbf{r}}}{r^2}\right) = 4\pi\,\delta^3(\mathbf{r}), \qquad \nabla^2\left(\frac{1}{r}\right) = -4\pi\,\delta^3(\mathbf{r}).$$

The first identity says exactly that the divergence theorem holds for $\mathbf{r}/r^3$ on every region: the flux 4π comes from $\iiint 4\pi\delta^3\,dV = 4\pi$ when the origin is enclosed. The second follows because $\nabla(1/r) = -\hat{\mathbf{r}}/r^2$. In electrostatics, it says that the

potential $\Phi = \frac{q}{4\pi\varepsilon_0 r}$ of a point charge solves Poisson's equation $\nabla^2\Phi = -\rho/\varepsilon_0$ with $\rho = q\,\delta^3(\mathbf{r})$.

LOOKING AHEAD. The delta function is indispensable in quantum mechanics. The position eigenstates $|x_0\rangle$ have "wavefunctions" $\delta(x - x_0)$. The orthonormality of continuous bases is written $\langle x|x'\rangle = \delta(x - x')$. The simplest bound-state problem is the attractive delta-function potential $V(x) = -\alpha\delta(x)$ (Chapter 14). And $\delta(x) = \frac{1}{2\pi}\int e^{ikx}dk$, the Fourier representation of Chapter 11, is the statement that plane waves form a complete set.

7.7 Curvilinear Coordinates

The Schrödinger equation for an atom involves the Laplacian, and the atom's symmetry is spherical. We therefore need ∇^2 in spherical coordinates. We derive the gradient, divergence, curl, and Laplacian in any *orthogonal* coordinate system at once.

7.7.1 Scale Factors

Let (q_1, q_2, q_3) be coordinates with $\mathbf{r} = \mathbf{r}(q_1, q_2, q_3)$, such that the coordinate tangent vectors $\partial\mathbf{r}/\partial q_i$ are mutually orthogonal. Write

$$\frac{\partial \mathbf{r}}{\partial q_i} = h_i\,\hat{\mathbf{e}}_i, \qquad h_i = \left|\frac{\partial \mathbf{r}}{\partial q_i}\right|,$$

where $\hat{\mathbf{e}}_i$ are orthonormal and the h_i are the *scale factors*. Then a displacement is $d\mathbf{r} = \sum_i h_i\,dq_i\,\hat{\mathbf{e}}_i$, so the arc-length and volume elements are

$$ds^2 = h_1^2 dq_1^2 + h_2^2 dq_2^2 + h_3^2 dq_3^2, \qquad dV = h_1 h_2 h_3\,dq_1\,dq_2\,dq_3.$$

Coordinates	(q_1, q_2, q_3)	(h_1, h_2, h_3)	dV
Cartesian	(x, y, z)	$(1, 1, 1)$	$dx\,dy\,dz$
Cylindrical	(ρ, ϕ, z)	$(1, \rho, 1)$	$\rho\,d\rho\,d\phi\,dz$
Spherical	(r, θ, ϕ)	$(1, r, r\sin\theta)$	$r^2\sin\theta\,dr\,d\theta\,d\phi$

Theorem 7.28 (Differential operators in orthogonal coordinates).

$$\nabla f = \sum_i \frac{1}{h_i}\frac{\partial f}{\partial q_i}\,\hat{\mathbf{e}}_i,$$
$$\boldsymbol{\nabla}\cdot\mathbf{F} = \frac{1}{h_1h_2h_3}\Big[\frac{\partial}{\partial q_1}(h_2h_3F_1) + \frac{\partial}{\partial q_2}(h_3h_1F_2) + \frac{\partial}{\partial q_3}(h_1h_2F_3)\Big],$$
$$\nabla^2 f = \frac{1}{h_1h_2h_3}\Big[\frac{\partial}{\partial q_1}\Big(\frac{h_2h_3}{h_1}\frac{\partial f}{\partial q_1}\Big) + \frac{\partial}{\partial q_2}\Big(\frac{h_3h_1}{h_2}\frac{\partial f}{\partial q_2}\Big) + \frac{\partial}{\partial q_3}\Big(\frac{h_1h_2}{h_3}\frac{\partial f}{\partial q_3}\Big)\Big],$$

and $\boldsymbol{\nabla}\times\mathbf{F}$ *is* $\dfrac{1}{h_1h_2h_3}$ *times the determinant with rows* $(h_1\hat{\mathbf{e}}_1,\ h_2\hat{\mathbf{e}}_2,\ h_3\hat{\mathbf{e}}_3)$, $(\partial_{q_1},\ \partial_{q_2},\ \partial_{q_3})$, *and* $(h_1F_1,\ h_2F_2,\ h_3F_3)$.

Proof. Gradient. The directional derivative along $\hat{\mathbf{e}}_i$ is the rate of change per unit distance. Moving distance ds along $\hat{\mathbf{e}}_i$ changes q_i by ds/h_i, so the $\hat{\mathbf{e}}_i$-component of ∇f is $h_i^{-1}\partial f/\partial q_i$.

Divergence. Use Theorem 7.22 with a small coordinate box with edges $h_i\,dq_i$. The two faces perpendicular to $\hat{\mathbf{e}}_1$ have area $h_2h_3\,dq_2\,dq_3$, and the net outward flux through them is $\big[(F_1h_2h_3)(q_1+dq_1) - (F_1h_2h_3)(q_1)\big]dq_2\,dq_3 \approx \partial/\partial q_1\ (F_1h_2h_3)\,dq_1dq_2dq_3$. Add the three pairs of faces and divide by the volume $h_1h_2h_3\,dq_1dq_2dq_3$.

Laplacian. $\nabla^2 f = \boldsymbol{\nabla}\cdot(\nabla f)$; substitute the gradient into the divergence.

Curl. Apply Theorem 7.17 to a small coordinate rectangle, by the same flux-counting argument. ∎

Corollary 7.29 (Spherical coordinates). *With* $(h_r, h_\theta, h_\phi) = (1, r, r\sin\theta)$,

$$\nabla f = \frac{\partial f}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial f}{\partial\theta}\hat{\boldsymbol{\theta}} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial\phi}\hat{\boldsymbol{\phi}},$$
$$\boldsymbol{\nabla}\cdot\mathbf{F} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2F_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\,F_\theta) + \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial\phi},$$
$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\Big(r^2\frac{\partial f}{\partial r}\Big) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial f}{\partial\theta}\Big) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2}.$$

In cylindrical coordinates, $\nabla^2 f = \frac{1}{\rho}\,\partial/\partial\rho\,(\rho\ \partial f/\partial\rho\,) + \frac{1}{\rho^2}\,\partial^2 f/\partial\phi^2 + \partial^2 f/\partial z^2$, *which reproduces Example 3.20.*

Example 7.30 (Checks). (a) For $f = 1/r$: $\nabla^2 f = \frac{1}{r^2}\,\mathrm{d}/\mathrm{d}r\,\big(r^2\cdot(-r^{-2})\big) = 0$ for $r \neq 0$, consistent with Theorem 7.27. (b) For $\mathbf{F} = \hat{\mathbf{r}}/r^2$: $\boldsymbol{\nabla}\cdot\mathbf{F} = \frac{1}{r^2}\,\partial/\partial r\,(1) = 0$, as in Example 6.22. (c) For $f = r^2 = x^2+y^2+z^2$: $\nabla^2 f = \frac{1}{r^2}\,\mathrm{d}/\mathrm{d}r\,\big(2r^3\big) = 6$, which agrees with the Cartesian computation $2+2+2$.

Example 7.31 (Spherically symmetric solutions). Find all radial functions $f(r)$ with $\nabla^2 f = -k^2 f$ in $\mathbb{R}^3$, for $k > 0$.

Solution. For radial f, $\nabla^2 f = f'' + \frac{2}{r}f' = \frac{1}{r}(rf)''$. So with $u = rf$ the equation becomes $u'' = -k^2 u$, whose solutions are $u = A\sin kr + B\cos kr$. Hence $f = \big(A\sin kr + B\cos kr\big)/r$. The solution $\sin(kr)/r$ is finite at the origin, while $\cos(kr)/r$ is singular there. This substitution $u = rf$ reappears for the radial equation of the hydrogen atom.

LOOKING AHEAD. The angular part of the spherical Laplacian,

$$\Lambda = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial}{\partial\theta}\Big) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2},$$

is, up to the factor $-\hbar^2$, the operator $\hat{L}^2$ for the square of orbital angular momentum. Its eigenfunctions, the *spherical harmonics* $Y_\ell^m(\theta,\phi)$, satisfy $\Lambda Y_\ell^m = -\ell(\ell+1)Y_\ell^m$. They give the shapes of the s, p, d, f orbitals of chemistry (Chapters 16 and 17).

7.8 The Unity of the Integral Theorems

Every theorem of this kind says that integrating a derivative over a region equals integrating the original function over the region's boundary.

Theorem	Region (dimension)	Statement
Fundamental theorem	interval $[a,b]$ (1)	$\int_a^b f'(x)\,dx = f(b) - f(a)$
Line integrals	curve C (1)	$\int_C \nabla f\cdot d\mathbf{r} = f(B) - f(A)$
Green	plane region D (2)	$\iint_D (Q_x - P_y)\,dA = \oint_{\partial D} P\,dx + Q\,dy$
Stokes	surface S (2)	$\iint_S \boldsymbol{\nabla}\times\mathbf{F}\cdot d\mathbf{S} = \oint_{\partial S}\mathbf{F}\cdot d\mathbf{r}$
Gauss	solid E (3)	$\iiint_E \boldsymbol{\nabla}\cdot\mathbf{F}\,dV = \oiint_{\partial E}\mathbf{F}\cdot d\mathbf{S}$

In the language of differential forms these are all one theorem, the *generalized*

Stokes theorem $\int_M d\omega = \int_{\partial M} \omega$. In it, d unifies gradient, curl, and divergence, and the identities $\boldsymbol{\nabla} \times \nabla = \mathbf{0}$ and $\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \times = 0$ both become $d(d\omega) = 0$. The boundary of a boundary is empty.

7.9 Worked Problems

The problems below compute surface areas and surface integrals, verify Stokes' theorem and the divergence theorem on explicit examples, and then use the theorems as tools: for the mean value property of harmonic functions, for the delta function in curvilinear coordinates, and for an electrostatic potential.

Problem 7.1 Area of a paraboloid cap

Find the area of the part of the paraboloid $z = x^2 + y^2$ that lies below the plane $z = 4$.

Solution. For a graph $z = g(x, y)$, $dS = \sqrt{1 + g_x^2 + g_y^2}\, dA = \sqrt{1 + 4x^2 + 4y^2}\, dA$. The surface lies over the disk $x^2 + y^2 \le 4$. In polar coordinates, with $w = 1 + 4r^2$,

$$A = \int_0^{2\pi} \int_0^2 \sqrt{1 + 4r^2}\, r\, dr\, d\theta = 2\pi \cdot \frac{1}{8} \cdot \frac{2}{3} \Big[w^{3/2} \Big]_1^{17} = \frac{\pi}{6} \left(17\sqrt{17} - 1 \right) \approx 36.18.$$

Sanity check: the area must exceed that of the disk it projects onto, $4\pi \approx 12.57$, and the cap is steep near its rim (slope 4 at $r = 2$), which explains why it is nearly three times as large.

Problem 7.2 Moment of inertia of a spherical shell

A thin uniform spherical shell of radius a has mass M. Find its moment of inertia about a diameter, first by direct integration and then by a symmetry argument.

Solution. The surface density is $\sigma = M/(4\pi a^2)$, and the distance from the z-axis is $a \sin\theta$. With $dS = a^2 \sin\theta\, d\theta\, d\phi$,

$$I_z = \sigma \int_0^{2\pi} \int_0^{\pi} (a \sin\theta)^2 a^2 \sin\theta\, d\theta\, d\phi = \sigma \cdot 2\pi a^4 \cdot \frac{4}{3} = \frac{M}{4\pi a^2} \cdot \frac{8\pi a^4}{3} = \frac{2}{3} M a^2.$$

By symmetry. $I_x + I_y + I_z = \iint \sigma \left[(y^2 + z^2) + (z^2 + x^2) + (x^2 + y^2) \right] dS = 2 \iint \sigma r^2\, dS = 2Ma^2$, since $r = a$ on the shell. The three moments are equal by symmetry, so each is $\frac{2}{3} Ma^2$. No integral was needed.

The same argument for the solid ball gives $I_x + I_y + I_z = 2 \iiint \frac{M}{V} r^2\, dV = 2M \cdot \frac{3}{5} a^2$, so each moment is $\frac{2}{5} Ma^2$, as in Example 5.21. The shell's moment is larger because all its mass sits at the maximum radius.

Problem 7.3 Stokes' theorem on a triangle

Let $\mathbf{F} = \langle z, x, y \rangle$, and let S be the triangle with vertices $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, oriented with upward normal. Verify Stokes' theorem by computing both sides.

Solution. *Surface side.* $\nabla \times \mathbf{F} = \langle \partial_y y - \partial_z x,\ \partial_z z - \partial_x y,\ \partial_x x - \partial_y z \rangle = \langle 1,1,1 \rangle$. The triangle lies in the plane $x + y + z = 1$ with unit normal $\hat{\mathbf{n}} = \frac{1}{\sqrt{3}} \langle 1,1,1 \rangle$, and its area is $\frac{1}{2} |\, \langle -1,1,0 \rangle \times \langle -1,0,1 \rangle \,| = \frac{1}{2} |\, \langle 1,1,1 \rangle \,| = \frac{\sqrt{3}}{2}$. Since $\nabla \times \mathbf{F}$ is constant,

$$\iint_S \nabla \times \mathbf{F} \cdot \hat{\mathbf{n}}\, dS = \frac{3}{\sqrt{3}} \cdot \frac{\sqrt{3}}{2} = \frac{3}{2}.$$

Boundary side. The upward normal induces the order $A = (1,0,0) \to B = (0,1,0) \to C = (0,0,1) \to A$. On each edge $\mathbf{r} = P + t(Q - P)$, $0 \le t \le 1$:

- $A \to B$: $\mathbf{r} = \langle 1 - t, t, 0 \rangle$, $d\mathbf{r} = \langle -1,1,0 \rangle\, dt$, $\mathbf{F} = \langle 0, 1 - t, t \rangle$, integrand $1 - t$, integral $\frac{1}{2}$.
- $B \to C$: $\mathbf{r} = \langle 0, 1 - t, t \rangle$, $d\mathbf{r} = \langle 0, -1, 1 \rangle\, dt$, $\mathbf{F} = \langle t, 0, 1 - t \rangle$, integrand $1 - t$, integral $\frac{1}{2}$.
- $C \to A$: $\mathbf{r} = \langle t, 0, 1 - t \rangle$, $d\mathbf{r} = \langle 1, 0, -1 \rangle\, dt$, $\mathbf{F} = \langle 1 - t, t, 0 \rangle$, integrand $1 - t$, integral $\frac{1}{2}$.

The total is $\frac{3}{2}$, in agreement. The three edges contribute equally because both $\mathbf{F}$ and the triangle are invariant under the cyclic permutation $(x, y, z) \mapsto (y, z, x)$.

Problem 7.4 The divergence theorem on a cylinder

Let E be the solid cylinder $x^2 + y^2 \le 1$, $0 \le z \le 2$, and $\mathbf{F} = \langle x, y, z \rangle$. Compute the outward flux of $\mathbf{F}$ through ∂E by integrating over each of its three faces, and compare with $\iiint_E \nabla \cdot \mathbf{F}\, dV$.

Solution. *Side.* The outward unit normal is $\langle x, y, 0 \rangle$, so $\mathbf{F} \cdot \hat{\mathbf{n}} = x^2 + y^2 = 1$, and the flux equals the side's area, $2\pi \cdot 1 \cdot 2 = 4\pi$.

Top ($z = 2$, $\hat{\mathbf{n}} = \hat{\mathbf{k}}$). $\mathbf{F} \cdot \hat{\mathbf{n}} = z = 2$, and the flux is $2 \cdot \pi = 2\pi$.

Bottom ($z = 0$, $\hat{\mathbf{n}} = -\hat{\mathbf{k}}$). $\mathbf{F} \cdot \hat{\mathbf{n}} = -z = 0$, so the flux is 0.

The total is 6π. On the other hand, $\nabla \cdot \mathbf{F} = 3$, and $\iiint_E 3\, dV = 3 \cdot 2\pi = 6\pi$. The two computations agree. Note that the flux per unit area varies by face: it is 1 on the side, 2 on the top, and 0 on the bottom, which passes through the origin, where $\mathbf{F}$ vanishes.

Problem 7.5 Flux out of a cylinder

Find the outward flux of $\mathbf{F} = \langle xy^2,\ yz^2,\ zx^2 \rangle$ through the boundary of the solid cylinder $x^2 + y^2 \le 1$, $0 \le z \le 1$.

Solution. Computing the flux through the side, top, and bottom directly would mean three surface integrals with awkward integrands. The divergence is simpler: $\nabla \cdot \mathbf{F} = y^2 + z^2 + x^2$. By the divergence theorem, in cylindrical coordinates,

$$\oiint \mathbf{F} \cdot d\mathbf{S} = \iiint (\rho^2 + z^2)\, dV = 2\pi \int_0^1 \int_0^1 (\rho^2 + z^2)\, \rho \, d\rho \, dz = 2\pi\left(\frac{1}{4} + \frac{1}{2}\cdot\frac{1}{3}\right) = \frac{5\pi}{6}.$$

As a partial check, the bottom face ($z = 0$, $\hat{\mathbf{n}} = -\hat{\mathbf{k}}$) contributes $-\iint zx^2\, dA = 0$, and the top face ($z = 1$) contributes $\iint x^2\, dA = \frac{\pi}{4}$. So the side must contribute $\frac{5\pi}{6} - \frac{\pi}{4} = \frac{7\pi}{12}$. On the side, $\hat{\mathbf{n}} = \langle x, y, 0 \rangle$ and $\mathbf{F} \cdot \hat{\mathbf{n}} = x^2y^2 + y^2z^2$. With $x = \cos\phi$, $y = \sin\phi$, $dS = d\phi\, dz$: $\int_0^1 \int_0^{2\pi} (\cos^2\phi \sin^2\phi + z^2 \sin^2\phi)\, d\phi\, dz = \frac{\pi}{4} + \frac{\pi}{3} = \frac{7\pi}{12}$. The theorem and the direct computation agree.

Problem 7.6 The mean value property of harmonic functions

Let u be harmonic ($\nabla^2 u = 0$) on a region containing the closed ball of radius R about $\mathbf{a}$. Prove that the average of u over the sphere S_R of radius R about $\mathbf{a}$ equals $u(\mathbf{a})$. Deduce that a nonconstant harmonic function has no local maximum or minimum.

Solution. Let $m(r) = \dfrac{1}{4\pi r^2} \oiint_{S_r} u\, dS$ be the average over the sphere of radius r. Parametrize S_r by the unit sphere: $\mathbf{x} = \mathbf{a} + r\hat{\boldsymbol{\ell}}$ with $dS = r^2\, d\Omega$, where $d\Omega = \sin\theta\, d\theta\, d\phi$. Then $m(r) = \frac{1}{4\pi} \oiint u(\mathbf{a} + r\hat{\boldsymbol{\ell}})\, d\Omega$, an integral over a fixed domain, and differentiation under the integral sign gives

$$m'(r) = \frac{1}{4\pi} \oiint \nabla u(\mathbf{a} + r\hat{\boldsymbol{\ell}}) \cdot \hat{\boldsymbol{\ell}}\, d\Omega = \frac{1}{4\pi r^2} \oiint_{S_r} \nabla u \cdot \hat{\mathbf{n}}\, dS = \frac{1}{4\pi r^2} \iiint_{B_r} \nabla^2 u\, dV = 0,$$

by the divergence theorem, since $\hat{\boldsymbol{\ell}}$ is the outward normal $\hat{\mathbf{n}}$. So m is constant on $(0, R]$. As $r \to 0$, continuity of u gives $m(r) \to u(\mathbf{a})$. Hence $m(R) = u(\mathbf{a})$.

For the second statement, suppose u has a local maximum at $\mathbf{a}$. Then $u \le u(\mathbf{a})$ on small spheres about $\mathbf{a}$, and the average equals $u(\mathbf{a})$, so $u = u(\mathbf{a})$ on each of these spheres (a continuous function that is $\le c$ with average c must equal c). Thus u is constant near $\mathbf{a}$. A connectedness argument extends this to the whole region. This is the *maximum principle*. An electrostatic potential in a charge-free region has no local extrema, so a charge cannot be held in stable equilibrium by electrostatic forces alone (*Earnshaw's theorem*).

Problem 7.7 The delta function in spherical coordinates

The three-dimensional delta function is defined by the property that $\iiint f\,\delta^3(\mathbf{r}-\mathbf{r}_0)\,dV = f(\mathbf{r}_0)$ for every continuous f. If $\mathbf{r}_0$ has spherical coordinates (r_0,θ_0,ϕ_0) with $r_0>0$ and $0<\theta_0<\pi$, show that

$$\delta^3(\mathbf{r}-\mathbf{r}_0)=\frac{\delta(r-r_0)\,\delta(\theta-\theta_0)\,\delta(\phi-\phi_0)}{r^2\sin\theta}.$$

Use it to write the charge density of a uniformly charged spherical shell of radius a and total charge Q.

Solution. Integrate the proposed expression against f, using $dV = r^2\sin\theta\,dr\,d\theta\,d\phi$:

$$\iiint f\,\frac{\delta(r-r_0)\delta(\theta-\theta_0)\delta(\phi-\phi_0)}{r^2\sin\theta}\,r^2\sin\theta\,dr\,d\theta\,d\phi$$
$$=\int\!\!\int\!\!\int f\,\delta(r-r_0)\delta(\theta-\theta_0)\delta(\phi-\phi_0)\,dr\,d\theta\,d\phi = f(r_0,\theta_0,\phi_0).$$

The Jacobian in the volume element exactly cancels the denominator. The rule: a product of one-dimensional deltas in curvilinear coordinates must be divided by the Jacobian $h_1h_2h_3$. (On the z-axis, where $\sin\theta_0=0$, the angle ϕ_0 is undefined and the formula must be modified, which is why that case was excluded.)

For the shell, the charge is spread uniformly in angle at $r=a$, so the density has the form $\rho = c\,\delta(r-a)$. The total charge fixes c: $\iiint c\,\delta(r-a)\,r^2\sin\theta\,dr\,d\theta\,d\phi = 4\pi a^2 c = Q$. Hence

$$\rho(\mathbf{r})=\frac{Q}{4\pi a^2}\,\delta(r-a),$$

which is the surface charge density $\sigma = Q/(4\pi a^2)$ times a delta function in the normal direction. Densities of this kind, concentrated on surfaces or at points, are how the delta function enters electrostatics and, through the potentials they create, the quantum mechanics of Chapter 14.

Problem 7.8 The coaxial capacitor

The region between two long coaxial conducting cylinders, of radii $a<b$, is free of charge. The inner cylinder is held at potential V_0 and the outer at 0. Find the potential Φ and the electric field between them, assuming Φ depends only on the distance ρ from the axis.

Solution. In a charge-free region $\nabla^2\Phi = 0$. With $\Phi=\Phi(\rho)$, the cylindrical Laplacian of Corollary 7.29 reduces to

$$\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}\left(\rho\frac{\mathrm{d}\Phi}{\mathrm{d}\rho}\right)=0 \quad\Longrightarrow\quad \rho\frac{\mathrm{d}\Phi}{\mathrm{d}\rho}=A \quad\Longrightarrow\quad \Phi = A\ln\rho + B.$$

The boundary conditions $\Phi(a) = V_0$ and $\Phi(b) = 0$ give

$$\Phi(\rho) = V_0 \frac{\ln(b/\rho)}{\ln(b/a)}, \qquad \mathbf{E} = -\nabla\Phi = \frac{V_0}{\rho \ln(b/a)}\,\hat{\boldsymbol{\rho}}.$$

Solutions of Laplace's equation with given boundary values are unique (an exercise of this chapter proves it), so the assumption of cylindrical symmetry cost nothing. The field falls off as $1/\rho$, the field of a line charge. By Gauss's law, the charge per unit length on the inner cylinder is $\lambda = 2\pi\varepsilon_0 \rho E = \frac{2\pi\varepsilon_0 V_0}{\ln(b/a)}$, so the capacitance per unit length is $\frac{2\pi\varepsilon_0}{\ln(b/a)}$, the standard result for a coaxial cable. Compare the spherically symmetric solutions earlier in this chapter, where the radial solutions are $A/r + B$: the logarithm is the two-dimensional replacement for $1/r$, as found in the previous chapter's problems.

EXERCISES

A. COMPUTATION

7.1 Find a parametrization of the part of the plane $x + 2y + 3z = 6$ in the first octant, and of the part of the cone $z = \sqrt{x^2 + y^2}$ with $1 \le z \le 3$.

7.2 Find the tangent plane to $\mathbf{r}(u, v) = \langle u + v,\ 3u^2,\ u - v \rangle$ at $(2, 3, 0)$.

7.3 Find the area of the part of the plane $3x + 2y + z = 6$ in the first octant.

7.4 Find the area of the part of $z = x^2 + y^2$ under the plane $z = 9$.

7.5 Find the area of the part of the sphere $x^2 + y^2 + z^2 = a^2$ inside the cylinder $x^2 + y^2 = ax$ (Viviani's problem).

7.6 Find the area of the helicoid $\mathbf{r}(u, v) = \langle u\cos v,\ u\sin v,\ v \rangle$ for $0 \le u \le 1$, $0 \le v \le \pi$.

7.7 Evaluate $\iint_S z\, dS$ over the hemisphere $x^2 + y^2 + z^2 = 4$, $z \ge 0$.

7.8 Evaluate $\iint_S (x^2 + y^2)\, dS$ over the sphere of radius a, and deduce the moment of inertia of a thin spherical shell.

7.9 Evaluate $\iint_S y\, dS$, where S is the part of the plane $3x + 2y + z = 6$ in the first octant.

7.10 Find the flux of $\mathbf{F} = \langle x, y, z^4 \rangle$ upward through the cone $z = \sqrt{x^2 + y^2}$ beneath $z = 1$.

7.11 Find the flux of $\mathbf{F} = \langle 0, 0, z \rangle$ outward through the sphere of radius a, directly and by the divergence theorem.

7.12 Find the flux of $\mathbf{F} = \langle xz,\ -2y,\ 3x \rangle$ outward through the sphere $x^2 + y^2 + z^2 = 4$.

7.13 Use Stokes' theorem to evaluate $\iint_S \nabla \times \mathbf{F} \cdot d\mathbf{S}$ for

$$\mathbf{F} = \langle 2y \cos z,\ e^x \sin z,\ xe^y \rangle,$$

where S is the hemisphere $x^2 + y^2 + z^2 = 9$, $z \geq 0$, oriented upward.

7.14 Use Stokes' theorem to evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$ for $\mathbf{F} = \langle x + y^2,\ y + z^2,\ z + x^2 \rangle$, where C is the triangle with vertices $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, oriented counterclockwise from above.

7.15 Use Stokes' theorem to evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$ for $\mathbf{F} = \langle yz,\ 2xz,\ e^{xy} \rangle$, where C is the circle $x^2 + y^2 = 16$, $z = 5$, counterclockwise from above.

7.16 Verify Stokes' theorem for $\mathbf{F} = \langle y,\ z,\ x \rangle$ on the hemisphere $x^2 + y^2 + z^2 = 1$, $z \geq 0$, oriented upward.

7.17 Use the divergence theorem to find the flux of $\mathbf{F} = \langle x^2 y,\ -x^2 z,\ z^2 y \rangle$ out of the box $[0,3] \times [0,2] \times [0,1]$.

7.18 Use the divergence theorem to find the flux of $\mathbf{F} = \langle 3xy^2,\ xe^z,\ z^3 \rangle$ out of the solid bounded by $y^2 + z^2 = 1$, $x = -1$, $x = 2$.

7.19 Use the divergence theorem to find the flux of $\mathbf{F} = \langle x^3,\ y^3,\ z^3 \rangle$ out of the sphere of radius a.

7.20 Find the flux of $\mathbf{F} = \langle z^2 x,\ \frac{1}{3}y^3 + \tan z,\ x^2 z + y^2 \rangle$ upward through the top half of the sphere $x^2 + y^2 + z^2 = 1$. (Close the surface with the unit disk.)

7.21 A charge distribution has density $\rho = \rho_0(1 - r/R)$ for $r \leq R$ and 0 outside. Use Gauss's law to find the electric field everywhere.

7.22 Evaluate $\int_{-\infty}^{\infty} (x^3 + 2x)\, \delta(x - 2)\, dx$, $\int_{-\infty}^{\infty} \cos x\, \delta(2x - \pi)\, dx$, and $\int_{-\infty}^{\infty} e^x \delta(x^2 - 1)\, dx$.

7.23 Evaluate $\iiint_{\mathbb{R}^3} (r^2 + \mathbf{r} \cdot \mathbf{a})\, \delta^3(\mathbf{r} - \mathbf{b})\, dV$ and $\iiint_{r<1} e^{-r}\, \nabla \cdot (\hat{\mathbf{r}}/r^2)\, dV$.

7.24 Compute $\nabla^2 f$ for $f = r^2 \cos\theta$ and for $f = r^{-2} \cos\theta$ using Corollary 7.29. Which is harmonic?

7.25 Compute $\nabla \cdot \mathbf{F}$ for $\mathbf{F} = r^n \hat{\mathbf{r}}$ in spherical coordinates. For which n is it divergence-free?

7.26 Compute ∇^2 of $\psi = e^{-r/a}$ and find the constant E for which $-\tfrac{1}{2}\nabla^2\psi - \tfrac{1}{r}\psi = E\psi$ when $a = 1$. (This is the hydrogen ground state in atomic units.)

7.27 Find the scale factors for parabolic coordinates $x = uv\cos\phi$, $y = uv\sin\phi$, $z = \tfrac{1}{2}(u^2 - v^2)$, verify they are orthogonal, and write the volume element.

B. THEORY AND PROOF

7.28 Prove that the flux of $\boldsymbol{\nabla} \times \mathbf{F}$ through any closed surface is zero, (a) using Stokes' theorem and (b) using the divergence theorem.

7.29 Prove that $\oiint_S \hat{\mathbf{n}}\, dS = \mathbf{0}$ for every closed surface S. (Apply the divergence theorem to $\mathbf{c}\cdot\hat{\mathbf{n}}$ for constant $\mathbf{c}$.)

7.30 Prove that the volume of a solid E is $\frac{1}{3}\oiint_{\partial E} \mathbf{r}\cdot d\mathbf{S}$, and use it to find the volume of a cone.

7.31 Prove Green's first identity in three dimensions: $\iiint_E f\nabla^2 g\, dV = \oiint_{\partial E} f\nabla g\cdot d\mathbf{S} - \iiint_E \nabla f\cdot\nabla g\, dV$.

7.32 (Archimedes' principle.) A body is submerged in a fluid of density ρ, where the pressure is $p = \rho g(h - z)$. Show that the net force $-\oiint p\,\hat{\mathbf{n}}\, dS$ equals $\rho g V\hat{\mathbf{k}}$.

7.33 Prove property (c) of Theorem 7.26 and evaluate $\delta(x^2 - a^2)$.

7.34 Derive the cylindrical-coordinate formulas for ∇f and $\boldsymbol{\nabla}\cdot\mathbf{F}$ from Theorem 7.28.

7.35 Show that $\mathbf{E} = -\nabla\Phi$ with $\nabla^2\Phi = -\rho/\varepsilon_0$ is consistent with Gauss's law, and that $\Phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\iiint \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}dV'$ solves Poisson's equation. (Use Theorem 7.27.)

7.36 Derive the continuity equation for the diffusion current $\mathbf{J} = -k\nabla u$ and show that it gives the heat equation $u_t = k\nabla^2 u$.

C. CHALLENGE

7.37 Show that the Möbius strip

$$\mathbf{r}(u,v) = \left\langle \left(2 + v\cos\tfrac{u}{2}\right)\cos u,\ \left(2 + v\cos\tfrac{u}{2}\right)\sin u,\ v\sin\tfrac{u}{2}\right\rangle,$$

$0 \le u \le 2\pi$, $-\tfrac{1}{2} \le v \le \tfrac{1}{2}$, is not orientable by computing $\mathbf{r}_u \times \mathbf{r}_v$ at $(0,0)$ and at $(2\pi,0)$.

7.38 Use Theorem 7.27 to show that $\nabla^2\left(\dfrac{e^{-\mu r}}{r}\right) = \mu^2\dfrac{e^{-\mu r}}{r} - 4\pi\delta^3(\mathbf{r})$. This is the Yukawa potential of the strong nuclear force, and $1/\mu$ is its range.

7.39 Derive the spherical Laplacian of Corollary 7.29 directly from the chain rule, without using scale factors, for a function $f(r)$ of r alone; then for $f(r,\theta)$.

7.40 A solid sphere of radius R carries uniform charge density. Find its electrostatic energy $U = \frac{\varepsilon_0}{2}\iiint_{\mathbb{R}^3}|\mathbf{E}|^2 dV$ in terms of its total charge Q. If an electron were such a sphere with $U = m_e c^2$, how large would it be? (This "classical electron radius" is about 2.8×10^{-15} m up to a numerical factor.)



PART II

The Mathematics of Quantum Theory

Complex analysis, linear algebra and Hilbert space, differential equations, and Fourier analysis: the language quantum mechanics is written in.

CHAPTER 8

Complex Numbers and Complex Analysis

LEARNING OBJECTIVES

- Perform complex arithmetic in rectangular and polar form, and prove Euler's formula from power series.
- Find all nth roots of a complex number and use complex exponentials to derive trigonometric identities.
- Define the complex exponential, logarithm, and powers, including branch cuts.
- Define analyticity, derive the Cauchy–Riemann equations, and show that analytic functions are harmonic.
- Evaluate contour integrals, and state and prove Cauchy's theorem and the Cauchy integral formula.
- Expand functions in Laurent series, classify singularities, and compute residues.
- Apply the residue theorem to evaluate real definite integrals.

Quantum mechanics is written in complex numbers. The Schrödinger equation contains $i = \sqrt{-1}$ explicitly, and wavefunctions are complex-valued. That is not a computational convenience that could be removed. The relative *phase* of two complex amplitudes is physically measurable through interference, and it has no real-number analog. This chapter develops complex numbers from the ground up and then proves the central theorems of complex analysis. Those theorems also provide one of the most powerful methods known for evaluating real integrals.

8.1 Complex Numbers

Definition 8.1 (Complex numbers). A *complex number* is an expression $z = x + iy$ with $x, y \in \mathbb{R}$, where $i^2 = -1$. Its *real part* is $\operatorname{Re} z = x$ and its *imaginary part* is $\operatorname{Im} z = y$. Addition and multiplication follow the rules of algebra with $i^2 = -1$:

$$(a + ib) + (c + id) = (a + c) + i(b + d),$$
$$(a + ib)(c + id) = (ac - bd) + i(ad + bc).$$

The set of complex numbers is $\mathbb{C}$. The *complex conjugate* of $z = x + iy$ is $\bar{z} = z^* = x - iy$, and its *modulus* is $|z| = \sqrt{x^2 + y^2}$.

Physicists usually write z^* for the conjugate, and we do so from here on. The key identity is

$$zz^* = (x + iy)(x - iy) = x^2 + y^2 = |z|^2, \tag{8.1}$$

which shows how to divide. For $w \neq 0$,

$$\frac{z}{w} = \frac{zw^*}{ww^*} = \frac{zw^*}{|w|^2}.$$

For example, $\dfrac{3 + 2i}{1 - i} = \dfrac{(3 + 2i)(1 + i)}{2} = \dfrac{1 + 5i}{2}$.

Theorem 8.2 (Properties of conjugate and modulus). *For $z, w \in \mathbb{C}$:*

(a) $(z + w)^* = z^* + w^*$, $(zw)^* = z^* w^*$, $z^{**} = z$;

(b) $\operatorname{Re} z = \frac{1}{2}(z + z^*)$, $\operatorname{Im} z = \frac{1}{2i}(z - z^*)$; *$z$ is real iff $z = z^*$;*

(c) $|zw| = |z||w|$, $|z^*| = |z|$, *and* $|\operatorname{Re} z| \leq |z|$;

(d) $|z + w| \leq |z| + |w|$ (*triangle inequality*).

Proof. (a) and (b) are direct computations. For (c), $|zw|^2 = zw(zw)^* = zz^* ww^* = |z|^2|w|^2$. For (d), $|z + w|^2 = |z|^2 + |w|^2 + 2\operatorname{Re}(zw^*) \leq |z|^2 + |w|^2 + 2|z||w|$, using (c). ∎

Geometrically, $\mathbb{C}$ is the plane $\mathbb{R}^2$, with $x + iy$ identified with the point (x, y): the *complex plane*, or *Argand diagram*. Addition is vector addition, $|z|$ is length, and $|z - w|$ is the distance between z and w. Conjugation is reflection in the real axis. Multiplication has no counterpart among the vector operations of Chapter 1, and it is what makes $\mathbb{C}$ special.

8.2 Polar Form and Euler's Formula

Writing $x = r\cos\theta$ and $y = r\sin\theta$ gives the *polar form* $z = r(\cos\theta + i\sin\theta)$, with $r = |z|$. The angle θ is an *argument* of z, written $\arg z$, and is determined only up to multiples of 2π. The *principal argument* $\operatorname{Arg} z \in (-\pi, \pi]$ picks one value.

Definition 8.3 (Complex exponential). For $z \in \mathbb{C}$, $e^z = \displaystyle\sum_{n=0}^{\infty} \frac{z^n}{n!}$.

The series converges absolutely for every z, by comparison with the real series for $e^{|z|}$. Multiplying two such series and collecting terms with the binomial theorem shows that $e^{z+w} = e^z e^w$.

Theorem 8.4 (Euler's formula). *For real θ, $e^{i\theta} = \cos\theta + i\sin\theta$.*

Proof. The powers of i cycle: $i^0 = 1$, $i^1 = i$, $i^2 = -1$, $i^3 = -i$. Separate the series into even and odd terms:

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(i\theta)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(i\theta)^{2k+1}}{(2k+1)!} = \sum_{k} \frac{(-1)^k \theta^{2k}}{(2k)!} + i\sum_{k} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} = \cos\theta + i\sin\theta.$$

∎

So every complex number is $z = re^{i\theta}$. At $\theta = \pi$, Euler's formula gives $e^{i\pi} + 1 = 0$, linking the five most important constants of mathematics. Multiplication in polar form is now transparent:

$$r_1 e^{i\theta_1} \cdot r_2 e^{i\theta_2} = r_1 r_2\, e^{i(\theta_1 + \theta_2)}.$$

To multiply complex numbers, multiply their moduli and add their arguments. In particular, multiplication by $e^{i\alpha}$ rotates the plane by α, and multiplication by $i = e^{i\pi/2}$ rotates by 90°.

Theorem 8.5 (De Moivre). $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ *for every integer n.*

Example 8.6 (Multiple-angle formulas). Expanding $(\cos\theta + i\sin\theta)^3$ by the binomial theorem and comparing real and imaginary parts gives

$$\cos 3\theta = \cos^3\theta - 3\cos\theta\sin^2\theta = 4\cos^3\theta - 3\cos\theta,$$
$$\sin 3\theta = 3\cos^2\theta\sin\theta - \sin^3\theta = 3\sin\theta - 4\sin^3\theta.$$

These are the triple-angle formulas, obtained with no trigonometric identities at all.

Example 8.7 (Inverting Euler). From $e^{\pm i\theta} = \cos\theta \pm i\sin\theta$:

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}, \qquad \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}.$$

Then $\cos^4\theta = \frac{1}{16}(e^{i\theta}+e^{-i\theta})^4 = \frac{1}{16}(2\cos 4\theta + 8\cos 2\theta + 6) = \frac{3}{8} + \frac{1}{2}\cos 2\theta + \frac{1}{8}\cos 4\theta$. Power-reduction formulas of any order follow in one line.

8.2.1 ROOTS

Theorem 8.8 (nth roots). *Every nonzero $w = \rho e^{i\alpha}$ has exactly n distinct nth roots,*

$$z_k = \rho^{1/n}\exp\Big(i\,\frac{\alpha+2\pi k}{n}\Big), \qquad k = 0, 1, \dots, n-1.$$

They lie equally spaced on the circle of radius $\rho^{1/n}$.

Proof. $z = re^{i\theta}$ satisfies $z^n = w$ iff $r^n = \rho$ and $n\theta = \alpha + 2\pi k$ for some integer k. Values of k differing by a multiple of n give the same z. ∎

The nth *roots of unity* are ω^k, with $\omega = e^{2\pi i/n}$ and $k = 0, \dots, n-1$. For $n \geq 2$ they sum to zero: $1 + \omega + \cdots + \omega^{n-1} = \frac{\omega^n - 1}{\omega - 1} = 0$. This identity underlies the discrete Fourier transform.

Example 8.9 (Cube roots). The cube roots of -8: $-8 = 8e^{i\pi}$, so the roots are $2e^{i\pi/3} = 1 + i\sqrt{3}$, $2e^{i\pi} = -2$, and $2e^{5i\pi/3} = 1 - i\sqrt{3}$.

8.2.2 PHASORS AND SUPERPOSITION

A real sinusoid $A\cos(\omega t + \varphi)$ is the real part of $Ae^{i\varphi}e^{i\omega t}$. The complex number $Ae^{i\varphi}$ is its *phasor*. Sums of sinusoids of a common frequency become sums of phasors.

Example 8.10 (Adding waves). Write $3\cos\omega t + 4\sin\omega t$ as $A\cos(\omega t + \varphi)$.

Solution. $\sin\omega t = \operatorname{Re}(-ie^{i\omega t})$, so the phasor is $3 - 4i = 5e^{-i\arctan(4/3)}$.

Hence $3\cos\omega t + 4\sin\omega t = 5\cos\left(\omega t - \arctan\frac{4}{3}\right)$.

LOOKING AHEAD. Two waves of equal amplitude with phasors 1 and $e^{i\varphi}$ superpose to a wave of amplitude $|1+e^{i\varphi}| = 2|\cos(\varphi/2)|$, and intensity $|1+e^{i\varphi}|^2 = 2+2\cos\varphi$. The intensity is maximal when $\varphi = 0$ and zero when $\varphi = \pi$: this is *interference*. In the double-slit experiment (Chapter 12), each electron's amplitude is a sum $\psi_1 + \psi_2$, and the probability $|\psi_1+\psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + 2\,\mathrm{Re}(\psi_1^*\psi_2)$ contains an interference term that classical probabilities never have.

8.3 Complex Functions

A *complex function* $f : \mathbb{C} \to \mathbb{C}$ can be written $f(x+iy) = u(x,y) + iv(x,y)$ with u, v real. Limits and continuity are defined exactly as for functions on $\mathbb{R}^2$ (Chapter 3), with $|z - z_0|$ as the distance.

Definition 8.11 (Elementary functions). For $z = x + iy$:

$$e^z = e^x(\cos y + i\sin y), \quad \cos z = \tfrac{1}{2}(e^{iz}+e^{-iz}), \quad \sin z = \tfrac{1}{2i}(e^{iz}-e^{-iz}),$$
$$\cosh z = \tfrac{1}{2}(e^z+e^{-z}), \qquad \sinh z = \tfrac{1}{2}(e^z-e^{-z}).$$

The complex exponential is periodic with period $2\pi i$, since $e^{z+2\pi i} = e^z$, and it is never zero. Complex sine and cosine are unbounded: $\cos(iy) = \cosh y$. The identities $\cos(iz) = \cosh z$ and $\sin(iz) = i\sinh z$ show that trigonometric and hyperbolic functions are the same functions, viewed along different directions in $\mathbb{C}$.

Definition 8.12 (Logarithm). For $z \neq 0$, the *logarithms* of z are the solutions w of $e^w = z$:

$$\log z = \ln|z| + i\arg z,$$

a multivalued expression whose values differ by integer multiples of $2\pi i$. The *principal logarithm* is $\mathrm{Log}\,z = \ln|z| + i\,\mathrm{Arg}\,z$. Complex powers are defined by $z^c = e^{c\log z}$.

For example, $\mathrm{Log}(-1) = i\pi$, and $i^i = e^{i\log i} = e^{i(i\pi/2+2\pi ik)} = e^{-\pi/2-2\pi k}$. Every value of i^i is a positive real number. The principal logarithm is continuous ev-

erywhere except on the negative real axis, where Arg jumps from π to $-\pi$. That ray is a *branch cut*. No continuous logarithm exists on any region that encircles the origin, which is the complex form of the winding-number phenomenon of Example 6.31.

8.4 Analytic Functions and the Cauchy–Riemann Equations

Definition 8.13 (Complex derivative). f is *complex-differentiable* at z_0 if

$$f'(z_0) = \lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

exists, where $h \to 0$ through complex values. f is *analytic* (or *holomorphic*) on an open set if it is complex-differentiable at every point of it. A function analytic on all of $\mathbb{C}$ is *entire*.

This looks like the one-variable definition. But h may approach 0 from any direction in the plane, and the limit must be the same from all of them. That is an enormously stronger requirement than real differentiability, and its consequences are what make complex analysis so powerful. The usual rules (sum, product, quotient, chain) hold with the usual proofs, and $\mathrm{d}/\mathrm{d}z\ z^n = nz^{n-1}$.

Theorem 8.14 (Cauchy–Riemann equations). *If $f = u + iv$ is complex-differentiable at $z_0 = x_0 + iy_0$, then at (x_0, y_0)*

$$u_x = v_y, \qquad u_y = -v_x,$$

and $f'(z_0) = u_x + iv_x$. Conversely, if u and v are C^1 near (x_0, y_0) and satisfy these equations there, then f is complex-differentiable at z_0.

Proof. Letting $h \to 0$ along the real axis ($h = t$) and along the imaginary axis ($h = it$) gives

$$f'(z_0) = \lim_{t \to 0} \frac{f(z_0 + t) - f(z_0)}{t} = u_x + iv_x,$$
$$f'(z_0) = \lim_{t \to 0} \frac{f(z_0 + it) - f(z_0)}{it} = \frac{1}{i}(u_y + iv_y) = v_y - iu_y.$$

Equating real and imaginary parts gives the equations. For the converse, differentiability of u and v (Theorem 3.15) gives $\Delta u = u_x \Delta x + u_y \Delta y + o(|h|)$, and like-

wise for v. Substituting the Cauchy–Riemann equations, $\Delta f = (u_x + iv_x)(\Delta x + i\Delta y) + o(|h|)$, which says exactly that $f'(z_0) = u_x + iv_x$. ∎

Example 8.15 (Analytic and non-analytic). (a) $f = z^2 = (x^2 - y^2) + i\,2xy$: $u_x = 2x = v_y$ and $u_y = -2y = -v_x$, so f is entire. (b) $f = z^* = x - iy$: $u_x = 1 \neq -1 = v_y$, so conjugation is nowhere complex-differentiable, although it is perfectly smooth as a map of $\mathbb{R}^2$. (c) $f = |z|^2$ satisfies the equations only at $z = 0$, so it is complex-differentiable there and nowhere else, and is analytic nowhere.

Corollary 8.16 (Analytic functions are harmonic). *If $f = u + iv$ is analytic and u, v are C^2, then $\nabla^2 u = \nabla^2 v = 0$.*

Proof. $u_{xx} = v_{yx}$ and $u_{yy} = -v_{xy}$; add and use Clairaut's theorem. Similarly for v. ∎

(C^2 is automatic: we will see that analytic functions are infinitely differentiable.) Conversely, every harmonic u on a simply connected region has a *harmonic conjugate* v, unique up to a constant, making $u + iv$ analytic. Solving the two-dimensional Laplace equation therefore reduces to finding analytic functions. That is why complex analysis is a standard tool in electrostatics, fluid flow, and heat conduction.

Example 8.17 (Finding a harmonic conjugate). $u = x^3 - 3xy^2$ is harmonic. From $v_y = u_x = 3x^2 - 3y^2$, $v = 3x^2 y - y^3 + g(x)$. Then $v_x = 6xy + g'(x)$ must equal $-u_y = 6xy$, so g is constant. So $u + iv = z^3$.

8.5 Contour Integrals

Definition 8.18 (Contour integral). Let C be a piecewise smooth oriented curve parametrized by $z(t) = x(t) + iy(t)$, $a \le t \le b$, and let f be continuous on C. Then

$$\int_C f(z)\,dz = \int_a^b f(z(t))\,z'(t)\,dt.$$

Writing $f = u + iv$ and $dz = dx + i\,dy$,

$$\int_C f\,dz = \int_C (u\,dx - v\,dy) + i\int_C (v\,dx + u\,dy), \tag{8.2}$$

so a contour integral is a pair of real line integrals (Chapter 6). It is linear, changes sign when the orientation is reversed, and is independent of orientation-preserving reparametrization.

Theorem 8.19 (The fundamental integral). *Let C_r be the circle $|z - z_0| = r$ traversed counterclockwise. For any integer n,*

$$\oint_{C_r} (z - z_0)^n\,dz = \begin{cases} 2\pi i, & n = -1, \\ 0, & n \neq -1. \end{cases}$$

Proof. With $z = z_0 + re^{it}$, $dz = ire^{it}\,dt$, the integral is $ir^{n+1}\int_0^{2\pi} e^{i(n+1)t}\,dt$. That is $2\pi i$ when $n = -1$, and zero otherwise, because $e^{i(n+1)t}$ then has period dividing 2π and integrates to zero. ∎

This single computation is the seed of everything that follows.

Theorem 8.20 (ML inequality). *If $|f(z)| \le M$ on C and C has length L, then $\left|\int_C f\,dz\right| \le ML$.*

Proof. $\left|\int_a^b f(z(t))z'(t)\,dt\right| \le \int_a^b |f(z(t))|\,|z'(t)|\,dt \le M\int_a^b |z'(t)|\,dt = ML$. The first inequality is the triangle inequality for integrals: write the integral as $Re^{i\alpha}$ and consider the real part of $e^{-i\alpha}\int f\,dz$. ∎

8.6 Cauchy's Theorem and Integral Formula

Theorem 8.21 (Cauchy's theorem). *If f is analytic on a simply connected open region D, then $\oint_C f(z)\,dz = 0$ for every closed piecewise smooth curve C in D.*

Proof assuming f' continuous. For a simple closed curve bounding a region $R \subseteq$

D, apply Green's theorem to each part of (8.2):

$$\oint_C (u\,dx - v\,dy) = \iint_R (-v_x - u_y)\,dA = 0,$$
$$\oint_C (v\,dx + u\,dy) = \iint_R (u_x - v_y)\,dA = 0,$$

both by the Cauchy–Riemann equations. Goursat later showed that the hypothesis that f' is continuous is unnecessary. ■

So contour integrals of analytic functions are path independent on simply connected regions, and analytic functions have antiderivatives there. As in Example 6.31, a curve can be deformed freely without crossing a singularity, and the integral does not change. Integrals around *different* curves enclosing the same singularities are equal.

Theorem 8.22 (Cauchy integral formula). *Let f be analytic on a simply connected region D, and let C be a positively oriented simple closed curve in D enclosing z_0. Then*

$$f(z_0) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z - z_0}\,dz, \qquad f^{(n)}(z_0) = \frac{n!}{2\pi i}\oint_C \frac{f(z)}{(z - z_0)^{n+1}}\,dz.$$

Proof. The integrand is analytic except at z_0, so C may be shrunk to a small circle C_r about z_0. Write

$$\oint_{C_r} \frac{f(z)}{z - z_0}dz = f(z_0)\oint_{C_r} \frac{dz}{z - z_0} + \oint_{C_r} \frac{f(z) - f(z_0)}{z - z_0}dz.$$

The first term is $2\pi i f(z_0)$ by Theorem 8.19. The second integrand is bounded near z_0 because f is differentiable there, so by the ML inequality the second term is $O(r) \to 0$. The derivative formula follows by differentiating under the integral sign with respect to z_0. ■

The Cauchy integral formula is astonishing. The values of an analytic function inside a curve are completely determined by its values on the curve. And because the right-hand side can be differentiated arbitrarily often, *an analytic function is infinitely differentiable*. Nothing remotely like this holds for real differentiable functions.

Example 8.23 (Using the integral formula). Evaluate $\oint_{|z|=2} \dfrac{e^z}{z^2 + 1}\,dz$.

Solution. Partial fractions give $\frac{1}{z^2+1} = \frac{1}{2i}\left(\frac{1}{z-i} - \frac{1}{z+i}\right)$. Both $\pm i$ lie inside the circle, so by Theorem 8.22 the integral is $\frac{2\pi i}{2i}(e^{i} - e^{-i}) = 2\pi i \sin 1$.

Theorem 8.24 (Liouville). *A bounded entire function is constant.*

Proof. If $|f| \le M$ on $\mathbb{C}$, the derivative formula on a circle of radius R about z_0, together with the ML inequality, gives $|f'(z_0)| \le \frac{1}{2\pi} \cdot \frac{M}{R^2} \cdot 2\pi R = M/R$. Let $R \to \infty$: $f' \equiv 0$. ∎

Corollary 8.25 (Fundamental theorem of algebra). *Every nonconstant polynomial with complex coefficients has a root in $\mathbb{C}$. Hence a polynomial of degree n factors as $c\prod_{k=1}^{n}(z - z_k)$.*

Proof. If p had no root, $1/p$ would be entire. It is also bounded, since $|p(z)| \to \infty$ as $|z| \to \infty$. By Liouville's theorem it would be constant, and so would p. Factoring out a root and inducting on the degree gives the factorization. ∎

8.7 Series, Singularities, and Residues

Theorem 8.26 (Taylor series). *If f is analytic on the disk $|z - z_0| < R$, then $f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$ there, with $a_n = f^{(n)}(z_0)/n!$.*

Proof. For $|z - z_0| < r < R$, expand the Cauchy kernel as a geometric series in $(z - z_0)/(w - z_0)$: $\frac{1}{w - z} = \sum_{n \ge 0} \frac{(z - z_0)^n}{(w - z_0)^{n+1}}$, uniformly for w on $|w - z_0| = r$. Substitute into the integral formula and integrate term by term. ∎

The radius of convergence is the distance to the nearest singularity. This explains a puzzle of real calculus: the series of $1/(1 + x^2)$ about 0 has radius 1 although the function is smooth on all of $\mathbb{R}$. Its singularities sit at $\pm i$, a distance 1 away in the complex plane.

Theorem 8.27 (Laurent series). *If f is analytic on the annulus $r < |z - z_0| < R$, then*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n, \qquad a_n = \frac{1}{2\pi i} \oint_C \frac{f(w)}{(w - z_0)^{n+1}}\, dw,$$

for any circle C in the annulus. The terms with $n<0$ form the principal part.

Definition 8.28 (Isolated singularities and residues). Let f be analytic on $0<|z-z_0|<R$, with Laurent coefficients a_n. The point z_0 is

- a *removable singularity* if $a_n=0$ for all $n<0$;
- a *pole of order m* if $a_{-m}\neq 0$ and $a_n=0$ for all $n<-m$ (a *simple pole* when $m=1$);
- an *essential singularity* if $a_n\neq 0$ for infinitely many $n<0$.

The *residue* of f at z_0 is $\operatorname{Res}_{z=z_0} f = a_{-1}$.

For example, $\frac{\sin z}{z}=1-\frac{z^2}{6}+\cdots$ has a removable singularity at 0; $\frac{1}{z^2(z-1)}$ has a pole of order 2 at 0 and a simple pole at 1; and $e^{1/z}=\sum_{n\ge 0}\frac{1}{n!z^n}$ has an essential singularity at 0, with residue 1.

By Theorem 8.19, integrating a Laurent series term by term around a small circle kills every term but one: $\oint f\,dz = 2\pi i\,a_{-1}$. The residue is what survives integration.

Theorem 8.29 (Computing residues). *(a) If z_0 is a simple pole, $\operatorname{Res}_{z_0} f = \lim_{z\to z_0}(z-z_0)f(z)$.*

(b) If $f=g/h$ with g,h analytic, $g(z_0)\neq 0$, $h(z_0)=0$, and $h'(z_0)\neq 0$, then $\operatorname{Res}_{z_0} f = g(z_0)/h'(z_0)$.

(c) If z_0 is a pole of order m, $\displaystyle \operatorname*{Res}_{z_0} f = \frac{1}{(m-1)!}\lim_{z\to z_0}\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}\big[(z-z_0)^m f(z)\big]$.

Proof. (c) $(z-z_0)^m f(z) = a_{-m}+\cdots+a_{-1}(z-z_0)^{m-1}+\cdots$; differentiate $m-1$ times and set $z=z_0$. Part (a) is the case $m=1$. For (b), $(z-z_0)\frac{g}{h} = \frac{g(z)}{(h(z)-h(z_0))/(z-z_0)}\to\frac{g(z_0)}{h'(z_0)}$. ∎

Theorem 8.30 (Residue theorem). *Let f be analytic inside and on a positively oriented simple closed curve C, except at finitely many isolated singularities $z_1,\dots,z_k$ inside C. Then*

$$\oint_C f(z)\,dz = 2\pi i\sum_{j=1}^{k}\operatorname*{Res}_{z=z_j} f.$$

Proof. Surround each z_j by a small circle. By Cauchy's theorem applied to the region between C and the circles (as in Example 6.31), $\oint_C$ equals the sum of the integrals over the circles, and each of those is $2\pi i$ times a residue. ∎

8.8 Evaluating Real Integrals

The residue theorem turns real integrals that resist all elementary methods into algebra. There are three standard types.

8.8.1 Rational functions on the real line

Theorem 8.31. *Let $f = p/q$ with polynomials p, q, $\deg q \geq \deg p + 2$, and q having no real zeros. Then*

$$\int_{-\infty}^{\infty} f(x)\,dx = 2\pi i \sum_{\operatorname{Im} z_j > 0} \operatorname*{Res}_{z=z_j} f.$$

Proof. Close the segment $[-R, R]$ with the semicircle Γ_R of radius R in the upper half-plane. For large R the closed contour encloses every upper-half-plane pole, so the residue theorem applies. On Γ_R, $|f| \leq K/R^2$, so by the ML inequality $\left|\int_{\Gamma_R} f\,dz\right| \leq \pi K/R \to 0$. ∎

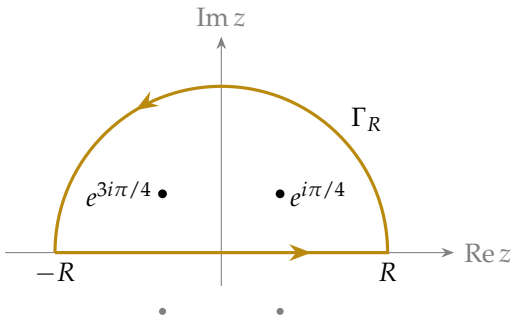


Figure 8.1: The upper semicircular contour for $\int dx/(1+x^4)$. Only the two poles in the upper half-plane contribute.

Example 8.32 ($\int dx/(1+x^4)$). The poles of $1/(1+z^4)$ are the fourth roots of -1. Those in the upper half-plane are $z_1 = e^{i\pi/4}$ and $z_2 = e^{3i\pi/4}$. By

Theorem 8.29(b), $\mathrm{Res}_{z_j} = \frac{1}{4z_j^3} = -\frac{z_j}{4}$, since $z_j^4 = -1$. So

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^4} = 2\pi i\Big(-\frac{z_1+z_2}{4}\Big) = -\frac{\pi i}{2}\cdot i\sqrt{2} = \frac{\pi}{\sqrt{2}},$$

using $z_1 + z_2 = i\sqrt{2}$. Hence $\int_0^{\infty} \frac{dx}{1+x^4} = \frac{\pi}{2\sqrt{2}}$, by symmetry.

8.8.2 TRIGONOMETRIC INTEGRALS OVER A PERIOD

For $\int_0^{2\pi} R(\cos\theta, \sin\theta)\, d\theta$, substitute $z = e^{i\theta}$. Then $\cos\theta = \frac{1}{2}(z + z^{-1})$, $\sin\theta = \frac{1}{2i}(z - z^{-1})$, and $d\theta = \frac{dz}{iz}$, so the integral becomes a contour integral over the unit circle.

Example 8.33 (A standard trigonometric integral). For $a > b > 0$, $\displaystyle\int_0^{2\pi} \frac{d\theta}{a + b\cos\theta} = \frac{2\pi}{\sqrt{a^2 - b^2}}$.

Solution. The integral becomes $\oint_{|z|=1} \frac{dz/(iz)}{a + \frac{b}{2}(z+z^{-1})} = \frac{2}{i}\oint \frac{dz}{bz^2 + 2az + b}$. The roots of the denominator are $z_{\pm} = \frac{-a \pm \sqrt{a^2-b^2}}{b}$. Their product is 1, so exactly one, z_+, lies inside the unit circle. Its residue is $\frac{1}{b(z_+ - z_-)} = \frac{1}{2\sqrt{a^2-b^2}}$. The integral is $\frac{2}{i}\cdot 2\pi i \cdot \frac{1}{2\sqrt{a^2-b^2}} = \frac{2\pi}{\sqrt{a^2-b^2}}$.

8.8.3 FOURIER-TYPE INTEGRALS

Lemma 8.34 (Jordan). *If $|f(z)| \le M_R \to 0$ on the upper semicircle Γ_R as $R \to \infty$, then $\int_{\Gamma_R} f(z)e^{ikz}\, dz \to 0$ for $k > 0$.*

The point is that $|e^{ikz}| = e^{-k\,\mathrm{Im}\, z}$ decays in the upper half-plane. The proof uses $\sin\theta \ge 2\theta/\pi$ on $[0, \pi/2]$ to show the integral is at most $\pi M_R/k$.

Example 8.35 (A Fourier transform). For $a > 0$ and $k \ge 0$, $\displaystyle\int_{-\infty}^{\infty} \frac{\cos kx}{x^2 + a^2}\, dx = \frac{\pi}{a}e^{-ka}$.

Solution. The integral is the real part of $\int \frac{e^{ikx}}{x^2+a^2}dx$. Close in the upper half-plane, which is allowed by Jordan's lemma. The only enclosed pole is $z = ia$, with residue $\frac{e^{ik(ia)}}{2ia} = \frac{e^{-ka}}{2ia}$. So the integral is $2\pi i \cdot \frac{e^{-ka}}{2ia} = \frac{\pi}{a}e^{-ka}$. The Lorentzian $1/(x^2 + a^2)$ is the line shape of a decaying quantum state,

and its Fourier transform is a decaying exponential: the time dependence of that decay.

LOOKING AHEAD. Contour integration is the workhorse of quantum scattering theory. Green's functions are built from integrals like those of Example 8.35, and the choice of whether to close the contour above or below a pole decides between outgoing and incoming waves. The analytic structure of amplitudes in the complex energy plane encodes physics directly: bound states appear as poles on the real axis, and unstable resonances appear as poles just below it.

8.9 Worked Problems

The problems below move from complex arithmetic and the multivalued logarithm, through Laurent expansions in several annuli, to four residue computations, each of which needs a different contour or a different trick. The final problem uses residues to sum an infinite series.

Problem 8.1 A quadratic with complex coefficients

Solve $z^2 - (3+i)z + (4+3i) = 0$.

Solution. The quadratic formula holds over $\mathbb{C}$, since its derivation (completing the square) uses only field operations. The discriminant is

$$\Delta = (3+i)^2 - 4(4+3i) = (8+6i) - (16+12i) = -8-6i.$$

To find $\sqrt{\Delta} = a+bi$, solve $a^2 - b^2 = -8$ and $2ab = -6$. Also $a^2 + b^2 = |\Delta| = 10$. Adding, $a^2 = 1$, and then $b^2 = 9$, with $ab < 0$. So $\sqrt{\Delta} = \pm(1-3i)$. Check: $(1-3i)^2 = 1 - 6i - 9 = -8 - 6i$. Then

$$z = \frac{(3+i) \pm (1-3i)}{2} = 2 - i \quad \text{or} \quad 1 + 2i.$$

Check with Vieta: the sum is $3+i$, and the product is $(2-i)(1+2i) = 2 + 4i - i + 2 = 4 + 3i$. Note that the roots are not conjugates. The conjugate-root theorem (Theorem 8.2) requires real coefficients.

Problem 8.2 A trigonometric sum by geometric series

Prove that for $\theta \notin 2\pi\mathbb{Z}$,

$$\sum_{k=0}^{n} \cos k\theta = \frac{\sin \frac{(n+1)\theta}{2} \cos \frac{n\theta}{2}}{\sin \frac{\theta}{2}}.$$

Solution. The sum is the real part of the geometric series $\sum_{k=0}^{n} e^{ik\theta} = \dfrac{e^{i(n+1)\theta} - 1}{e^{i\theta} - 1}$, valid since $e^{i\theta} \neq 1$. Factor out half-angles from numerator and denominator, using $e^{i\alpha} - 1 = e^{i\alpha/2} \cdot 2i \sin \frac{\alpha}{2}$:

$$\frac{e^{i(n+1)\theta/2} \cdot 2i \sin \frac{(n+1)\theta}{2}}{e^{i\theta/2} \cdot 2i \sin \frac{\theta}{2}} = e^{in\theta/2} \frac{\sin \frac{(n+1)\theta}{2}}{\sin \frac{\theta}{2}}.$$

The real part is the stated formula, and the imaginary part gives $\sum_{k=0}^{n} \sin k\theta$ with $\sin \frac{n\theta}{2}$ in place of $\cos \frac{n\theta}{2}$. This sum, in the form $\sum_{k=-n}^{n} e^{ik\theta} = \frac{\sin\left(n+\frac{1}{2}\right)\theta}{\sin \frac{\theta}{2}}$, is the *Dirichlet kernel*, which controls the convergence of Fourier series in Chapter 11. Real-variable proofs require repeated product-to-sum identities. Euler's formula turns them into the algebra of a geometric series.

Problem 8.3 The many values of i^i

Define $z^w = e^{w \log z}$, where $\log z = \ln|z| + i \arg z$ is multivalued. Find all values of i^i, and show by example that the principal logarithm Log does not satisfy $\operatorname{Log}(z_1 z_2) = \operatorname{Log} z_1 + \operatorname{Log} z_2$.

Solution. $\log i = \ln 1 + i\left(\frac{\pi}{2} + 2\pi k\right)$ for $k \in \mathbb{Z}$. So

$$i^i = e^{i \log i} = e^{i \cdot i(\pi/2 + 2\pi k)} = e^{-\pi/2 - 2\pi k}, \qquad k \in \mathbb{Z}.$$

Every value is a positive real number. The principal value ($k = 0$) is $e^{-\pi/2} \approx 0.2079$.

For the second part, take $z_1 = z_2 = -1$, with $\operatorname{Log}(-1) = i\pi$. Then $\operatorname{Log} z_1 + \operatorname{Log} z_2 = 2\pi i$, but $\operatorname{Log}(z_1 z_2) = \operatorname{Log} 1 = 0$. The identity holds only up to a multiple of $2\pi i$, because arguments add, but the principal argument is confined to $(-\pi, \pi]$. The same issue appears in the false "proof" $1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = i^2 = -1$: the rule $\sqrt{ab} = \sqrt{a}\sqrt{b}$ for principal roots fails off the positive reals.

Problem 8.4 One function, three Laurent series

Expand $f(z) = \dfrac{1}{(z-1)(z-2)}$ in Laurent series about 0 valid in (a) $|z| < 1$, (b) $1 < |z| < 2$, (c) $|z| > 2$.

Solution. By partial fractions $f = \dfrac{1}{z-2} - \dfrac{1}{z-1}$. Each term is expanded as a geometric series in whichever ratio is less than 1 in modulus on the given annulus.

$$\frac{1}{z-2} = \begin{cases} -\dfrac{1}{2}\dfrac{1}{1-z/2} = -\displaystyle\sum_{n\ge 0} \frac{z^n}{2^{n+1}}, & |z| < 2, \\ \dfrac{1}{z}\dfrac{1}{1-2/z} = \displaystyle\sum_{n\ge 1} \frac{2^{n-1}}{z^n}, & |z| > 2, \end{cases} \qquad \frac{1}{z-1} = \begin{cases} -\displaystyle\sum_{n\ge 0} z^n, & |z| < 1, \\ \displaystyle\sum_{n\ge 1} \frac{1}{z^n}, & |z| > 1. \end{cases}$$

Combining:

(a) $|z| < 1$: $f = \displaystyle\sum_{n\ge 0} \Big(1 - \frac{1}{2^{n+1}}\Big) z^n = \frac{1}{2} + \frac{3}{4}z + \frac{7}{8}z^2 + \cdots$, an ordinary Taylor series.

(b) $1 < |z| < 2$: $f = -\displaystyle\sum_{n\ge 0} \frac{z^n}{2^{n+1}} - \sum_{n\ge 1} \frac{1}{z^n}$, with infinitely many terms of both kinds.

(c) $|z| > 2$: $f = \displaystyle\sum_{n\ge 1} \frac{2^{n-1} - 1}{z^n} = \frac{1}{z^2} + \frac{3}{z^3} + \frac{7}{z^4} + \cdots$.

The coefficient of $1/z$ in (c) is 0. That is consistent with the residue theorem: a circle $|z| = R > 2$ encloses both poles, whose residues are -1 (at 1) and $+1$ (at 2), so $\oint_{|z|=R} f\,dz = 0 = 2\pi i\, c_{-1}$. In (b), the coefficient of $1/z$ is -1, the residue at the single enclosed pole $z = 1$. The Laurent coefficients depend on the annulus, and the coefficient of $1/z$ always equals the sum of the residues inside.

Problem 8.5 A sixth-degree denominator

Evaluate $\displaystyle\int_0^\infty \frac{dx}{1+x^6}$.

Solution. The integrand is even, so the integral is half of $\int_{-\infty}^{\infty}$, which by Theorem 8.31 is $2\pi i$ times the sum of the residues in the upper half-plane. The poles are the sixth roots of -1, namely $z_k = e^{i\pi(2k+1)/6}$, and the ones in the upper half-plane are $e^{i\pi/6}$, i, $e^{5i\pi/6}$. At a simple zero of the denominator,

$$\operatorname{Res}_{z=z_k} \frac{1}{1+z^6} = \frac{1}{6z_k^5} = \frac{z_k}{6z_k^6} = -\frac{z_k}{6},$$

using $z_k^6 = -1$. This trick avoids computing fifth powers. The sum of the three upper roots is $e^{i\pi/6} + e^{5i\pi/6} + i = (\cos\frac{\pi}{6} + \cos\frac{5\pi}{6}) + i(\sin\frac{\pi}{6} + \sin\frac{5\pi}{6} + 1) = 2i$, so the residues sum to $-\frac{i}{3}$. Hence

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^6} = 2\pi i \cdot \Big(-\frac{i}{3}\Big) = \frac{2\pi}{3}, \qquad \int_0^{\infty} \frac{dx}{1+x^6} = \frac{\pi}{3}.$$

In general $\int_0^\infty \frac{dx}{1+x^n} = \frac{\pi/n}{\sin(\pi/n)}$, which gives $\frac{\pi}{2}$ for $n=2$, $\frac{\pi}{2\sqrt{2}}$ for $n=4$ (Example 8.32), and $\frac{\pi}{3}$ here.

Problem 8.6 A trigonometric integral with a higher harmonic

Evaluate $\displaystyle\int_0^{2\pi} \frac{\cos 2\theta}{5 - 4\cos\theta}\, d\theta$.

Solution. Set $z = e^{i\theta}$, so $d\theta = \frac{dz}{iz}$ and $\cos\theta = \frac{1}{2}(z + z^{-1})$. Write $\cos 2\theta = \operatorname{Re} z^2$ and compute $J = \oint_{|z|=1} \dfrac{z^2}{5 - 2(z+z^{-1})} \dfrac{dz}{iz}$. Simplifying,

$$\frac{z^2}{5 - 2z - 2z^{-1}} \cdot \frac{1}{iz} = \frac{z^2}{i(5z - 2z^2 - 2)} = \frac{iz^2}{2z^2 - 5z + 2} = \frac{iz^2}{(2z-1)(z-2)}.$$

Only the pole $z = \frac{1}{2}$ lies inside the unit circle, with residue $\dfrac{i(1/4)}{2(\frac{1}{2} - 2)} = -\dfrac{i}{12}$. Then $J = 2\pi i \cdot (-\frac{i}{12}) = \frac{\pi}{6}$. This is real, so its imaginary part, $\int \frac{\sin 2\theta}{5-4\cos\theta} d\theta$, is 0, as oddness about $\theta = \pi$ requires, and

$$\int_0^{2\pi} \frac{\cos 2\theta}{5 - 4\cos\theta}\, d\theta = \frac{\pi}{6}.$$

Replacing $\cos 2\theta$ by $\operatorname{Re} z^2$ rather than $\frac{1}{2}(z^2 + z^{-2})$ was essential. The latter introduces a pole of order two at the origin and doubles the work.

Problem 8.7 A rectangular contour

Show that for $0 < a < 1$,

$$\int_{-\infty}^{\infty} \frac{e^{ax}}{1+e^x}\, dx = \frac{\pi}{\sin \pi a}.$$

Solution. Let $f(z) = \dfrac{e^{az}}{1+e^z}$. The key property is that $1 + e^z$ is periodic with period $2\pi i$, while e^{az} picks up the factor $e^{2\pi i a}$: $f(z + 2\pi i) = e^{2\pi i a} f(z)$. Integrate around the rectangle with vertices $\pm R$ and $\pm R + 2\pi i$, counterclockwise.

- The bottom edge gives $I_R = \int_{-R}^{R} f(x)\,dx$. The top edge, traversed from right to left, gives $-e^{2\pi i a} I_R$.
- On the right edge, $|f| \le \dfrac{e^{aR}}{e^R - 1} \to 0$ since $a < 1$. On the left edge, $|f| \le \dfrac{e^{-aR}}{1 - e^{-R}} \to 0$ since $a > 0$. Both edges have fixed length 2π, so their contributions vanish.
- Inside, $1 + e^z = 0$ only at $z = i\pi$, a simple pole with residue $\dfrac{e^{ai\pi}}{e^{i\pi}} = -e^{i\pi a}$.

Letting $R \to \infty$,

$$(1 - e^{2\pi i a})I = -2\pi i e^{i\pi a} \quad \Longrightarrow \quad I = \frac{-2\pi i e^{i\pi a}}{1 - e^{2\pi i a}} = \frac{2\pi i}{e^{i\pi a} - e^{-i\pi a}} = \frac{\pi}{\sin \pi a}.$$

Substituting $x = \ln t$ turns this into $\int_0^\infty \frac{t^{a-1}}{1+t}\,dt$, the integral that one of this chapter's exercises evaluates with a keyhole contour. The two contours are the same contour seen through the exponential map, which wraps the strip $0 \le \operatorname{Im} z \le 2\pi$ onto the plane cut along the positive axis.

Problem 8.8 Summing a series by residues

Show that for $a > 0$,

$$\sum_{n=-\infty}^{\infty} \frac{1}{n^2 + a^2} = \frac{\pi}{a} \coth \pi a,$$

and deduce $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Solution. The function $\pi \cot \pi z$ has simple poles at every integer n, with residue $\frac{\pi \cos \pi n}{\pi \cos \pi n} = 1$. So $g(z) = \dfrac{\pi \cot \pi z}{z^2 + a^2}$ has residue $\dfrac{1}{n^2 + a^2}$ at $z = n$. It also has simple poles at $z = \pm ia$. Since $\cot(\pm i\pi a) = \mp i \coth \pi a$,

$$\operatorname{Res}_{z=\pm ia} g = \frac{\pi \cot(\pm i\pi a)}{\pm 2ia} = \frac{\mp i\pi \coth \pi a}{\pm 2ia} = -\frac{\pi \coth \pi a}{2a} \quad \text{(both signs)}.$$

Integrate g around the square C_N with corners $(N + \frac{1}{2})(\pm 1 \pm i)$. On these squares $|\cot \pi z|$ is bounded by a constant independent of N (on the vertical sides $|\cot \pi z| = |\tanh \pi y| \le 1$, and on the horizontal sides $|\cot \pi z| \le \coth \frac{\pi}{2}$). Since $|z^2 + a^2| \ge (N + \frac{1}{2})^2 - a^2$ and the perimeter is $8(N + \frac{1}{2})$, the ML estimate (Theorem 8.20) gives $\oint_{C_N} g\,dz \to 0$. By the residue theorem, the sum of all residues is zero:

$$\sum_{n=-\infty}^{\infty} \frac{1}{n^2 + a^2} - \frac{\pi \coth \pi a}{a} = 0.$$

For the second part, separate the $n = 0$ term and use symmetry: $\sum_{n\geq 1} \frac{1}{n^2+a^2} = \frac{\pi a \coth \pi a - 1}{2a^2}$. As $a \to 0$, $\pi a \coth \pi a = 1 + \frac{(\pi a)^2}{3} + O(a^4)$, so the right side tends to $\frac{\pi^2}{6}$. The interchange of the limit and the sum is justified because the series converges uniformly in a by comparison with $\sum 1/n^2$. This is Euler's 1734 result, here as a by-product of a contour integral.

Problem 8.9 The integral formula for derivatives

Evaluate $\displaystyle\oint_{|z|=2} \frac{e^z}{(z-1)^3}\,dz$ and $\displaystyle\oint_{|z|=1} z^3 \cos\frac{1}{z}\,dz$.

Solution. For the first, the generalized Cauchy integral formula $f^{(n)}(a) = \frac{n!}{2\pi i}\oint \frac{f(z)}{(z-a)^{n+1}}dz$ with $f = e^z$, $a = 1$, $n = 2$ gives $\oint \frac{e^z}{(z-1)^3}dz = \frac{2\pi i}{2!}e^1 = \pi i e$.

For the second, the singularity at 0 is essential, so no derivative formula applies. Instead read off the residue from the Laurent series: $z^3\cos\frac{1}{z} = z^3\sum_{k\geq 0}\frac{(-1)^k}{(2k)!\,z^{2k}} = z^3 - \frac{z}{2} + \frac{1}{24z} - \frac{1}{720z^3} + \cdots$. The coefficient of $1/z$ comes from $k = 2$ and is $\frac{1}{24}$, so $\oint z^3\cos\frac{1}{z}\,dz = 2\pi i \cdot \frac{1}{24} = \frac{\pi i}{12}$. The residue theorem needs only the coefficient c_{-1}, which is defined for any isolated singularity, however wild.

EXERCISES

A. COMPUTATION

8.1 Compute $(2-3i)(4+i)$, $\dfrac{1+2i}{3-4i}$, and $|(1+i)^{10}|$.

8.2 Find all z with $z^2 = -5 + 12i$, in rectangular form.

8.3 Write $1 - i\sqrt{3}$ and $-2-2i$ in polar form, and compute $(1-i\sqrt{3})^9$.

8.4 Find all solutions of $z^4 = -16$ and plot them.

8.5 Find all solutions of $z^6 = 1$ and verify that they sum to zero.

8.6 Use De Moivre's theorem to express $\cos 5\theta$ as a polynomial in $\cos\theta$.

8.7 Express $\sin^5\theta$ as a linear combination of $\sin\theta$, $\sin 3\theta$, $\sin 5\theta$.

8.8 Evaluate $\sum_{k=0}^{n}\cos k\theta$ in closed form by summing a geometric series of complex exponentials.

8.9 Write $2\cos(\omega t) - 2\sqrt{3}\sin(\omega t)$ in the form $A\cos(\omega t + \varphi)$.

8.10 Find all values of $\log(1+i)$, $(1+i)^i$, and $i^{1/3}$.

8.11 Solve $e^z = -2$ and $\cos z = 2$.

8.12 Compute $|e^{iz}|$ for $z = x + iy$, and show that $|\sin z|^2 = \sin^2 x + \sinh^2 y$.

8.13 Determine where $f(z) = x^2 + iy^2$ is complex-differentiable. Where is it analytic?

8.14 Verify the Cauchy–Riemann equations for $f(z) = e^z$ and $f(z) = 1/z$ ($z \neq 0$).

8.15 Show that $u = e^x \cos y + x^2 - y^2$ is harmonic, find a harmonic conjugate, and express $u + iv$ in terms of z.

8.16 Evaluate $\int_C z^* \, dz$ along (a) the segment from 0 to $1+i$ and (b) the path $0 \to 1 \to 1+i$. Explain why they differ.

8.17 Evaluate $\oint_{|z|=1} \dfrac{dz}{z^2}$, $\oint_{|z|=1} \dfrac{\cos z}{z} dz$, and $\oint_{|z|=1} \bar{z}\, dz$.

8.18 Evaluate $\oint_{|z|=3} \dfrac{z^2}{z-2}\, dz$ and $\oint_{|z|=3} \dfrac{e^{2z}}{(z-1)^3}\, dz$.

8.19 Evaluate $\oint_{|z-i|=1} \dfrac{dz}{z^2+1}$.

8.20 Find the Laurent series of $\dfrac{1}{z(z-1)}$ valid in (a) $0 < |z| < 1$ and (b) $|z| > 1$.

8.21 Classify the singularities and find the residues of $\dfrac{z}{(z-1)(z+2)^2}$.

8.22 Find the residue at 0 of $\dfrac{1}{z^2 \sin z}$, of $\dfrac{e^z}{z^3}$, and of $z^2 e^{1/z}$.

8.23 Evaluate $\oint_{|z|=2} \tan z \, dz$.

8.24 Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{(x^2+1)(x^2+4)}$.

8.25 Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{x^2}{(x^2+1)^2}\, dx$.

8.26 Evaluate $\displaystyle\int_0^{\infty} \frac{dx}{1+x^6}$.

8.27 Evaluate $\displaystyle\int_0^{2\pi} \frac{d\theta}{5 - 4\cos\theta}$ and $\displaystyle\int_0^{2\pi} \frac{d\theta}{(2+\cos\theta)^2}$.

8.28 Evaluate $\displaystyle\int_0^{2\pi} \cos^{2n}\theta\, d\theta$ by the substitution $z = e^{i\theta}$.

8.29 Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{x \sin x}{x^2+4}\, dx$.

8.30 Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{\cos x}{(x^2+1)^2}\, dx$.

B. THEORY AND PROOF

8.31 Prove that $e^{z+w} = e^z e^w$ from the power series.

8.32 Prove that if f is analytic and real-valued on a connected open set, it is constant.

8.33 Prove that if f is analytic and $|f|$ is constant on a connected open set, then f is constant.

8.34 Show that an analytic function preserves angles between curves wherever $f' \neq 0$ (it is *conformal*).

8.35 Prove the Cauchy estimate $|f^{(n)}(z_0)| \leq n!M/r^n$, where $|f| \leq M$ on $|z - z_0| = r$.

8.36 Prove the *mean value property*: $f(z_0) = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{it})\, dt$. Deduce the maximum modulus principle.

8.37 Show that the nth roots of unity satisfy $\prod_{k=1}^{n-1}(z - \omega^k) = 1 + z + \cdots + z^{n-1}$, and deduce $\prod_{k=1}^{n-1}|1 - \omega^k| = n$.

8.38 Prove Jordan's lemma (Lemma 8.34).

8.39 Show that if p is a polynomial with real coefficients, its nonreal roots occur in conjugate pairs.

C. CHALLENGE

8.40 Evaluate $\displaystyle\int_0^{\infty} \frac{\sin x}{x}\, dx = \frac{\pi}{2}$ by integrating e^{iz}/z around a contour consisting of $[-R, -\epsilon]$, a small upper semicircle about 0, $[\epsilon, R]$, and the large semicircle. Compare with Example 5.33.

8.41 Show that $\displaystyle\int_0^{\infty} \frac{x^{a-1}}{1+x}\, dx = \frac{\pi}{\sin \pi a}$ for $0 < a < 1$, using a keyhole contour and the branch cut of z^{a-1}.

8.42 Use the residue theorem with $\pi \cot \pi z$ to prove $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. (Integrate $\pi \cot(\pi z)/z^2$ around a large square, and show the integral tends to zero.)

8.43 Evaluate the Fresnel integrals $\int_0^\infty \cos(x^2)\,dx = \int_0^\infty \sin(x^2)\,dx = \frac{1}{2}\sqrt{\pi/2}$ by integrating e^{iz^2} around the boundary of the sector $0 \le \arg z \le \pi/4$, $|z| \le R$.

CHAPTER 9

Linear Algebra and Hilbert Space

LEARNING OBJECTIVES

- State the axioms of a vector space over $\mathbb{R}$ or $\mathbb{C}$, and work with subspaces, span, linear independence, bases, and dimension.
- Represent linear maps by matrices, and use kernel, image, rank–nullity, determinants, and inverses.
- Work in complex inner-product spaces: prove Cauchy–Schwarz, construct orthonormal bases by Gram–Schmidt, and compute projections.
- Find eigenvalues and eigenvectors, and diagonalize matrices.
- Prove that Hermitian operators have real eigenvalues and orthogonal eigenvectors, and prove the spectral theorem.
- Prove that commuting Hermitian operators can be diagonalized simultaneously.
- Use Dirac notation, including outer products, projectors, and completeness relations.
- Extend these ideas to spaces of functions, and show that $-i\,d/dx$ is Hermitian.

If a single branch of mathematics had to be named as the language of quantum mechanics, it would be linear algebra. A quantum state is a vector. A measurable quantity is a Hermitian operator. The possible results of a measurement are that operator's eigenvalues, the probabilities are squared inner products, and time evolution is a unitary transformation. Everything in Part III is an application of this chapter, usually in infinite dimensions. We develop the theory in finite dimensions, where every statement can be proved completely, and then show how it extends to spaces of functions.

9.1 VECTOR SPACES

In Chapter 1 we observed that the arrows of $\mathbb{R}^3$ obey eight algebraic rules. Here we turn those rules into a definition.

Definition 9.1 (Vector space). Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. A *vector space over* $\mathbb{F}$ is a set V with an addition $V \times V \to V$ and a scalar multiplication $\mathbb{F} \times V \to V$ such that for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in \mathbb{F}$: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$; $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$; there is a zero vector $\mathbf{0}$ with $\mathbf{u} + \mathbf{0} = \mathbf{u}$; each $\mathbf{u}$ has an additive inverse $-\mathbf{u}$; $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$; $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$; $(ab)\mathbf{u} = a(b\mathbf{u})$; and $1\mathbf{u} = \mathbf{u}$.

Example 9.2 (Vector spaces). (a) $\mathbb{R}^n$ and $\mathbb{C}^n$, with componentwise operations. In quantum mechanics a spin-$\frac{1}{2}$ state lives in $\mathbb{C}^2$.

(b) $\mathcal{P}_n$, the polynomials of degree at most n, with ordinary addition and scalar multiplication.

(c) $C[a, b]$, the continuous functions on $[a, b]$. A "vector" here is an entire function.

(d) The set of solutions of a homogeneous linear differential equation such as $y'' + \omega^2 y = 0$. A sum or scalar multiple of solutions is again a solution: this is the *superposition principle*.

(e) The $m \times n$ matrices $M_{m\times n}(\mathbb{F})$.

Definition 9.3 (Subspace, span, independence, basis). A *subspace* of V is a nonempty subset closed under addition and scalar multiplication. The *span* of $\mathbf{v}_1, \dots, \mathbf{v}_k$ is the set of all linear combinations $\sum c_i \mathbf{v}_i$. The vectors are *linearly independent* if $\sum c_i \mathbf{v}_i = \mathbf{0}$ only when every $c_i = 0$. A *basis* is a linearly independent spanning set.

Theorem 9.4 (Dimension). *If V has a finite spanning set, then every basis of V has the same number of elements, the* dimension $\dim V$. *Every vector has* unique *coordinates with respect to a basis.*

Proof. The key step is the *exchange lemma*: if $\mathbf{w}_1, \dots, \mathbf{w}_m$ are independent and $\mathbf{v}_1, \dots, \mathbf{v}_n$ span V, then $m \le n$. To see it, $\mathbf{w}_1$ is a combination of the $\mathbf{v}_j$ with some nonzero coefficient, say on $\mathbf{v}_1$. Solve for $\mathbf{v}_1$; then $\mathbf{w}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ still span.

Repeat with $\mathbf{w}_2$, which must involve some remaining $\mathbf{v}_j$, since otherwise it would depend on $\mathbf{w}_1$. If $m > n$, after n steps the $\mathbf{w}$'s alone would span, making $\mathbf{w}_{n+1}$ dependent on them, a contradiction. Applying the lemma in both directions to two bases shows they have equal size. If $\mathbf{v} = \sum a_i \mathbf{e}_i = \sum b_i \mathbf{e}_i$, then $\sum (a_i - b_i)\mathbf{e}_i = \mathbf{0}$, and independence forces $a_i = b_i$. ∎

So $\dim \mathbb{R}^n = n$, $\dim \mathbb{C}^n = n$ over $\mathbb{C}$, and $\dim \mathcal{P}_n = n+1$ (basis $1, x, \dots, x^n$). The space $C[a,b]$ contains the independent functions $1, x, x^2, \dots$ for every n, so it is *infinite-dimensional*. So is every state space for a particle moving in space.

9.2 Linear Maps and Matrices

Definition 9.5 (Linear map). A map $T : V \to W$ between vector spaces is *linear* if $T(a\mathbf{u} + b\mathbf{v}) = aT\mathbf{u} + bT\mathbf{v}$. A linear map $V \to V$ is a *linear operator*. The *kernel* is $\ker T = \{\mathbf{v} : T\mathbf{v} = \mathbf{0}\}$ and the *image* is $\operatorname{im} T = \{T\mathbf{v}\}$.

Differentiation $D = \mathrm{d}/\mathrm{d}x$ is a linear operator on smooth functions, as are multiplication by x and the integral $f \mapsto \int_0^x f$. In quantum mechanics every observable, from position and momentum to energy and spin, is a linear operator.

Fix bases $\{\mathbf{e}_j\}$ of V and $\{\mathbf{f}_i\}$ of W. A linear map is determined by what it does to basis vectors. Write $T\mathbf{e}_j = \sum_i A_{ij}\mathbf{f}_i$. Then for $\mathbf{v} = \sum_j v_j \mathbf{e}_j$,

$$T\mathbf{v} = \sum_j v_j T\mathbf{e}_j = \sum_i \Big(\sum_j A_{ij} v_j \Big) \mathbf{f}_i .$$

In coordinates, T acts by the matrix $A = [A_{ij}]$, whose jth column is the coordinate vector of $T\mathbf{e}_j$. Composition of maps corresponds to multiplication of matrices, $(AB)_{ik} = \sum_j A_{ij} B_{jk}$. That is why matrix multiplication is defined this way, and why it is associative but not commutative.

Example 9.6 (Differentiation as a matrix). On $\mathcal{P}_3$ with basis $1, x, x^2, x^3$, $D(x^k) = k x^{k-1}$, so

$$[D] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This matrix is nilpotent, $[D]^4 = 0$, reflecting the fact that four derivatives annihilate a cubic. Its kernel, the constants, is one-dimensional.

Theorem 9.7 (Rank–nullity). *For a linear map $T : V \to W$ with V finite-dimensional,* $\dim \ker T + \dim \operatorname{im} T = \dim V$.

Proof. Choose a basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of $\ker T$ and extend it to a basis $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_r$ of V. The vectors $T\mathbf{v}_1, \dots, T\mathbf{v}_r$ span $\operatorname{im} T$, since $T\mathbf{u}_i = \mathbf{0}$. They are independent: if $\sum c_j T\mathbf{v}_j = \mathbf{0}$, then $\sum c_j \mathbf{v}_j \in \ker T$ is a combination of the $\mathbf{u}_i$, and independence of the whole basis forces every $c_j = 0$. So $\dim \operatorname{im} T = r = \dim V - k$. ∎

Corollary 9.8. *For an operator T on a finite-dimensional space, T is injective $\iff$ surjective $\iff$ invertible. For a square matrix, $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ $\iff$ $A\mathbf{x} = \mathbf{0}$ has only the zero solution.*

9.2.1 DETERMINANTS

The determinant of an $n \times n$ matrix is the unique function of its columns that is linear in each column, changes sign when two columns are swapped, and equals 1 on the identity. Explicitly,

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)} = \varepsilon_{i_1 \cdots i_n} A_{1 i_1} \cdots A_{n i_n},$$

where the sum runs over all permutations and ε is the n-index Levi-Civita symbol. Geometrically, $|\det A|$ is the factor by which A scales volumes (Lemma 5.24).

Theorem 9.9 (Properties of determinants). $\det(AB) = \det A \det B$; $\det A^T = \det A$; *A is invertible iff* $\det A \neq 0$; *and* $\det A^{-1} = 1/\det A$. *Row reduction changes the determinant in a predictable way: swapping rows negates it, scaling a row scales it, and adding a multiple of one row to another leaves it unchanged.*

9.3 INNER PRODUCT SPACES

On $\mathbb{C}^n$, the formula $\sum u_i v_i$ will not do as an inner product: $\sum z_i z_i$ can be zero or negative for nonzero $\mathbf{z}$. The fix is to conjugate one factor, which makes $\langle \mathbf{z}, \mathbf{z} \rangle = \sum |z_i|^2 \geq 0$.

Definition 9.10 (Inner product). An *inner product* on a complex vector space V is a map $\langle \cdot,\cdot\rangle : V\times V\to\mathbb{C}$ such that for all $\mathbf{u},\mathbf{v},\mathbf{w}$ and $a,b\in\mathbb{C}$:

(i) $\langle \mathbf{u},\mathbf{v}\rangle = \langle \mathbf{v},\mathbf{u}\rangle^{*}$ (conjugate symmetry)

(ii) $\langle \mathbf{u}, a\mathbf{v}+b\mathbf{w}\rangle = a\langle \mathbf{u},\mathbf{v}\rangle + b\langle \mathbf{u},\mathbf{w}\rangle$ (linearity in the second slot)

(iii) $\langle \mathbf{v},\mathbf{v}\rangle > 0$ for $\mathbf{v}\neq\mathbf{0}$ (positive definiteness)

The *norm* is $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v},\mathbf{v}\rangle}$. A complete inner-product space is a *Hilbert space*.

By (i) and (ii), $\langle a\mathbf{u},\mathbf{v}\rangle = a^{*}\langle \mathbf{u},\mathbf{v}\rangle$: the inner product is *antilinear* in the first slot. This is the physics convention. Many mathematics texts conjugate the second slot instead. Every finite-dimensional inner-product space is complete, so it is a Hilbert space automatically.

Example 9.11 (Standard inner products). (a) On $\mathbb{C}^n$, $\langle \mathbf{u},\mathbf{v}\rangle = \sum_i u_i^{*}v_i = \mathbf{u}^{\dagger}\mathbf{v}$, where $\mathbf{u}^{\dagger}$ is the conjugate transpose. (b) On complex-valued continuous functions on $[a,b]$, $\langle f,g\rangle = \int_a^b f(x)^{*}g(x)\,dx$. (c) With a positive *weight* $w(x)$, $\langle f,g\rangle_w = \int_a^b f^{*}g\,w\,dx$. Weighted inner products define the classical orthogonal polynomials of Chapter 10.

Theorem 9.12 (Cauchy–Schwarz in inner-product spaces). $|\langle \mathbf{u},\mathbf{v}\rangle| \le \|\mathbf{u}\|\,\|\mathbf{v}\|$, *with equality iff* $\mathbf{u}$ *and* $\mathbf{v}$ *are linearly dependent.*

Proof. If $\mathbf{u}=\mathbf{0}$ there is nothing to prove. Otherwise let $\mathbf{w} = \mathbf{v} - \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|^2}\mathbf{u}$, the component of $\mathbf{v}$ orthogonal to $\mathbf{u}$, so that $\langle \mathbf{u},\mathbf{w}\rangle = 0$. Then

$$0 \le \|\mathbf{w}\|^2 = \langle \mathbf{w},\mathbf{v}\rangle = \|\mathbf{v}\|^2 - \frac{\langle \mathbf{u},\mathbf{v}\rangle^{*}\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|^2} = \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{u},\mathbf{v}\rangle|^2}{\|\mathbf{u}\|^2},$$

which is the inequality. Equality holds iff $\mathbf{w}=\mathbf{0}$, that is, iff $\mathbf{v}$ is a multiple of $\mathbf{u}$. ∎

This proof is the orthogonal decomposition of Theorem 1.18, now in complex form. Applied to functions it gives $\left|\int f^{*}g\right|^2 \le \int|f|^2\int|g|^2$, the inequality behind the Heisenberg uncertainty principle (Chapter 15).

9.3.1 ORTHONORMAL BASES

Definition 9.13 (Orthonormal set). Vectors $\mathbf{e}_1, \dots, \mathbf{e}_k$ are *orthonormal* if $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$.

Theorem 9.14 (Expansion in an orthonormal basis). *If $\{\mathbf{e}_i\}$ is an orthonormal basis, then for every $\mathbf{v}$,*

$$\mathbf{v} = \sum_i \langle \mathbf{e}_i, \mathbf{v} \rangle \, \mathbf{e}_i, \qquad \|\mathbf{v}\|^2 = \sum_i |\langle \mathbf{e}_i, \mathbf{v} \rangle|^2, \qquad \langle \mathbf{u}, \mathbf{v} \rangle = \sum_i \langle \mathbf{u}, \mathbf{e}_i \rangle \langle \mathbf{e}_i, \mathbf{v} \rangle.$$

Proof. Write $\mathbf{v} = \sum c_j \mathbf{e}_j$ and take the inner product with $\mathbf{e}_i$: $\langle \mathbf{e}_i, \mathbf{v} \rangle = \sum c_j \delta_{ij} = c_i$. The other identities follow by expanding both vectors and using orthonormality. $\blacksquare$

The coefficients are simply inner products, with no system of equations to solve. The second identity, *Parseval's identity*, is the Pythagorean theorem in n dimensions. In quantum mechanics it says that the probabilities of the possible outcomes of a measurement sum to one.

Theorem 9.15 (Gram–Schmidt). *Given linearly independent $\mathbf{v}_1, \dots, \mathbf{v}_k$, define recursively*

$$\mathbf{w}_j = \mathbf{v}_j - \sum_{i<j} \langle \mathbf{e}_i, \mathbf{v}_j \rangle \, \mathbf{e}_i, \qquad \mathbf{e}_j = \frac{\mathbf{w}_j}{\|\mathbf{w}_j\|}.$$

Then $\mathbf{e}_1, \dots, \mathbf{e}_k$ are orthonormal, and for each j they span the same subspace as $\mathbf{v}_1, \dots, \mathbf{v}_j$.

Proof. By induction. $\mathbf{w}_j$ is $\mathbf{v}_j$ minus its projection onto the span of the earlier $\mathbf{e}_i$, so it is orthogonal to each of them. It is nonzero because $\mathbf{v}_j$ is not in the span of $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}$. $\blacksquare$

Example 9.16 (Legendre polynomials). Apply Gram–Schmidt to $1, x, x^2$ with $\langle f, g \rangle = \int_{-1}^{1} f g \, dx$. Here $\langle 1, 1 \rangle = 2$, so $e_1 = 1/\sqrt{2}$. Next $\langle 1, x \rangle = 0$ and $\langle x, x \rangle = \tfrac{2}{3}$, so $e_2 = \sqrt{3/2}\, x$. Finally, $w_3 = x^2 - \langle e_1, x^2 \rangle e_1 - \langle e_2, x^2 \rangle e_2 = x^2 - \tfrac{1}{3}$, with $\|w_3\|^2 = \tfrac{8}{45}$. Up to normalization these are $P_0 = 1$, $P_1 = x$, and $P_2 = \tfrac{1}{2}(3x^2 - 1)$, the *Legendre polynomials*, standardized so that $P_n(1) = 1$. They reappear as the θ-dependence of the hydrogen wavefunctions (Chapter 16).

Definition 9.17 (Orthogonal projection). If W is a subspace with orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_k$, the *orthogonal projection* onto W is $P\mathbf{v} = \sum_{i=1}^{k} \langle \mathbf{e}_i, \mathbf{v} \rangle \mathbf{e}_i$.

$P\mathbf{v}$ is the point of W closest to $\mathbf{v}$, and $\mathbf{v} - P\mathbf{v}$ is orthogonal to W. The projection satisfies $P^2 = P$. Least squares (Chapter 4) is orthogonal projection onto a column space.

9.4 Eigenvalues and Eigenvectors

Definition 9.18 (Eigenvalue, eigenvector). A scalar λ is an *eigenvalue* of an operator A if $A\mathbf{v} = \lambda\mathbf{v}$ for some nonzero $\mathbf{v}$, an *eigenvector*. The *eigenspace* of λ is $\ker(A - \lambda I)$. For a matrix, the eigenvalues are the roots of the *characteristic polynomial* $p(\lambda) = \det(A - \lambda I)$.

Over $\mathbb{C}$, the fundamental theorem of algebra guarantees that every operator on a finite-dimensional space has at least one eigenvalue. Over $\mathbb{R}$ this can fail. A rotation of the plane by 90°, $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, moves every real vector off its own line, and its eigenvalues are $\pm i$. This is one more reason quantum mechanics is set over $\mathbb{C}$.

Theorem 9.19 (Trace and determinant). *If A is $n \times n$ with eigenvalues $\lambda_1, \dots, \lambda_n$ (with multiplicity), then $\operatorname{tr} A = \sum \lambda_i$ and $\det A = \prod \lambda_i$.*

Proof. Factor $\det(A - \lambda I) = \prod(\lambda_i - \lambda)$ over $\mathbb{C}$. Setting $\lambda = 0$ gives the determinant. Comparing the coefficients of λ^{n-1} gives the trace, since only the product of diagonal entries contributes to that coefficient. ∎

Theorem 9.20 (Diagonalization). *Eigenvectors belonging to distinct eigenvalues are linearly independent. If A has n independent eigenvectors, collected as the columns of S, then $S^{-1}AS = \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.*

Proof. Suppose $\sum_{i=1}^{k} c_i \mathbf{v}_i = \mathbf{0}$ is a shortest nontrivial dependence among eigenvectors with distinct eigenvalues. Apply $A - \lambda_k I$: $\sum_{i<k} c_i(\lambda_i - \lambda_k)\mathbf{v}_i = \mathbf{0}$ is a shorter one, and it is nontrivial unless all $c_i = 0$ for $i < k$, which would force $c_k \mathbf{v}_k = \mathbf{0}$. This contradiction proves independence. For the second claim, $AS = S\Lambda$ column by column. ∎

Example 9.21 (Diagonalizing a symmetric matrix). $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ has $p(\lambda) = (2-\lambda)^2 - 1 = (\lambda - 1)(\lambda - 3)$. The eigenvectors are $\frac{1}{\sqrt{2}}\langle 1, -1\rangle$ for $\lambda = 1$ and $\frac{1}{\sqrt{2}}\langle 1, 1\rangle$ for $\lambda = 3$, and they are orthogonal. With these as columns, S is a rotation and $S^T A S = \operatorname{diag}(1, 3)$. Powers are now easy: $A^n = S \operatorname{diag}(1, 3^n) S^T$.

A matrix need not be diagonalizable. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has the single eigenvalue 1 with a one-dimensional eigenspace. Such matrices are called *defective*. The point of the next section is that the matrices of quantum mechanics are never defective.

9.5 Hermitian and Unitary Operators

Definition 9.22 (Adjoint). The *adjoint* of an operator A on an inner-product space is the operator $A^\dagger$ satisfying

$$\langle \mathbf{u}, A\mathbf{v}\rangle = \langle A^\dagger \mathbf{u}, \mathbf{v}\rangle \qquad \text{for all } \mathbf{u}, \mathbf{v}.$$

For a matrix with the standard inner product on $\mathbb{C}^n$, $A^\dagger = (A^T)^*$ is the conjugate transpose.

Adjoints satisfy $(AB)^\dagger = B^\dagger A^\dagger$, $(cA)^\dagger = c^* A^\dagger$, and $A^{\dagger\dagger} = A$.

Definition 9.23 (Hermitian, unitary, normal). An operator is *Hermitian* (or *self-adjoint*) if $A^\dagger = A$, *unitary* if $U^\dagger U = U U^\dagger = I$, and *normal* if $AA^\dagger = A^\dagger A$. Real Hermitian matrices are *symmetric*, and real unitary matrices are *orthogonal*.

Example 9.24 (The Pauli matrices).

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are Hermitian and unitary, and each squares to I. They satisfy $\sigma_x \sigma_y = i\sigma_z$ and cyclic permutations, so $\sigma_x \sigma_y - \sigma_y \sigma_x = 2i\sigma_z$. In Chapter 17 they represent the spin of an electron.

Theorem 9.25 (Hermitian operators). *Let A be Hermitian. Then*

(a) *every eigenvalue of A is real;*

(b) *eigenvectors with distinct eigenvalues are orthogonal;*

(c) $\langle \mathbf{v}, A\mathbf{v} \rangle$ *is real for every* $\mathbf{v}$.

Proof. (c) $\langle \mathbf{v}, A\mathbf{v} \rangle = \langle A\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, A\mathbf{v} \rangle^*$. (a) If $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $\lambda\|\mathbf{v}\|^2 = \langle \mathbf{v}, A\mathbf{v} \rangle$ is real by (c), so λ is real. (b) If $A\mathbf{u} = \mu\mathbf{u}$ and $A\mathbf{v} = \lambda\mathbf{v}$ with $\mu \neq \lambda$, then $\lambda\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, A\mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = \mu^*\langle \mathbf{u}, \mathbf{v} \rangle = \mu\langle \mathbf{u}, \mathbf{v} \rangle$, using (a). Since $\lambda \neq \mu$, $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. ∎

This theorem is why observables in quantum mechanics are represented by Hermitian operators. Measurement outcomes are real numbers, and outcomes that can be told apart with certainty must correspond to orthogonal states.

Theorem 9.26 (Unitary operators). *U is unitary iff it preserves inner products, $\langle U\mathbf{u}, U\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$ for all $\mathbf{u}, \mathbf{v}$. A unitary matrix has orthonormal columns, its eigenvalues have modulus 1, and $|\det U| = 1$.*

Proof. $\langle U\mathbf{u}, U\mathbf{v} \rangle = \langle \mathbf{u}, U^\dagger U\mathbf{v} \rangle$, which equals $\langle \mathbf{u}, \mathbf{v} \rangle$ for all vectors iff $U^\dagger U = I$. The (i, j) entry of $U^\dagger U$ is the inner product of columns i and j. If $U\mathbf{v} = \lambda\mathbf{v}$, then $\|\mathbf{v}\| = \|U\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$. ∎

Unitary operators are the symmetries of a complex inner-product space, the analogs of rotations. A change from one orthonormal basis to another is unitary. In quantum mechanics, time evolution is unitary, which is why total probability is conserved.

9.5.1 THE SPECTRAL THEOREM

Theorem 9.27 (Spectral theorem). *Let A be a Hermitian operator on a finite-dimensional complex inner-product space V. Then V has an orthonormal basis consisting of eigenvectors of A, and all the eigenvalues are real. In matrix form, $A = U\Lambda U^\dagger$ with U unitary and Λ real diagonal. A real symmetric matrix can be diagonalized this way with U real orthogonal.*

Proof. Induct on $n = \dim V$. Over $\mathbb{C}$, A has an eigenvalue λ_1, which is real by Theorem 9.25, with a unit eigenvector $\mathbf{e}_1$. Let $W = \{\mathbf{w} : \langle \mathbf{e}_1, \mathbf{w} \rangle = 0\}$, a subspace

of dimension $n-1$. Then W is *invariant* under A: for $\mathbf{w} \in W$,

$$\langle \mathbf{e}_1, A\mathbf{w} \rangle = \langle A\mathbf{e}_1, \mathbf{w} \rangle = \lambda_1 \langle \mathbf{e}_1, \mathbf{w} \rangle = 0.$$

The restriction of A to W is still Hermitian, so by induction W has an orthonormal basis of eigenvectors $\mathbf{e}_2, \ldots, \mathbf{e}_n$. Together with $\mathbf{e}_1$ these form the required basis. For real symmetric A, the eigenvalues are real and every eigenspace $\ker(A - \lambda I)$ is defined by real equations, so real eigenvectors can be chosen throughout. ■

The argument uses only one property of A: that the orthogonal complement of an eigenvector is invariant. The same holds for every normal operator, so normal operators, including unitary ones, are also unitarily diagonalizable. The cost is that their eigenvalues need not be real.

Corollary 9.28 (Spectral decomposition). *A Hermitian A with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ can be written $A = \sum_j \lambda_j P_j$, where P_j is the orthogonal projection onto the λ_j-eigenspace. The projections satisfy $P_i P_j = \delta_{ij} P_j$ and $\sum_j P_j = I$.*

9.5.2 COMMUTING OPERATORS

Definition 9.29 (Commutator). The *commutator* of operators A and B is $[A, B] = AB - BA$.

Commutators are bilinear and antisymmetric, and they satisfy $[A, BC] = [A, B]C + B[A, C]$ and the Jacobi identity $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.

Theorem 9.30 (Simultaneous diagonalization). *Hermitian operators A and B on a finite-dimensional space have a common orthonormal basis of eigenvectors if and only if $[A, B] = 0$.*

Proof. If both are diagonal in one basis, they commute, because diagonal matrices commute. Conversely, suppose $[A, B] = 0$, and let V_λ be an eigenspace of A. For $\mathbf{v} \in V_\lambda$, $A(B\mathbf{v}) = BA\mathbf{v} = \lambda B\mathbf{v}$, so B maps V_λ into itself. Restricted to V_λ, B is still Hermitian, so by the spectral theorem V_λ has an orthonormal basis of eigenvectors of B. Each of those is also an eigenvector of A, with eigenvalue λ. Do this in every eigenspace of A; the eigenspaces are mutually orthogonal, so the union of these bases is the required basis. ■

LOOKING AHEAD. In quantum mechanics, observables whose operators commute are *compatible*. They can be measured simultaneously with definite values, and states can be labeled by all their eigenvalues at once. A hydrogen state is labeled $|n, \ell, m\rangle$ precisely because $\hat{H}$, $\hat{L}^2$, and $\hat{L}_z$ commute. Observables that do not commute, such as position and momentum with $[\hat{x}, \hat{p}] = i\hbar$, cannot share a basis of eigenstates. That is the origin of the uncertainty principle.

9.5.3 FUNCTIONS OF OPERATORS

If $A = U\Lambda U^{\dagger}$, define $f(A) = U f(\Lambda) U^{\dagger}$, where $f(\Lambda) = \operatorname{diag}(f(\lambda_i))$. For polynomials this agrees with ordinary substitution, since $A^n = U\Lambda^n U^{\dagger}$. For the exponential it agrees with the series $e^{A} = \sum A^n/n!$, which converges for every square matrix.

Theorem 9.31 (Exponentials of Hermitian operators are unitary). *If H is Hermitian and t is real, then $U(t) = e^{-iHt}$ is unitary, $U(t+s) = U(t)U(s)$, and $U'(t) = -iHU(t)$.*

Proof. With $H = V\Lambda V^{\dagger}$, $U(t) = V \operatorname{diag}(e^{-i\lambda_j t}) V^{\dagger}$. The diagonal entries have modulus 1 because λ_j is real. The group property and the derivative are checked on the diagonal. ∎

Example 9.32 (Rotation generated by a Pauli matrix). Since $\sigma_y^2 = I$, the exponential series splits into even and odd parts, exactly as in the proof of Euler's formula:

$$e^{-i\theta\sigma_y/2} = \cos\tfrac{\theta}{2}\, I - i \sin\tfrac{\theta}{2}\, \sigma_y = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}.$$

The half-angles are not a typo. They mean that rotating a spin-$\tfrac{1}{2}$ state by 2π multiplies it by -1, a fact that has been confirmed by neutron interferometry (Chapter 17).

The Schrödinger equation $i\hbar\,\partial_t \psi = \hat{H}\psi$ has solution $\psi(t) = e^{-i\hat{H}t/\hbar}\psi(0)$. Theorem 9.31 is the reason probability is conserved.

9.6 Dirac Notation

Physicists write vectors and inner products in a notation designed for quantum mechanics.

- A vector $\mathbf{v}$ is written $|v\rangle$, a *ket*.
- The linear functional $\mathbf{w} \mapsto \langle \mathbf{u}, \mathbf{w} \rangle$ is written $\langle u|$, a *bra*. In $\mathbb{C}^n$ it is the row vector $\mathbf{u}^\dagger$.
- Their *bracket* is the inner product: $\langle u|v\rangle = \langle \mathbf{u}, \mathbf{v} \rangle$. Also $\langle u|A|v\rangle = \langle \mathbf{u}, A\mathbf{v} \rangle$, a *matrix element*.
- The *outer product* $|u\rangle\langle w|$ is the operator $|v\rangle \mapsto |u\rangle\langle w|v\rangle$. In $\mathbb{C}^n$ it is the matrix $\mathbf{u}\mathbf{w}^\dagger$.

The notation makes many identities automatic. If $\{|n\rangle\}$ is an orthonormal basis, then $|n\rangle\langle n|$ is the projection onto $|n\rangle$, and Theorem 9.14 becomes

$$\sum_n |n\rangle\langle n| = I \qquad \text{(completeness, or resolution of the identity).} \tag{9.1}$$

Inserting (9.1) anywhere is a legal move. For example, $|\psi\rangle = \sum_n |n\rangle\langle n|\psi\rangle$ expands a state in a basis, and $\langle \phi|\psi\rangle = \sum_n \langle \phi|n\rangle\langle n|\psi\rangle$ computes an inner product from components. The spectral decomposition of a Hermitian operator with eigenbasis $\{|n\rangle\}$ is

$$A = \sum_n \lambda_n |n\rangle\langle n|, \qquad f(A) = \sum_n f(\lambda_n)|n\rangle\langle n|.$$

Matrix elements in a basis are $A_{mn} = \langle m|A|n\rangle$, and the adjoint satisfies $\langle m|A^\dagger|n\rangle = \langle n|A|m\rangle^*$.

Example 9.33 (A spin measurement in Dirac notation). In $\mathbb{C}^2$ let $|\uparrow\rangle = \langle 1, 0\rangle$ and $|\downarrow\rangle = \langle 0, 1\rangle$, the eigenvectors of σ_z. The eigenvectors of σ_x are $|\pm\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle \pm |\downarrow\rangle)$, with eigenvalues ± 1. For the state $|\psi\rangle = \frac{3}{5}|\uparrow\rangle + \frac{4}{5}|\downarrow\rangle$,

$$\langle +|\psi\rangle = \tfrac{1}{\sqrt{2}}\left(\tfrac{3}{5} + \tfrac{4}{5}\right) = \tfrac{7}{5\sqrt{2}}, \qquad \langle -|\psi\rangle = \tfrac{1}{\sqrt{2}}\left(\tfrac{3}{5} - \tfrac{4}{5}\right) = -\tfrac{1}{5\sqrt{2}}.$$

By the Born rule, a measurement of σ_x gives $+1$ with probability $\frac{49}{50}$ and -1 with probability $\frac{1}{50}$. The expected value is $\langle \psi|\sigma_x|\psi\rangle = \frac{49}{50} - \frac{1}{50} = \frac{24}{25}$, which agrees with direct matrix multiplication: $2 \cdot \frac{3}{5} \cdot \frac{4}{5} = \frac{24}{25}$.

9.7 Infinite Dimensions: Spaces of Functions

The state of a particle on a line is a function $\psi(x)$, and the relevant space is infinite-dimensional.

Definition 9.34 (The space L^2). $L^2(a,b)$ is the space of complex-valued functions on (a,b) with $\int_a^b |f|^2 dx < \infty$, with inner product $\langle f|g\rangle = \int_a^b f^* g\, dx$. Two functions that differ only on a set of zero length are regarded as equal.

L^2 is a Hilbert space, a fact proved in measure theory: every Cauchy sequence converges in it. Most of the finite-dimensional theory carries over, but with care. Bases become infinite, and sums become series that converge in norm. One example of an orthonormal basis of $L^2(-\pi,\pi)$ is $\{e_n(x) = e^{inx}/\sqrt{2\pi} : n \in \mathbb{Z}\}$:

$$\langle e_m|e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)x}dx = \delta_{mn}.$$

That this set is *complete*, meaning every L^2 function is the norm limit of its expansion, is the theorem of Chapter 11. The expansion coefficients $\langle e_n|f\rangle$ are the Fourier coefficients of f.

Theorem 9.35 (Momentum is Hermitian). *On differentiable functions in $L^2(\mathbb{R})$ that vanish at $\pm\infty$, or that are periodic on $[-\pi,\pi]$, the operator $\hat{p} = -i\hbar\, \mathrm{d}/\mathrm{d}x$ satisfies $\langle f|\hat{p}g\rangle = \langle \hat{p}f|g\rangle$.*

Proof. Integrate by parts:

$$\langle f|\hat{p}g\rangle = \int f^*(-i\hbar g')\,dx = \big[-i\hbar f^* g\big] + \int i\hbar f'^* g\,dx = \int (-i\hbar f')^* g\,dx = \langle \hat{p}f|g\rangle.$$

The boundary term vanishes under either hypothesis. The factor $-i$ is essential: d/dx by itself is *anti*-Hermitian, and multiplying by i makes it Hermitian. ∎

The eigenfunctions of $\hat{p}$ are e^{ikx}, with eigenvalue $\hbar k$, and they are not in $L^2(\mathbb{R})$: $\int |e^{ikx}|^2 dx = \infty$. Hermitian operators in infinite dimensions can have *continuous spectra* whose "eigenvectors" are not normalizable. The position operator $\hat{x}$ (multiplication by x) has eigenfunctions $\delta(x - x_0)$, which are not even functions. Making this rigorous requires the spectral theorem for unbounded self-adjoint operators, a major result of twentieth-century analysis due to von Neumann. For our purposes the formal rules suffice, with completeness written as $\int |x\rangle\langle x|\,dx = I$ and orthonormality as $\langle x|x'\rangle = \delta(x - x')$.

COMMON ERROR (Hermitian on which domain?). Whether an operator is Hermitian depends on its domain. On $[0,1]$ with no boundary conditions, the boundary term $-i\hbar[f^*g]_0^1$ need not vanish, and $-i\hbar\, d/dx$ is not Hermitian. Boundary conditions are part of the definition of a quantum observable. This is why the energy levels of a particle in a box (Chapter 14) depend on what happens at its walls.

9.8 Worked Problems

The problems below use the spectral theorem as a computational tool: for powers and exponentials of matrices, for degenerate eigenvalues, for rotations and projections, and for commuting observables. The last two move into the complex and infinite-dimensional settings of quantum mechanics.

Problem 9.1 Powers and exponentials through projectors

Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Find its spectral decomposition $A = \sum \lambda_i P_i$, and use it to compute A^{10} and e^{At}.

Solution. A is real symmetric, so by Theorem 9.27 it has an orthonormal eigenbasis. Its eigenvalues have sum $\operatorname{tr} A = 4$ and product $\det A = 3$, so they are 3 and 1, with unit eigenvectors $\frac{1}{\sqrt{2}}\langle 1,1\rangle$ and $\frac{1}{\sqrt{2}}\langle 1,-1\rangle$. The projectors onto these eigenspaces are

$$P_+ = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad P_- = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \qquad A = 3P_+ + P_-.$$

They satisfy $P_\pm^2 = P_\pm$, $P_+P_- = 0$, and $P_+ + P_- = I$. These relations make any polynomial or power series in A act on the eigenvalues alone: $f(A) = f(3)P_+ + f(1)P_-$ (Corollary 9.28). Hence

$$A^{10} = 3^{10}P_+ + P_- = \begin{pmatrix} 29525 & 29524 \\ 29524 & 29525 \end{pmatrix}, \qquad e^{At} = \frac{1}{2}\begin{pmatrix} e^{3t}+e^t & e^{3t}-e^t \\ e^{3t}-e^t & e^{3t}+e^t \end{pmatrix}.$$

Computing A^{10} by repeated multiplication would take four squarings and a product. The spectral form gives every power at once, and shows that for large n, $A^n \approx 3^n P_+$: repeated application of A projects any vector onto the dominant eigenvector. That is the *power method* for computing eigenvectors numerically.

Problem 9.2 A degenerate eigenvalue

Find the spectral decomposition of $B = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, including an orthonormal eigenbasis. In what sense is the eigenbasis not unique?

Solution. Write $B = I + J$, where J is the all-ones matrix. Since $J\langle 1,1,1\rangle = 3\langle 1,1,1\rangle$ and $J\mathbf{v} = \mathbf{0}$ for every $\mathbf{v} \perp \langle 1,1,1\rangle$, the eigenvalues of J are $3, 0, 0$, and those of B are $4, 1, 1$. The eigenvalue 1 is degenerate, with a two-dimensional eigenspace: the plane $x + y + z = 0$.

The projector onto the eigenspace for 4 is $P_4 = \hat{\mathbf{u}}\hat{\mathbf{u}}^T$ with $\hat{\mathbf{u}} = \frac{1}{\sqrt{3}}\langle 1,1,1\rangle$, which is $\frac{1}{3}J$. The projector onto the plane is $P_1 = I - \frac{1}{3}J$. So

$$B = 4 \cdot \tfrac{1}{3}J + 1 \cdot \left(I - \tfrac{1}{3}J\right).$$

For an orthonormal eigenbasis, any orthonormal basis of the plane will do, for instance $\frac{1}{\sqrt{2}}\langle 1,-1,0\rangle$ and $\frac{1}{\sqrt{6}}\langle 1,1,-2\rangle$ (Gram–Schmidt applied to any two independent vectors in the plane). This is the non-uniqueness: within a degenerate eigenspace, every orthonormal basis is an eigenbasis. The projectors, however, are unique. In quantum mechanics, the projectors determine the probabilities of measurement outcomes, and the freedom to rotate within a degenerate eigenspace is resolved only by measuring a second, compatible observable, as in Problem 9.5.

Problem 9.3 The axis and angle of a rotation

Show that $R = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ is a rotation of $\mathbb{R}^3$, and find its axis and angle.

Solution. The columns of R are $\hat{\jmath}, \hat{\mathbf{k}}, \hat{\imath}$, an orthonormal set, so $R^TR = I$. And $\det R = +1$ (one cyclic permutation of the columns of I, which is two transpositions). So $R \in SO(3)$ is a rotation.

The axis is the eigenvector with eigenvalue 1: $R\langle x,y,z\rangle = \langle z,x,y\rangle$, which equals $\langle x,y,z\rangle$ only when $x = y = z$. So the axis is along $\langle 1,1,1\rangle$. For the angle, a rotation by θ about any axis has, in an orthonormal basis adapted to the axis, the matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, with trace $1 + 2\cos\theta$. The trace is basis-independent (Theorem 9.19), and $\operatorname{tr} R = 0$. So $\cos\theta = -\frac{1}{2}$, and $\theta = \frac{2\pi}{3}$.

The sense of rotation: $\mathbf{v} = \langle 1, -1, 0\rangle$ is perpendicular to the axis, and $R\mathbf{v} = \langle 0, 1, -1\rangle$. Since $\mathbf{v} \times R\mathbf{v} = \langle 1, 1, 1\rangle$ points along the axis, R rotates counterclockwise by 120° when viewed from the tip of $\langle 1, 1, 1\rangle$. Consistently, $R^3 = I$, and the other eigenvalues are $e^{\pm 2\pi i/3}$, with complex eigenvectors. A real rotation always has eigenvalues $1, e^{i\theta}, e^{-i\theta}$.

Problem 9.4 Projection onto a plane

Find the matrix of the orthogonal projection P of $\mathbb{R}^3$ onto the plane $x + y + z = 0$, and verify its algebraic characterization. Find the reflection across the plane.

Solution. Let $\hat{\mathbf{n}} = \frac{1}{\sqrt{3}}\langle 1, 1, 1\rangle$. Projecting onto the plane removes the component along $\hat{\mathbf{n}}$: $P\mathbf{v} = \mathbf{v} - (\hat{\mathbf{n}} \cdot \mathbf{v})\hat{\mathbf{n}}$, so

$$P = I - \hat{\mathbf{n}}\hat{\mathbf{n}}^T = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

This is P_1 from Problem 9.2, as it must be. Checks: $P^T = P$, and $P^2 = I - 2\hat{\mathbf{n}}\hat{\mathbf{n}}^T + \hat{\mathbf{n}}(\hat{\mathbf{n}}^T\hat{\mathbf{n}})\hat{\mathbf{n}}^T = I - \hat{\mathbf{n}}\hat{\mathbf{n}}^T = P$, since $\hat{\mathbf{n}}^T\hat{\mathbf{n}} = 1$. The eigenvalues are $1, 1, 0$: vectors in the plane are unchanged, and $\hat{\mathbf{n}}$ is annihilated. The trace, 2, equals the dimension of the plane, as it does for every orthogonal projection, since the trace is the sum of the eigenvalues.

The reflection keeps the plane and reverses $\hat{\mathbf{n}}$, so $S = I - 2\hat{\mathbf{n}}\hat{\mathbf{n}}^T = 2P - I$. Then $S^2 = 4P^2 - 4P + I = I$, $S^T = S$, and $\det S = -1$. A reflection is orthogonal and symmetric, with eigenvalues $1, 1, -1$. In quantum computing the same construction appears as the *Householder* or *Grover* reflection $I - 2|\psi\rangle\langle\psi|$.

Problem 9.5 Commuting observables

Let $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. Show that A and B commute, that A has a degenerate eigenvalue, and find an orthonormal basis of simultaneous eigenvectors. Which pairs of eigenvalues (a, b) occur?

Solution. Direct multiplication gives $AB = BA = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$.

The eigenvalues of A: $\mathbf{e}_1$ has eigenvalue 2, and the lower block $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has eigenvalues 2 and 0, with eigenvectors $\langle 0, 1, 1\rangle$ and $\langle 0, 1, -1\rangle$. So $a = 2$ is doubly de-

generate, with eigenspace spanned by $\mathbf{e}_1$ and $\mathbf{f} = \frac{1}{\sqrt{2}}\langle 0,1,1\rangle$, and $a = 0$ is simple, with eigenvector $\mathbf{g} = \frac{1}{\sqrt{2}}\langle 0,1,-1\rangle$.

By Theorem 9.30, B maps each eigenspace of A into itself. Indeed $B\mathbf{e}_1 = \langle 0,1,1\rangle = \sqrt{2}\,\mathbf{f}$ and $B\mathbf{f} = \frac{1}{\sqrt{2}}\langle 2,0,0\rangle = \sqrt{2}\,\mathbf{e}_1$. Restricted to the degenerate eigenspace, in the basis $\mathbf{e}_1, \mathbf{f}$, B is $\begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}$, with eigenvalues $\pm\sqrt{2}$ and eigenvectors $\frac{1}{\sqrt{2}}(\mathbf{e}_1 \pm \mathbf{f})$. Also $B\mathbf{g} = \mathbf{0}$. The simultaneous eigenbasis is

$$\tfrac{1}{2}\left\langle \sqrt{2},1,1\right\rangle \ (a=2,\ b=\sqrt{2}), \qquad \tfrac{1}{2}\left\langle \sqrt{2},-1,-1\right\rangle \ (a=2,\ b=-\sqrt{2}),$$
$$\tfrac{1}{\sqrt{2}}\langle 0,1,-1\rangle \ (a=0,\ b=0).$$

The pairs (a,b) are all distinct, so they label the basis vectors uniquely: $\{A, B\}$ is a *complete set of commuting observables*. Measuring A alone cannot distinguish the first two states. Measuring B as well can. The quantum numbers (n,ℓ,m) of hydrogen in Chapter 16 arise exactly this way, from the commuting operators $\hat{H}$, $\hat{L}^2$, $\hat{L}_z$.

Problem 9.6 The determinant of an exponential

Prove that $\det e^A = e^{\operatorname{tr} A}$ for every complex square matrix A. Conclude that e^A is always invertible, and that if H is Hermitian and traceless then e^{iH} is unitary with determinant 1.

Solution. *Diagonalizable case.* If $A = SDS^{-1}$ with $D = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$, then $e^A = Se^DS^{-1}$, since $(SDS^{-1})^k = SD^kS^{-1}$ term by term in the exponential series. So $\det e^A = \det e^D = \prod e^{\lambda_i} = e^{\sum \lambda_i} = e^{\operatorname{tr} A}$.

General case. Diagonalizable matrices are dense: perturbing the entries of A by an arbitrarily small amount makes its eigenvalues distinct, and a matrix with distinct eigenvalues is diagonalizable. Both sides of the identity are continuous functions of the entries of A. An identity between continuous functions that holds on a dense set holds everywhere.

Since $e^{\operatorname{tr} A} \neq 0$, e^A is invertible, with inverse e^{-A}. If H is Hermitian, e^{iH} is unitary (Theorem 9.31), and if also $\operatorname{tr} H = 0$ then $\det e^{iH} = e^{i\cdot 0} = 1$. Example: for $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$, the eigenvalues are 1 and 3, $\det e^A = e \cdot e^3 = e^4 = e^{\operatorname{tr} A}$. The traceless Hermitian matrices are the real combinations of the Pauli matrices, and their exponentials form the group $SU(2)$ that describes spin rotations in Chapter 17.

Problem 9.7 Diagonalizable is not the same as unitarily diagonalizable

Let $N = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Show that N is diagonalizable but that its eigenvectors are not orthogonal, and that N does not commute with $N^\dagger$. Prove that a matrix is unitarily diagonalizable only if it is *normal*, $[A, A^\dagger] = 0$.

Solution. N is upper triangular, so its eigenvalues are its diagonal entries, 1 and 2. They are distinct, so N is diagonalizable. The eigenvectors are $\langle 1, 0\rangle$ for $\lambda = 1$ and $\langle 1, 1\rangle$ for $\lambda = 2$ (from $(N - 2I)\mathbf{v} = \mathbf{0}$, that is, $-v_1 + v_2 = 0$). They make an angle of 45°, so no orthonormal eigenbasis exists, and $N = SDS^{-1}$ only with a non-unitary $S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Directly,

$$NN^\dagger = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}, \qquad N^\dagger N = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}, \qquad [N, N^\dagger] = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \neq 0.$$

For the theorem, suppose $A = UDU^\dagger$ with U unitary and D diagonal. Then $A^\dagger = UD^*U^\dagger$, and $AA^\dagger = UDD^*U^\dagger = UD^*DU^\dagger = A^\dagger A$, since diagonal matrices commute. So A is normal. (The converse also holds: every normal matrix is unitarily diagonalizable, which generalizes Theorem 9.27 from Hermitian and unitary matrices to all normal ones.)

The distinction matters in physics. Observables are Hermitian, hence normal, so their eigenstates for distinct eigenvalues are orthogonal and measurement outcomes are mutually exclusive. Non-normal matrices arise for open and driven systems, such as lasers and fluid flows, and their non-orthogonal eigenvectors cause transient growth: even when all eigenvalues signal decay, $\|e^{Nt}\mathbf{x}\|$ can grow for a while before it decays, an effect invisible in the spectrum alone.

Problem 9.8 The generator of dilations

On smooth, rapidly decreasing functions in $L^2(\mathbb{R})$, consider $\hat{D} = x\frac{\mathrm{d}}{\mathrm{d}x}$. Show that $\hat{D}$ is not Hermitian, but that $\hat{G} = -i\big(x\frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{2}\big)$ is. Show that $\hat{G}$ generates dilations: $e^{-i\alpha\hat{G}}f(x) = e^{-\alpha/2}f(e^{-\alpha}x)$.

Solution. Integrate by parts. The boundary terms vanish because the functions decrease rapidly:

$$\langle f|\hat{D}g\rangle = \int \bar{f}\, x g'\, dx = -\int (x\bar{f})' g\, dx = -\int \overline{(f + xf')}\, g\, dx = \langle -(1 + \hat{D})f \,|\, g\rangle.$$

So $\hat{D}^\dagger = -1 - \hat{D} \neq \hat{D}$. Then $\big(\hat{D} + \frac{1}{2}\big)^\dagger = -1 - \hat{D} + \frac{1}{2} = -\big(\hat{D} + \frac{1}{2}\big)$: the operator $\hat{D} + \frac{1}{2}$ is anti-Hermitian, and multiplying by $-i$ makes it Hermitian, since $(-iK)^\dagger = iK^\dagger = -iK$ when $K^\dagger = -K$.

For the second part, let $f_\alpha(x) = e^{-\alpha/2} f(e^{-\alpha}x)$. Differentiating in α,

$$\frac{\partial f_\alpha}{\partial \alpha} = -\frac{1}{2} f_\alpha - e^{-\alpha/2} e^{-\alpha} x f'(e^{-\alpha}x) = -\Big(\frac{1}{2} + x\frac{\partial}{\partial x}\Big) f_\alpha = -i\hat{G} f_\alpha.$$

This is the equation solved by $e^{-i\alpha\hat{G}} f$, with the same initial value f at $\alpha = 0$, so the two agree. The prefactor $e^{-\alpha/2}$ preserves the norm, $\int |f_\alpha|^2 dx = \int |f|^2 dx$, as a unitary operator must. Compare $\hat{p} = -i\hbar\, d/dx$, which generates translations (Chapter 15). The $\frac{1}{2}$ is the price of Hermiticity, the same symmetrization that makes $\frac{1}{2}(\hat{x}\hat{p} + \hat{p}\hat{x})$, which is $\hbar\hat{G}$, the Hermitian version of the classical product xp.

EXERCISES

A. COMPUTATION

9.1 Determine whether $\langle 1,2,3\rangle$, $\langle 4,5,6\rangle$, $\langle 2,1,0\rangle$ are linearly independent. Find the dimension of their span.

9.2 Find a basis for the subspace of $\mathbb{R}^4$ defined by $x_1 + x_2 - x_3 = 0$ and $x_2 + x_4 = 0$.

9.3 Show that $1,\ 1+x,\ 1+x+x^2$ is a basis for $\mathcal{P}_2$, and find the coordinates of $3 - 2x + 5x^2$ in it.

9.4 Find the matrix of $T(p) = p' + 2p$ on $\mathcal{P}_2$ in the basis $1, x, x^2$. Is T invertible?

9.5 Find the kernel and image of $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{pmatrix}$ and verify rank–nullity.

9.6 Compute $\det \begin{pmatrix} 2 & 0 & 1 \\ 1 & 3 & 2 \\ 1 & 1 & 2 \end{pmatrix}$ and the inverse of the matrix.

9.7 For $\mathbf{u} = \langle 1+i,\ 2\rangle$ and $\mathbf{v} = \langle 3,\ 1-i\rangle$ in $\mathbb{C}^2$, compute $\langle \mathbf{u}, \mathbf{v}\rangle$, $\langle \mathbf{v}, \mathbf{u}\rangle$, $\|\mathbf{u}\|$, and $\|\mathbf{v}\|$.

9.8 Normalize $\mathbf{v} = \langle 1,\ i,\ 1-i\rangle$ in $\mathbb{C}^3$.

9.9 Apply Gram–Schmidt to $\langle 1,1,0\rangle$, $\langle 1,0,1\rangle$, $\langle 0,1,1\rangle$ in $\mathbb{R}^3$.

9.10 Apply Gram–Schmidt to $1, x, x^2$ on $[0,1]$ with $\langle f, g\rangle = \int_0^1 f g\, dx$.

9.11 Find the orthogonal projection of $f(x) = x^3$ onto $\mathcal{P}_1$ in $L^2(-1,1)$, the best linear approximation in the mean-square sense.

9.12 Find the eigenvalues and eigenvectors of $\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$ and diagonalize it.

9.13 Find the eigenvalues and eigenvectors of $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{pmatrix}$.

9.14 Find the eigenvalues of $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ and their eigenvectors in $\mathbb{C}^2$.

9.15 Show that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable.

9.16 Find a unitary U and real diagonal Λ with $A = U\Lambda U^\dagger$ for $A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$.

9.17 Find the eigenvalues and normalized eigenvectors of σ_y and of $\sigma_x + \sigma_z$.

9.18 Verify $\sigma_x\sigma_y = i\sigma_z$ and compute $[\sigma_y, \sigma_z]$ and $\{\sigma_x, \sigma_y\} = \sigma_x\sigma_y + \sigma_y\sigma_x$.

9.19 Compute e^{At} for $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and interpret.

9.20 Compute e^{A} for $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ by writing $A = I + N$ with $N^2 = 0$.

9.21 For $\mathbf{n}$ a unit vector and $\boldsymbol{\sigma}\cdot\mathbf{n} = n_x\sigma_x + n_y\sigma_y + n_z\sigma_z$, show that $(\boldsymbol{\sigma}\cdot\mathbf{n})^2 = I$ and find its eigenvalues.

9.22 In Dirac notation with $|\psi\rangle = \frac{1}{\sqrt{3}}|\uparrow\rangle + i\sqrt{\frac{2}{3}}|\downarrow\rangle$, find the probabilities of each outcome when σ_x, σ_y, and σ_z are measured, and the three expectation values.

9.23 Write the projector $P = |+\rangle\langle+|$ as a 2×2 matrix and verify $P^2 = P$ and $P^\dagger = P$.

9.24 Verify that $\{\cos x/\sqrt{\pi},\ \sin x/\sqrt{\pi},\ \cos 2x/\sqrt{\pi}\}$ is orthonormal in $L^2(-\pi,\pi)$.

9.25 Compute $\langle f|g\rangle$, $\|f\|$, and $\|g\|$ for $f = e^{ix}$ and $g = x$ on $[-\pi,\pi]$.

B. THEORY AND PROOF

9.26 Prove that the intersection of two subspaces is a subspace, but the union generally is not.

9.27 Prove that $(AB)^\dagger = B^\dagger A^\dagger$, and that the product of two Hermitian operators is Hermitian iff they commute.

9.28 Prove the *polarization identity* for an inner product that is antilinear in its first slot: $\langle \mathbf{u}, \mathbf{v}\rangle = \frac{1}{4}\sum_{k=0}^{3} i^{-k}\,\|\mathbf{u} + i^k \mathbf{v}\|^2$. It shows that the norm determines the inner product.

9.29 Prove that an orthogonal projection P satisfies $P^2 = P = P^\dagger$, and conversely that any operator with $P^2 = P = P^\dagger$ is the orthogonal projection onto its image.

9.30 Prove that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, and deduce that $[A, B] = I$ is impossible for finite matrices. (So $[\hat{x}, \hat{p}] = i\hbar$ requires infinite dimensions.)

9.31 Prove that a Hermitian operator is positive definite iff all its eigenvalues are positive.

9.32 Prove that every operator can be written uniquely as $A = H_1 + iH_2$ with H_1, H_2 Hermitian.

9.33 Prove that if H is Hermitian then e^{iH} is unitary, and that every unitary matrix has this form.

9.34 Prove that the eigenvalues of an anti-Hermitian operator ($A^\dagger = -A$) are purely imaginary.

9.35 Prove that $e^A e^B = e^{A+B}$ when $[A, B] = 0$, and give a counterexample when $[A, B] \neq 0$.

C. CHALLENGE

9.36 (Uncertainty for matrices.) For Hermitian A, B and a unit vector $\mathbf{v}$, let $\Delta A^2 = \langle \mathbf{v}, A^2\mathbf{v}\rangle - \langle \mathbf{v}, A\mathbf{v}\rangle^2$. Prove $\Delta A\,\Delta B \geq \frac{1}{2}|\langle \mathbf{v}, [A, B]\mathbf{v}\rangle|$. Check it for $A = \sigma_x$, $B = \sigma_y$, $\mathbf{v} = |\uparrow\rangle$.

9.37 (Hermite functions as eigenvectors.) Show that $\hat{a} = \frac{1}{\sqrt{2}}\big(x + \mathrm{d}/\mathrm{d}x\,\big)$ and $\hat{a}^\dagger = \frac{1}{\sqrt{2}}\big(x - \mathrm{d}/\mathrm{d}x\,\big)$ are adjoint to each other on $L^2(\mathbb{R})$, that $[\hat{a}, \hat{a}^\dagger] = 1$, and that $\hat{a}\, e^{-x^2/2} = 0$.

9.38 Show that the matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has no square root, but $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has infinitely many.

9.39 (Density matrices.) For a unit vector $|\psi\rangle$, let $\rho = |\psi\rangle\langle\psi|$. Show $\operatorname{tr}\rho = 1$, $\rho^2 = \rho$, and $\langle\psi|A|\psi\rangle = \operatorname{tr}(\rho A)$. A *mixed state* is a convex combination $\rho = \sum p_i|\psi_i\rangle\langle\psi_i|$; show $\operatorname{tr}\rho^2 \leq 1$ with equality iff ρ is a pure state.

CHAPTER 10

Ordinary and Partial Differential Equations

LEARNING OBJECTIVES

- Solve separable and linear first-order equations, and state the existence and uniqueness theorem.
- Describe the solution space of a linear second-order equation, and solve constant-coefficient equations in all three cases.
- Analyze free, damped, and driven oscillators, including resonance, using complex exponentials.
- Solve linear systems $\mathbf{x}' = A\mathbf{x}$ by eigenvectors and find the normal modes of coupled oscillators.
- Solve equations by power series, and derive the Hermite and Legendre polynomials from series that terminate.
- Develop Sturm–Liouville theory: real eigenvalues, orthogonal eigenfunctions, and expansions.
- Solve the heat, wave, and Laplace equations by separation of variables.

Newton's second law is a differential equation, and so is the Schrödinger equation. In the time-independent form that determines the energy levels of an atom,

$$-\frac{\hbar^2}{2m}\psi''(x) + V(x)\psi(x) = E\psi(x),$$

it is a linear second-order ordinary differential equation with an unknown parameter E. This chapter develops the theory that equation requires. That includes the remarkable mechanism by which demanding a well-behaved solution forces E to take discrete values. Quantization in quantum mechanics is, mathematically, an eigenvalue problem for a differential equation.

10.1 First-Order Equations

An *ordinary differential equation* (ODE) relates an unknown function $y(x)$ to its derivatives. Its *order* is the highest derivative that appears. An *initial value problem* adds a condition such as $y(x_0) = y_0$.

Theorem 10.1 (Existence and uniqueness (Picard–Lindelöf)). *If $f(x,y)$ and $\partial f/\partial y$ are continuous on a rectangle containing (x_0,y_0), then the problem $y' = f(x,y)$, $y(x_0) = y_0$ has a unique solution on some interval about x_0. The same holds for systems $\mathbf{y}' = \mathbf{f}(x,\mathbf{y})$, and hence for equations of any order rewritten as first-order systems.*

The proof constructs the solution as the limit of the iteration $y_{n+1}(x) = y_0 + \int_{x_0}^{x} f(t, y_n(t))\,dt$ and shows that the iteration is a contraction. Uniqueness is what makes physics deterministic: the present state fixes the future.

10.1.1 Separable Equations

If $y' = g(x)h(y)$, then separating the variables and integrating gives $\int \frac{dy}{h(y)} = \int g(x)\,dx$.

Example 10.2 (Exponential growth and decay). $N' = -\lambda N$ gives $\int dN/N = -\lambda \int dt$, so $N(t) = N_0 e^{-\lambda t}$. The *half-life* is $\tau_{1/2} = \ln 2/\lambda$. In Chapter 13 we will see that an unstable quantum state decays exactly this way, and that its energy is then uncertain by an amount of order $\hbar\lambda$.

10.1.2 Linear Equations

Theorem 10.3 (Integrating factor). *The solution of $y' + p(x)y = q(x)$ is*

$$y = \frac{1}{\mu(x)}\Big(\int \mu(x) q(x)\,dx + C\Big), \qquad \mu(x) = e^{\int p(x)\,dx}.$$

Proof. Multiplying by μ turns the left side into $(\mu y)'$, because $\mu' = p\mu$. Integrate. ∎

Example 10.4 (An RC circuit). A resistor R and capacitor C in series with a source $V_0 \cos \omega t$ obey $RQ' + Q/C = V_0 \cos \omega t$. With $\tau = RC$, the steady

state is found most easily with complex exponentials. Try $Q = \mathrm{Re}(Ae^{i\omega t})$: then $(i\omega R + 1/C)A = V_0$, so $A = \dfrac{CV_0}{1 + i\omega\tau}$. The charge oscillates with amplitude $CV_0/\sqrt{1+\omega^2\tau^2}$ and lags the source by the phase $\arctan(\omega\tau)$. The general solution adds the transient $Ke^{-t/\tau}$.

10.2 Second-Order Linear Equations

Definition 10.5 (Linear second-order equation). $y'' + p(x)y' + q(x)y = r(x)$ is *homogeneous* if $r \equiv 0$. The *Wronskian* of two functions is $W(y_1, y_2) = y_1 y_2' - y_1' y_2$.

Theorem 10.6 (Structure of solutions). *Let p, q be continuous on an interval I.*

(a) *The solutions of the homogeneous equation form a two-dimensional vector space.*

(b) *Two solutions y_1, y_2 form a basis iff $W(y_1, y_2) \neq 0$, in which case W is nonzero everywhere on I.*

(c) *Every solution of the inhomogeneous equation is $y = y_p + c_1 y_1 + c_2 y_2$, where y_p is any particular solution.*

Proof. (a) Linear combinations of solutions are solutions, so the solutions form a vector space. The map $y \mapsto (y(x_0), y'(x_0)) \in \mathbb{R}^2$ is linear. By Theorem 10.1 it is a bijection: every initial pair has exactly one solution. So the dimension is 2. (b) Differentiating W and using the equation gives Abel's identity $W' = -pW$, so $W(x) = W(x_0)e^{-\int p}$, which is either always zero or never zero. And $W(x_0) \neq 0$ iff the vectors $(y_i(x_0), y_i'(x_0))$ are independent. (c) The difference of two solutions of the inhomogeneous equation solves the homogeneous one. ∎

10.2.1 Constant Coefficients

For $ay'' + by' + cy = 0$, try $y = e^{rx}$. Substituting gives the *characteristic equation* $ar^2 + br + c = 0$.

Theorem 10.7 (Constant-coefficient solutions). *Let r_1, r_2 be the roots of $ar^2 + br + c = 0$.*

(a) *Distinct real roots:* $y = c_1 e^{r_1 x} + c_2 e^{r_2 x}$.

(b) *A repeated root* r: $y = (c_1 + c_2 x) e^{rx}$.

(c) *Complex roots* $\alpha \pm i\beta$: $y = e^{\alpha x}(c_1 \cos \beta x + c_2 \sin \beta x)$.

Proof. In each case the two functions are solutions with nonzero Wronskian. In (c), $e^{(\alpha \pm i\beta)x}$ are complex solutions, and their real and imaginary parts are real solutions by linearity, using Euler's formula. In (b), xe^{rx} is a solution because $(D - r)^2 (xe^{rx}) = 0$, where $D = d/dx$. ■

10.2.2 OSCILLATORS

The *harmonic oscillator* $m\ddot{x} = -kx$, or $\ddot{x} + \omega_0^2 x = 0$ with $\omega_0 = \sqrt{k/m}$, has solutions $x = A\cos(\omega_0 t + \varphi)$. Every system near a stable equilibrium is approximately a harmonic oscillator. If $U(x)$ has a minimum at x_0, then $U \approx U(x_0) + \frac{1}{2}U''(x_0)(x - x_0)^2$, so $k = U''(x_0)$. This is why the quantum harmonic oscillator (Chapter 14) describes molecular vibrations, phonons in solids, and every mode of the electromagnetic field.

Adding friction proportional to velocity gives $\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0$, with roots $r = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2}$:

- *underdamped* ($\gamma < \omega_0$): $x = Ae^{-\gamma t}\cos(\omega_d t + \varphi)$ with $\omega_d = \sqrt{\omega_0^2 - \gamma^2}$;
- *critically damped* ($\gamma = \omega_0$): $x = (c_1 + c_2 t)e^{-\gamma t}$, the fastest return to equilibrium without oscillation;
- *overdamped* ($\gamma > \omega_0$): a sum of two decaying exponentials.

Example 10.8 (Driven oscillator and resonance). Solve $\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = F_0 \cos \omega t$ in steady state.

Solution. Try $x = \mathrm{Re}(Ae^{i\omega t})$: $(\omega_0^2 - \omega^2 + 2i\gamma\omega)A = F_0$. So

$$|A| = \frac{F_0}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\gamma^2\omega^2}}, \qquad \tan\delta = \frac{2\gamma\omega}{\omega_0^2 - \omega^2},$$

where x lags the force by δ. For light damping, $|A|$ peaks sharply near $\omega = \omega_0$, with peak height of order $F_0/(2\gamma\omega_0)$ and width of order γ. Near resonance $|A|^2$ is approximately a *Lorentzian* proportional to $1/[(\omega - \omega_0)^2 + \gamma^2]$, the same line shape as in Example 8.35.

10.2.3 VARIATION OF PARAMETERS

Theorem 10.9. *If y_1, y_2 are independent solutions of $y'' + py' + qy = 0$ with Wronskian W, a particular solution of $y'' + py' + qy = r$ is*

$$y_p = -y_1 \int \frac{y_2 r}{W}\,dx + y_2 \int \frac{y_1 r}{W}\,dx.$$

Proof. Seek $y_p = u_1 y_1 + u_2 y_2$ subject to $u_1' y_1 + u_2' y_2 = 0$. Substituting into the equation gives $u_1' y_1' + u_2' y_2' = r$. Solve this linear system by Cramer's rule. ∎

Example 10.10. For $y'' + y = \sec x$: $y_1 = \cos x$, $y_2 = \sin x$, $W = 1$, so $y_p = -\cos x \int \sin x \sec x\,dx + \sin x \int 1\,dx = \cos x \ln|\cos x| + x \sin x$.

10.3 Linear Systems and Normal Modes

A system $\mathbf{x}' = A\mathbf{x}$ with constant matrix A has the solution $\mathbf{x}(t) = e^{At}\mathbf{x}(0)$. If A has an eigenbasis $A\mathbf{v}_k = \lambda_k \mathbf{v}_k$, then

$$\mathbf{x}(t) = \sum_k c_k e^{\lambda_k t}\mathbf{v}_k,$$

with the c_k the coordinates of $\mathbf{x}(0)$ in the eigenbasis. Each eigenvector evolves independently: the eigenbasis *decouples* the system.

Example 10.11 (Coupled oscillators). Two equal masses m on a line are joined to walls by springs of constant k and to each other by a spring of constant k'. With displacements x_1, x_2,

$$m\ddot{\mathbf{x}} = -K\mathbf{x}, \qquad K = \begin{pmatrix} k + k' & -k' \\ -k' & k + k' \end{pmatrix}.$$

Try $\mathbf{x} = \mathbf{v}\cos \omega t$: then $K\mathbf{v} = m\omega^2 \mathbf{v}$, an eigenvalue problem. K is symmetric with eigenvalues k, for $\mathbf{v} = \langle 1, 1\rangle$, and $k + 2k'$, for $\mathbf{v} = \langle 1, -1\rangle$. These are the *normal modes*. In the first the masses move together with $\omega_1 = \sqrt{k/m}$, and the middle spring is never stretched. In the second they move oppositely with $\omega_2 = \sqrt{(k + 2k')/m}$. Every motion is a superposition of the two. If one mass starts displaced and the other at rest, the energy sloshes back and forth between them at the *beat frequency* $\omega_2 - \omega_1$.

The spectral theorem guarantees that this works for any number of masses. K is symmetric, so there is an orthonormal basis of normal modes with real frequencies. In the limit of infinitely many masses, the normal modes become the standing waves of a string (Section 10.6). In quantum field theory, each normal mode becomes a quantum harmonic oscillator.

10.4 Series Solutions and Special Functions

When the coefficients are not constant, there is usually no closed-form solution. If p and q are analytic at x_0, an *ordinary point*, then the solutions are analytic there too, and can be found as power series $y = \sum a_n (x - x_0)^n$. Substituting gives a *recurrence relation* for the coefficients.

10.4.1 HERMITE'S EQUATION

The equation

$$y'' - 2xy' + 2\nu y = 0 \tag{10.1}$$

arises from the quantum harmonic oscillator, as we show in Chapter 14. Substituting $y = \sum a_n x^n$:

$$\sum_n \left[(n+2)(n+1)a_{n+2} - 2n a_n + 2\nu a_n\right] x^n = 0 \quad \Longrightarrow \quad a_{n+2} = \frac{2(n-\nu)}{(n+1)(n+2)}\, a_n.$$

The even coefficients are determined by a_0 and the odd ones by a_1, which gives two independent solutions.

Theorem 10.12 (Termination and growth). *If ν is a nonnegative integer n, one of the two series terminates, giving a polynomial of degree n. Otherwise both series are infinite, and their terms behave for large index like those of e^{x^2}, so the solutions grow like e^{x^2}.*

Proof. If $\nu = n$, then $a_{n+2} = 0$, and all subsequent coefficients of that parity vanish. If the series does not terminate, the ratio $a_{k+2}/a_k \approx 2/k$ for large k, which matches the series $e^{x^2} = \sum x^{2j}/j!$, whose ratio of successive coefficients is $1/(j+1) = 2/k$ with $k = 2j$. ∎

The polynomial solutions, normalized so that the leading coefficient is 2^n, are the *Hermite polynomials*:

$$H_0 = 1, \quad H_1 = 2x, \quad H_2 = 4x^2 - 2, \quad H_3 = 8x^3 - 12x, \quad H_4 = 16x^4 - 48x^2 + 12.$$

They are generated by Rodrigues' formula $H_n(x) = (-1)^n e^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}e^{-x^2}$, and they are orthogonal with weight e^{-x^2}:

$$\int_{-\infty}^{\infty} H_m H_n e^{-x^2}\,dx = 2^n n!\sqrt{\pi}\,\delta_{mn}.$$

Theorem 10.12 is the mathematical heart of energy quantization in the harmonic oscillator. In Chapter 14 the wavefunction is $\psi = y(x)e^{-x^2/2}$. If y grew like e^{x^2}, then $\psi \sim e^{x^2/2}$ could not be normalized. Only the terminating case is allowed, which forces ν to be an integer, and that forces the energy to be $E = (\nu + \tfrac{1}{2})\hbar\omega$.

10.4.2 LEGENDRE'S EQUATION

The equation

$$(1-x^2)y'' - 2xy' + \ell(\ell+1)y = 0 \tag{10.2}$$

arises when Laplace's equation is separated in spherical coordinates, with $x = \cos\theta$. The series method gives

$$a_{n+2} = \frac{n(n+1)-\ell(\ell+1)}{(n+1)(n+2)}\,a_n = \frac{(n-\ell)(n+\ell+1)}{(n+1)(n+2)}\,a_n.$$

For ℓ a nonnegative integer, one solution terminates. These are the *Legendre polynomials* P_ℓ, normalized by $P_\ell(1) = 1$:

$$P_0 = 1, \quad P_1 = x, \quad P_2 = \tfrac{1}{2}(3x^2-1),$$
$$P_3 = \tfrac{1}{2}(5x^3-3x), \quad P_4 = \tfrac{1}{8}(35x^4-30x^2+3).$$

These agree with the Gram–Schmidt computation of Example 9.16. A nonterminating solution has ratio $a_{n+2}/a_n \to 1$, and it diverges logarithmically at $x = \pm 1$, which are the poles $\theta = 0, \pi$ of the sphere. Requiring a solution finite on the whole sphere therefore forces ℓ to be an integer. This is the origin of the quantization of angular momentum (Chapter 16). Rodrigues' formula is $P_\ell(x) = \dfrac{1}{2^\ell \ell!}\dfrac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}(x^2-1)^\ell$.

10.4.3 REGULAR SINGULAR POINTS AND THE FROBENIUS METHOD

At a *regular singular point*, where $(x-x_0)p$ and $(x-x_0)^2 q$ are analytic, one tries $y = (x-x_0)^s \sum a_n (x-x_0)^n$. The lowest power gives the *indicial equation* for s. Bessel's equation $x^2y'' + xy' + (x^2-\nu^2)y = 0$ has indicial equation $s^2 = \nu^2$, and the root $s = \nu$ gives the Bessel function

$$J_\nu(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\,\Gamma(k+\nu+1)}\left(\frac{x}{2}\right)^{2k+\nu}.$$

Bessel functions describe the vibrations of a drum, waves in cylinders, and, in their spherical form $j_\ell(x)$, the radial wavefunctions of a free particle in three dimensions.

10.5 Sturm–Liouville Theory

The Hermite and Legendre polynomials are both orthogonal families, and both arise as the polynomial solutions of a differential equation with a parameter. This is no coincidence. Both equations are eigenvalue problems for Hermitian differential operators, and Theorem 9.25 applies to them.

Definition 10.13 (Sturm–Liouville problem). A *Sturm–Liouville problem* on $[a,b]$ is

$$-\big(p(x)y'\big)' + q(x)y = \lambda\, w(x)\, y, \tag{10.3}$$

with $p>0$, $w>0$ on (a,b), and boundary conditions that make the boundary term $\big[p\,(f^{*}g' - f'^{*}g)\big]_a^b$ vanish for all admissible f and g. Typical conditions are $y(a)=y(b)=0$ (*Dirichlet*), $y'(a)=y'(b)=0$ (*Neumann*), periodicity, or, when p vanishes at an endpoint, mere boundedness there. Write $L = -\frac{d}{dx}p\frac{d}{dx} + q$.

Theorem 10.14 (Sturm–Liouville). *Under these conditions:*

(a) *L is Hermitian:* $\int_a^b f^{*}(Lg)\,dx = \int_a^b (Lf)^{*} g\,dx$;

(b) *the eigenvalues λ are real;*

(c) *eigenfunctions with distinct eigenvalues are orthogonal with weight w:* $\int_a^b y_m^{*} y_n\, w\,dx = 0$ *for* $\lambda_m \neq \lambda_n$.

For a regular problem on a finite interval, the eigenvalues form an increasing sequence $\lambda_1 < \lambda_2 < \cdots \to \infty$, the nth eigenfunction has exactly $n-1$ zeros in (a,b), and the eigenfunctions form a complete orthogonal basis of the weighted space $L^2_w(a,b)$.

Proof. (a) Integrate by parts twice: $\int f^{*}(-(pg')')\,dx = -[pf^{*}g'] + \int pf'^{*}g'\,dx = -[pf^{*}g'] + [pf'^{*}g] - \int (pf'^{*})'g\,dx$. The boundary terms are exactly those assumed to vanish. (b) and (c) then follow as in Theorem 9.25, with the inner product $\int f^{*}g\,w\,dx$. If $Ly=\lambda w y$, then $\lambda\int |y|^2 w = \int y^{*}Ly$ is real, and for two eigenfunctions $(\lambda_n - \lambda_m^{*})\int y_m^{*}y_n w = \int y_m^{*}Ly_n - \int (Ly_m)^{*}y_n = 0$. The statements

about ordering, zeros, and completeness require more analysis, and we omit their proofs. ■

Example 10.15 (The simplest case). $-y'' = \lambda y$ on $[0, L]$ with $y(0) = y(L) = 0$. For $\lambda = k^2 > 0$, $y = \sin kx$ satisfies $y(L) = 0$ iff $kL = n\pi$. Checking $\lambda \le 0$ gives only the zero solution. So the eigenvalues are $\lambda_n = (n\pi/L)^2$ for $n = 1, 2, \dots$, with eigenfunctions $\sin(n\pi x/L)$. These are orthogonal on $[0, L]$, and $\sin(n\pi x/L)$ has $n - 1$ interior zeros. The time-independent Schrödinger equation for a particle in a box is exactly this problem, and $E_n = \hbar^2 \lambda_n / 2m$ (Chapter 14).

Example 10.16 (Legendre and Hermite in Sturm–Liouville form). Legendre's equation (10.2) is $-\big((1 - x^2)y'\big)' = \ell(\ell + 1)y$ on $[-1, 1]$, with $p = 1 - x^2$ vanishing at the endpoints and $w = 1$. Boundedness at ± 1 is the boundary condition, so $\int_{-1}^{1} P_m P_n \, dx = 0$ for $m \neq n$. Multiplying Hermite's equation by e^{-x^2} gives $-\big(e^{-x^2} y'\big)' = 2\nu e^{-x^2} y$, so $p = w = e^{-x^2}$ on $\mathbb{R}$, and the Hermite polynomials are orthogonal with weight e^{-x^2}.

LOOKING AHEAD. Theorem 10.14 is a special case of the principle that organizes all of quantum mechanics. The Hamiltonian $\hat{H} = -\frac{\hbar^2}{2m} \, \mathrm{d}^2/\mathrm{d}x^2 + V(x)$ is a Sturm–Liouville operator with $p = \frac{\hbar^2}{2m}$, $q = V$, and $w = 1$. So energies are real, stationary states with different energies are orthogonal, and the stationary states form a complete basis in which any state can be expanded. The theorem that the nth eigenfunction has $n - 1$ nodes explains the node counting of atomic orbitals.

10.6 Partial Differential Equations: Separation of Variables

10.6.1 THE HEAT EQUATION

Consider a rod of length L whose ends are held at temperature zero:

$$u_t = k u_{xx}, \qquad u(0, t) = u(L, t) = 0, \qquad u(x, 0) = f(x).$$

Seek solutions of the form $u = X(x)T(t)$. Substituting and dividing by XT,

$$\frac{T'}{kT} = \frac{X''}{X} = -\lambda.$$

The left side depends only on t and the middle only on x, so both equal a constant, which we call $-\lambda$. The spatial problem $-X'' = \lambda X$, $X(0) = X(L) = 0$, is Example 10.15, giving $X_n = \sin(n\pi x/L)$ and $\lambda_n = (n\pi/L)^2$. Then $T_n = e^{-k\lambda_n t}$. By linearity,

$$u(x,t) = \sum_{n=1}^{\infty} b_n\, e^{-k(n\pi/L)^2 t} \sin\frac{n\pi x}{L}, \qquad b_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}\,dx, \tag{10.4}$$

where the b_n come from orthogonality at $t = 0$. Higher modes decay faster, so the temperature profile smooths out over time, and for large t it is dominated by the $n = 1$ mode.

Example 10.17 (Heat flow from a uniform start). For $f(x) = u_0$ on $[0, \pi]$ with $k = 1$: $b_n = \frac{2u_0}{\pi}\int_0^\pi \sin nx\,dx = \frac{2u_0}{\pi n}(1 - (-1)^n)$, which is $\frac{4u_0}{\pi n}$ for odd n and 0 for even n. So $u = \frac{4u_0}{\pi}\sum_{n\text{ odd}} \frac{e^{-n^2 t}\sin nx}{n}$.

10.6.2 THE WAVE EQUATION

A string of length L fixed at both ends obeys $u_{tt} = c^2 u_{xx}$. Separation gives the same spatial modes, with temporal factors $\cos\omega_n t$ and $\sin\omega_n t$, where $\omega_n = n\pi c/L$:

$$u(x,t) = \sum_{n=1}^{\infty} \left(a_n\cos\omega_n t + b_n \sin\omega_n t\right)\sin\frac{n\pi x}{L}.$$

These are the *harmonics* of the string, with frequencies that are integer multiples of the fundamental $\omega_1 = \pi c/L$. That is why a plucked string sounds a definite pitch. A drumhead, whose modes are Bessel functions with irregularly spaced frequencies, does not.

10.6.3 LAPLACE'S EQUATION IN SPHERICAL COORDINATES

Seek solutions of $\nabla^2 u = 0$ of the form $u = R(r)\Theta(\theta)$, independent of ϕ. By Corollary 7.29,

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) = -\frac{1}{\Theta\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\right) = \ell(\ell+1).$$

With $x = \cos\theta$ the angular equation becomes Legendre's equation (10.2), so $\Theta = P_\ell(\cos\theta)$ for integer ℓ. The radial equation $(r^2R')' = \ell(\ell+1)R$ is equidimensional, and $R = r^s$ gives $s = \ell$ or $s = -(\ell+1)$. So

$$u(r,\theta) = \sum_{\ell=0}^{\infty}\left(A_\ell r^\ell + \frac{B_\ell}{r^{\ell+1}}\right)P_\ell(\cos\theta).$$

The B_ℓ terms are the multipole potentials of electrostatics: monopole $1/r$, dipole $\cos\theta/r^2$, quadrupole, and so on.

Example 10.18 (A sphere with prescribed surface potential). Find the potential inside a sphere of radius a whose surface is held at $V_0\cos^2\theta$.

Solution. Inside, finiteness at $r=0$ forces $B_\ell=0$. Write $\cos^2\theta = x^2 = \frac{1}{3}P_0+\frac{2}{3}P_2$. Matching at $r=a$: $u=\frac{V_0}{3}+\frac{2V_0}{3}\frac{r^2}{a^2}P_2(\cos\theta)=\frac{V_0}{3}+\frac{V_0 r^2}{3a^2}(3\cos^2\theta-1)$.

LOOKING AHEAD. The hydrogen atom is solved by exactly this separation, applied to the Schrödinger equation rather than to Laplace's equation, and allowing dependence on ϕ. The angular equation produces the *associated* Legendre functions and the spherical harmonics Y_ℓ^m. The radial equation, which now includes the Coulomb potential and the energy, produces Laguerre polynomials, and the requirement that the series terminate yields the Bohr energy levels $E_n=-13.6\,\text{eV}/n^2$.

10.7 Worked Problems

The problems below run from nonlinear first-order equations through forced and defective linear systems to the Frobenius method, an eigenvalue problem that must be solved numerically, and two boundary-value problems for partial differential equations.

Problem 10.1 The logistic equation

Solve $y'=ry\left(1-\frac{y}{K}\right)$ with $y(0)=y_0$, where $r,K>0$ and $0<y_0<K$. Show that $y\to K$, and that the growth rate is largest when $y=K/2$.

Solution. The equation is separable. Using partial fractions, $\dfrac{K}{y(K-y)}=\dfrac{1}{y}+\dfrac{1}{K-y}$, so

$$\int\left(\frac{1}{y}+\frac{1}{K-y}\right)dy=\int r\,dt \quad\Longrightarrow\quad \ln\frac{y}{K-y}=rt+c.$$

At $t=0$, $c=\ln\frac{y_0}{K-y_0}$. Solving for y,

$$y(t)=\frac{K}{1+Ae^{-rt}},\qquad A=\frac{K-y_0}{y_0}>0.$$

As $t \to \infty$, $e^{-rt} \to 0$ and $y \to K$, the *carrying capacity*. The growth rate $y' = ry(1 - y/K)$ is a downward parabola in y, maximal at $y = K/2$, where it equals $rK/4$. That happens at $t^* = \frac{1}{r}\ln A$, the inflection point of the S-shaped curve. The constant solutions $y = 0$ and $y = K$ are equilibria: 0 is unstable (small populations grow) and K is stable. Uniqueness (Theorem 10.1) guarantees that a solution starting between them stays between them forever, since it cannot cross an equilibrium solution.

Problem 10.2 A Bernoulli equation

Solve $xy' + y = x^2y^2$ for $x > 0$.

Solution. Divide by x: $y' + \frac{1}{x}y = xy^2$. This is a *Bernoulli equation* $y' + p(x)y = q(x)y^n$ with $n = 2$. The substitution $v = y^{1-n} = 1/y$ makes it linear: $v' = -y'/y^2$, so dividing the equation by $-y^2$ gives

$$v' - \frac{1}{x}v = -x.$$

The integrating factor (Theorem 10.3) is $e^{-\int dx/x} = \frac{1}{x}$, and $\left(\frac{v}{x}\right)' = -1$. So $\frac{v}{x} = -x + C$, $v = Cx - x^2$, and

$$y = \frac{1}{Cx - x^2},$$

together with the solution $y \equiv 0$, which was lost when dividing by y^2. Check: $y' = -\frac{C-2x}{(Cx-x^2)^2}$, so $xy' + y = \frac{-Cx+2x^2+Cx-x^2}{(Cx-x^2)^2} = \frac{x^2}{(Cx-x^2)^2} = x^2y^2$. Each nonzero solution blows up at $x = C$ (if $C > 0$), a feature of nonlinear equations: the interval of existence depends on the initial data.

Problem 10.3 Transient and steady state

Solve $\ddot{x} + 2\dot{x} + 5x = 10\cos t$ with $x(0) = 0$, $\dot{x}(0) = 0$. Identify the transient and the steady state, and write the steady state as $A\cos(t - \delta)$.

Solution. *Homogeneous part.* The characteristic equation $s^2 + 2s + 5 = 0$ has roots $s = -1 \pm 2i$, so $x_h = e^{-t}(c_1\cos 2t + c_2\sin 2t)$.

Particular solution. Try $x_p = a\cos t + b\sin t$. Then $\ddot{x}_p + 2\dot{x}_p + 5x_p = (4a + 2b)\cos t + (4b - 2a)\sin t$, so $4a + 2b = 10$ and $4b - 2a = 0$. Hence $a = 2b$, $10b = 10$, so $b = 1$ and $a = 2$: $x_p = 2\cos t + \sin t$.

Initial conditions. $x(0) = c_1 + 2 = 0$ gives $c_1 = -2$. $\dot{x}(0) = -c_1 + 2c_2 + 1 = 0$ gives $c_2 = -\frac{3}{2}$. So

$$x(t) = \underbrace{e^{-t}\left(-2\cos 2t - \tfrac{3}{2}\sin 2t\right)}_{\text{transient}} + \underbrace{2\cos t + \sin t}_{\text{steady state}}.$$

The transient decays like e^{-t} and oscillates at the natural damped frequency 2. The steady state oscillates at the driving frequency 1, with amplitude $A = \sqrt{2^2+1^2} = \sqrt{5}$ and phase lag $\delta = \arctan\frac{1}{2} \approx 26.6°$: $x_{ss} = \sqrt{5}\cos(t-\delta)$. In the complex method of Example 10.8, $x_{ss} = \operatorname{Re}\frac{10e^{it}}{5-1+2i} = \operatorname{Re}\frac{10e^{it}}{4+2i}$, and $\frac{10}{4+2i} = 2-i$, which gives $2\cos t + \sin t$ in one line. After a few time constants, the system forgets its initial conditions completely.

Problem 10.4 A defective system

Solve $\mathbf{x}' = A\mathbf{x}$ with $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$ and $\mathbf{x}(0) = \langle 1, 0\rangle$.

Solution. The characteristic polynomial is $\lambda^2 - 4\lambda + 4 = (\lambda-2)^2$, a double eigenvalue $\lambda = 2$. But $A - 2I = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ has rank 1, so the eigenspace is one-dimensional, spanned by $\langle 1, -1\rangle$. A is not diagonalizable, and the method of normal modes fails: it produces only one solution.

The matrix exponential still works. Write $A = 2I + N$ with $N = A - 2I$. Then $N^2 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, and since $2I$ commutes with N,

$$e^{At} = e^{2t}e^{Nt} = e^{2t}(I + tN) = e^{2t}\begin{pmatrix} 1+t & t \\ -t & 1-t \end{pmatrix}.$$

The series for e^{Nt} stops after two terms because N is nilpotent. So $\mathbf{x}(t) = e^{At}\langle 1, 0\rangle = e^{2t}\langle 1+t,\ -t\rangle$.

The factor te^{2t} is the signature of a defective eigenvalue, just as te^{rx} appears for a repeated root of a scalar equation. The scalar equation $y'' - 4y' + 4y = 0$ becomes, with $\mathbf{x} = \langle y, y'\rangle$, a system whose matrix is defective in the same way. In the phase plane, every trajectory leaves the origin tangent to the single eigenvector direction $\langle 1, -1\rangle$ and turns toward it again as $t \to \infty$.

Problem 10.5 Bessel's equation of order zero by Frobenius

Find a series solution of $x^2y'' + xy' + x^2y = 0$ about $x = 0$ with $y(0) = 1$.

Solution. The point $x = 0$ is a regular singular point, so try $y = \sum_{n\ge 0} a_n x^{n+r}$ with $a_0 \neq 0$. Substituting,

$$\sum_{n\ge 0}\big[(n+r)(n+r-1) + (n+r)\big]a_n x^{n+r} + \sum_{n\ge 0} a_n x^{n+r+2} = 0.$$

The bracket simplifies to $(n+r)^2$. The lowest power, x^r, gives the *indicial equation* $r^2 = 0$: a double root $r = 0$. The coefficient of x^{r+1} gives $(1+r)^2 a_1 = 0$, so $a_1 = 0$.

For $n \geq 2$, matching x^{n+r} gives the recurrence

$$n^2 a_n = -a_{n-2} \qquad (r = 0).$$

So all odd coefficients vanish, and $a_{2k} = -\frac{a_{2k-2}}{(2k)^2} = \frac{(-1)^k}{2^{2k}(k!)^2} a_0$. With $a_0 = 1$,

$$y = J_0(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k!)^2} \left(\frac{x}{2}\right)^{2k} = 1 - \frac{x^2}{4} + \frac{x^4}{64} - \frac{x^6}{2304} + \cdots .$$

The ratio test shows the series converges for all x. This is the *Bessel function* J_0, which describes the vibrations of a circular drum and the radial dependence of waves in a cylinder. Because the indicial root is double, Frobenius' method gives only one solution. The second solution, Y_0, contains $J_0(x)\ln x$ and is singular at the origin. In physical problems that include the axis, it is excluded by requiring finiteness.

Problem 10.6 A Sturm–Liouville problem with a Robin condition

Find the eigenvalues and eigenfunctions of $-y'' = \lambda y$ on $[0, 1]$ with $y(0) = 0$ and $y'(1) + y(1) = 0$. Compute the first two eigenvalues numerically, and find their asymptotic form.

Solution. First, $\lambda > 0$. Multiply by y and integrate by parts: $\lambda \int_0^1 y^2 = \int_0^1 y'^2 - [yy']_0^1 = \int_0^1 y'^2 + y(1)^2 > 0$, using $y(0) = 0$ and $y'(1) = -y(1)$. So write $\lambda = k^2$ with $k > 0$. The condition at 0 gives $y = \sin kx$, and the condition at 1 gives

$$k\cos k + \sin k = 0, \qquad \text{that is,} \qquad \tan k = -k.$$

This transcendental equation has one root in each interval $\left((n - \tfrac{1}{2})\pi,\, n\pi\right)$, $n = 1, 2, \ldots$, where the falling branches of $\tan k$ cross the line $-k$. Numerically (by bisection),

$$k_1 \approx 2.0288, \quad \lambda_1 \approx 4.1159; \qquad k_2 \approx 4.9132, \quad \lambda_2 \approx 24.139.$$

For large n the line $-k$ is steep relative to the spacing, and the roots approach the asymptotes of tan from above: $k_n \approx (n - \tfrac{1}{2})\pi + \frac{1}{(n-\frac{1}{2})\pi}$. Already $k_2 \approx 4.712 + 0.212 = 4.924$, close to 4.913.

The eigenfunctions $\sin k_n x$ are orthogonal on $[0, 1]$ by Theorem 10.14, although they are not the sines of a Fourier series. Robin conditions model heat loss through an end by Newton cooling, and in quantum mechanics they arise when matching a wavefunction to a decaying exponential outside a finite well (Chapter 14). There, too, the energies are roots of a transcendental equation.

Problem 10.7 Heat flow with a heated end

A rod of length 1 is initially at temperature 0. For $t > 0$ its left end is held at 0 and its right end at 1. Find the temperature $u(x,t)$, where $u_t = u_{xx}$.

Solution. Separation of variables requires homogeneous boundary conditions, so first remove the inhomogeneity. The *steady state* $U(x)$ satisfies $U'' = 0$, $U(0) = 0$, $U(1) = 1$, so $U = x$. The deviation $v = u - U$ satisfies $v_t = v_{xx}$ with $v(0,t) = v(1,t) = 0$ and $v(x,0) = -x$. By (10.4),

$$v = \sum_{n\geq 1} b_n e^{-n^2\pi^2 t}\sin n\pi x, \qquad b_n = 2\int_0^1 (-x)\sin n\pi x\,dx = \frac{2(-1)^n}{n\pi},$$

using $\int_0^1 x\sin n\pi x\,dx = \frac{(-1)^{n+1}}{n\pi}$. Therefore

$$u(x,t) = x + \frac{2}{\pi}\sum_{n\geq 1}\frac{(-1)^n}{n}e^{-n^2\pi^2 t}\sin n\pi x.$$

At $t = 0$ the series is the Fourier sine series of $-x$, so $u(x,0) = 0$ for $0 \leq x < 1$. The slowest-decaying mode has time constant $1/\pi^2 \approx 0.10$, so after $t \approx 0.5$ the rod is within 1% of the linear steady state everywhere. The method, subtracting a steady state to make the boundary conditions homogeneous, applies to any linear problem with time-independent boundary data.

Problem 10.8 Laplace's equation in a square

Solve $\nabla^2 u = 0$ in the unit square with $u = 1$ on the top edge $y = 1$ and $u = 0$ on the other three edges. Find u at the center exactly, without summing the series.

Solution. Separate $u = X(x)Y(y)$: $X''/X = -Y''/Y = -\lambda$. The homogeneous conditions $X(0) = X(1) = 0$ give $X = \sin n\pi x$ with $\lambda = n^2\pi^2$, and then $Y'' = n^2\pi^2 Y$ with $Y(0) = 0$ gives $Y = \sinh n\pi y$. Superpose, $u = \sum c_n \sin n\pi x \sinh n\pi y$, and impose $u(x,1) = 1$: $c_n \sinh n\pi$ must be the sine coefficient of 1, which is $\frac{4}{n\pi}$ for odd n and 0 for even n. So

$$u(x,y) = \frac{4}{\pi}\sum_{n\ \text{odd}}\frac{\sin n\pi x\,\sinh n\pi y}{n\,\sinh n\pi}.$$

At the center, use symmetry instead. Let u_T be this solution, and u_B, u_L, u_R the solutions with 1 on the bottom, left, or right edge, each obtained from u_T by a rotation of the square. Their sum is harmonic and equals 1 on the entire boundary (except at corners), so by uniqueness it is the constant 1. All four are equal at the center, by symmetry. Hence $u_T(\frac{1}{2},\frac{1}{2}) = \frac{1}{4}$. Summing the series numerically confirms it: the terms decay like $e^{-n\pi/2}$, and 100 terms give 0.25000 to all printed digits.

EXERCISES

A. COMPUTATION

10.1 Solve $y' = xy^2$, $y(0) = 1$, and find the interval on which the solution exists.

10.2 Solve $y' + 2xy = x$, $y(0) = 0$.

10.3 Solve $xy' + 2y = \sin x$ for $x > 0$.

10.4 A radioactive sample loses 30% of its mass in 10 years. Find its half-life.

10.5 Newton's law of cooling: $T' = -k(T - T_a)$. A cup of coffee at 90°C in a 20°C room cools to 60°C in 10 minutes. When does it reach 40°C?

10.6 Solve $y'' - 5y' + 6y = 0$, $y(0) = 1$, $y'(0) = 0$.

10.7 Solve $y'' + 4y' + 4y = 0$, $y(0) = 1$, $y'(0) = 1$.

10.8 Solve $y'' + 2y' + 5y = 0$, $y(0) = 1$, $y'(0) = -1$.

10.9 Find the general solution of $y'' - y = e^{2x}$.

10.10 Find the general solution of $y'' + 4y = \cos 2x$ and explain the term that grows with x.

10.11 Use variation of parameters to solve $y'' + y = \tan x$.

10.12 A mass of 0.5 kg on a spring with $k = 8$ N/m is released from 0.1 m with no initial velocity. Find $x(t)$, the period, and the maximum speed.

10.13 For $\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0$ with $\omega_0 = 5$, find γ for critical damping and solve with $x(0) = 1$, $\dot{x}(0) = 0$.

10.14 Find the frequency at which the steady-state amplitude of $\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = F_0 \cos \omega t$ is largest, and the maximum amplitude.

10.15 Solve $\mathbf{x}' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\mathbf{x}$ with $\mathbf{x}(0) = \langle 1, 0 \rangle$.

10.16 Solve $\mathbf{x}' = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix}\mathbf{x}$, and describe the trajectories in the phase plane.

10.17 Find the normal modes and frequencies of three equal masses connected in a line by four identical springs between two walls.

10.18 Find the power-series solution of $y'' + xy = 0$ (Airy's equation) through the x^7 term.

10.19 Find the first four nonzero terms of the two independent series solutions of $y'' - 2xy' + y = 0$.

10.20 Verify $H_3(x) = 8x^3 - 12x$ from Rodrigues' formula, and check that it satisfies (10.1) with $\nu = 3$.

10.21 Verify $\int_{-\infty}^{\infty} H_2(x)^2 e^{-x^2} dx = 8\sqrt{\pi}$.

10.22 Compute P_3 from Rodrigues' formula, and verify $\int_{-1}^{1} P_\ell^2 \, dx = \frac{2}{2\ell+1}$ for $\ell = 0, 1, 2, 3$.

10.23 Expand $f(x) = x^3$ in Legendre polynomials.

10.24 Find the eigenvalues and eigenfunctions of $-y'' = \lambda y$ on $[0, L]$ with $y'(0) = y'(L) = 0$.

10.25 Find the eigenvalues and eigenfunctions of $-y'' = \lambda y$ on $[0, 1]$ with $y(0) = 0$, $y'(1) = 0$.

10.26 Solve $u_t = u_{xx}$ on $[0, \pi]$ with $u(0,t) = u(\pi,t) = 0$ and $u(x,0) = \sin x + 3\sin 4x$.

10.27 Solve the heat equation on $[0, 1]$ with zero boundary values and $u(x,0) = x(1-x)$. Give the first two nonzero terms.

10.28 A string of length π with $c = 1$ starts from rest with displacement $u(x,0) = \sin x + \frac{1}{2}\sin 3x$. Find $u(x,t)$.

10.29 Find the potential outside a sphere of radius a held at $V_0 \cos\theta$, with $u \to 0$ at infinity.

B. THEORY AND PROOF

10.30 Prove Abel's identity $W' = -pW$.

10.31 Show that $e^{r_1 x}$ and $e^{r_2 x}$ are independent iff $r_1 \neq r_2$.

10.32 Show that the substitution $x = e^t$ converts the Euler equation $x^2y'' + axy' + by = 0$ into a constant-coefficient equation, and solve $x^2y'' - 2y = 0$.

10.33 Prove that the Hermite polynomials defined by Rodrigues' formula satisfy $H_n' = 2nH_{n-1}$ and $H_{n+1} = 2xH_n - 2nH_{n-1}$.

10.34 Prove the orthogonality of Hermite polynomials directly from Rodrigues' formula by integrating by parts.

10.35 Show that the generating function $e^{2xt-t^2}=\sum_{n\ge 0} H_n(x)t^n/n!$ and use it to compute $H_0,\dots,H_3$.

10.36 Show that the eigenvalues of a Sturm–Liouville problem with $q\ge 0$ and Dirichlet conditions are positive. (Multiply by y^* and integrate.)

10.37 Show that $u=\frac{1}{\sqrt{t}}e^{-x^2/(4kt)}$ cannot be obtained by separation of variables, yet solves the heat equation.

10.38 Derive (10.4) for the coefficients b_n using orthogonality.

C. CHALLENGE

10.39 (Quantization by shooting.) For $-y''+x^2y=Ey$, solve numerically from $x=0$ with $y(0)=1$, $y'(0)=0$ for $E=0.9, 1.0, 1.1$, and observe the behavior at large x. Explain why only $E=1$ gives a normalizable solution, and relate this to Theorem 10.12.

10.40 Show that the substitution $y=e^{-x^2/2}h(x)$ transforms $-y''+x^2y=Ey$ into Hermite's equation with $2\nu=E-1$.

10.41 The generating function of Legendre polynomials is $(1-2xt+t^2)^{-1/2}=\sum_{\ell} P_\ell(x)t^\ell$. Use it to show that $\dfrac{1}{|\mathbf{r}-\mathbf{r}'|}=\sum_\ell \dfrac{r'^\ell}{r^{\ell+1}}P_\ell(\cos\gamma)$ for $r'<r$, where γ is the angle between $\mathbf{r}$ and $\mathbf{r}'$: the multipole expansion.

10.42 Show that $J_{1/2}(x)=\sqrt{\dfrac{2}{\pi x}}\sin x$.

CHAPTER 11

Fourier Analysis

LEARNING OBJECTIVES

- Compute Fourier series in real and complex form, including half-range sine and cosine series.
- State the convergence theorems for Fourier series, explain the Gibbs phenomenon, and relate smoothness to decay of coefficients.
- Prove Bessel's inequality, state Parseval's identity, and use it to sum series exactly.
- Define the Fourier transform and its inverse, and compute transforms of Gaussians, boxes, and exponentials.
- Use the shift, scaling, derivative, and convolution properties, and solve the heat equation on the line.
- Represent the delta function as a Fourier integral and interpret it as completeness of plane waves.
- Prove the Fourier uncertainty relation $\sigma_x \sigma_k \geq \frac{1}{2}$, and analyze wave packets, phase velocity, and group velocity.

A musical chord is a sum of pure tones, and an ear decomposes sound into frequencies. Fourier's insight, made in 1807 while studying heat flow and greeted with skepticism, was that *every* reasonable function can be decomposed into sinusoids. In the language of Chapter 9, the functions e^{inx} form an orthonormal basis of a Hilbert space, and Fourier analysis is expansion in that basis. For quantum mechanics, Fourier analysis is the bridge between position and momentum. A wavefunction $\psi(x)$ and its Fourier transform $\phi(k)$ describe the same state, one in position space and one in momentum space. The mathematical fact that a function and its transform cannot both be sharply localized *is* the Heisenberg uncertainty principle.

11.1 FOURIER SERIES

Definition 11.1 (Fourier series). Let f be integrable on $[-\pi,\pi]$. Its *Fourier series* is

$$f(x)\sim\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos nx+b_n\sin nx\right),$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,dx,\qquad b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\,dx.$$

The symbol $\sim$ records that convergence is a separate question.

The formulas come from orthogonality. For integers $m,n\geq 1$,

$$\int_{-\pi}^{\pi}\cos mx\cos nx\,dx=\pi\delta_{mn},\qquad \int_{-\pi}^{\pi}\sin mx\sin nx\,dx=\pi\delta_{mn},$$

$$\int_{-\pi}^{\pi}\cos mx\sin nx\,dx=0,$$

so multiplying the series by $\cos nx$ or $\sin nx$ and integrating isolates one coefficient (Theorem 9.14). For a function of period $2L$, replace nx by $n\pi x/L$ and $\frac{1}{\pi}\int_{-\pi}^{\pi}$ by $\frac{1}{L}\int_{-L}^{L}$.

Theorem 11.2 (Complex form). *Equivalently,* $f(x)\sim\sum_{n=-\infty}^{\infty}c_ne^{inx}$ *with* $c_n=\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi}f(x)e^{-inx}dx$. *For real* f, $c_{-n}=c_n^{*}$, $c_0=a_0/2$, *and* $c_n=\frac{1}{2}(a_n-ib_n)$ *for* $n\geq 1$.

The complex form is the expansion of Chapter 9 in the orthonormal basis $e^{inx}/\sqrt{2\pi}$. It is the one physicists use.

Example 11.3 (Square wave). Let $f(x)=-1$ on $(-\pi,0)$ and $f(x)=1$ on $(0,\pi)$, extended periodically.

Solution. f is odd, so $a_n=0$, and $b_n=\frac{2}{\pi}\int_0^{\pi}\sin nx\,dx=\frac{2}{\pi n}(1-(-1)^n)$. Thus

$$f(x)\sim\frac{4}{\pi}\left(\sin x+\frac{\sin 3x}{3}+\frac{\sin 5x}{5}+\cdots\right).$$

At $x=\pi/2$ this gives $1=\frac{4}{\pi}\left(1-\frac{1}{3}+\frac{1}{5}-\cdots\right)$, which is Leibniz's series $\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\cdots$.

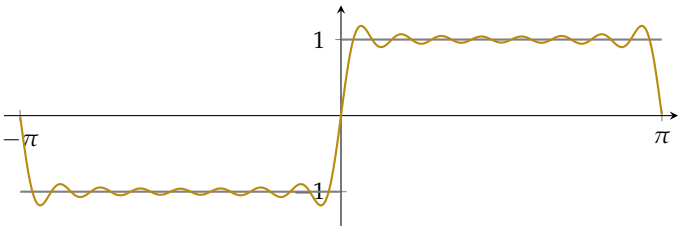


Figure 11.1: The square wave and its Fourier partial sum through $\sin 15x$. The overshoot near each jump is the Gibbs phenomenon.

Example 11.4 (The sawtooth and the parabola). For $f(x) = x$ on $(-\pi, \pi)$, which is odd: $b_n = \frac{2}{\pi}\int_0^\pi x \sin nx\,dx = \frac{2(-1)^{n+1}}{n}$, so

$$x \sim 2\Big(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \cdots\Big).$$

For $f(x) = x^2$, which is even: $a_0 = \frac{2\pi^2}{3}$ and $a_n = \frac{4(-1)^n}{n^2}$, so

$$x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}\cos nx \qquad (-\pi \le x \le \pi).$$

Setting $x = \pi$ gives $\pi^2 = \frac{\pi^2}{3} + 4\sum \frac{1}{n^2}$, so $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, Euler's solution of the Basel problem.

11.1.1 HALF-RANGE EXPANSIONS

A function given only on $[0, L]$ can be extended to $[-L, L]$ as an odd function, giving a *sine series* $\sum b_n \sin(n\pi x/L)$ with $b_n = \frac{2}{L}\int_0^L f \sin(n\pi x/L)\,dx$, or as an even function, giving a *cosine series*. The sine series is the expansion in the Dirichlet eigenfunctions of Example 10.15, and it is what solved the heat equation in (10.4).

11.2 CONVERGENCE AND PARSEVAL'S IDENTITY

Theorem 11.5 (Dirichlet's theorem). *If f is 2π-periodic and piecewise smooth, its Fourier series converges at every x to $\frac{1}{2}\big[f(x^+) + f(x^-)\big]$. In particular it converges to $f(x)$ wherever f is continuous. If f is also continuous everywhere, the*

convergence is uniform.

At a jump, the series converges to the midpoint. The square wave's series gives 0 at $x = 0$. Near a jump, however, the partial sums overshoot by about 9% of the jump, and the overshoot does not shrink as more terms are added. It only moves closer to the jump. This is the *Gibbs phenomenon* (Fig. 11.1).

Theorem 11.6 (Smoothness and decay). *If f is 2π-periodic and $f, f', \ldots, f^{(k-1)}$ are continuous with $f^{(k)}$ piecewise continuous, then $|c_n| \leq C/|n|^k$. Conversely, rapidly decaying coefficients give a smooth function.*

Proof. Integrating by parts k times, with periodicity killing the boundary terms, $c_n = \frac{1}{(in)^k} \cdot \frac{1}{2\pi} \int f^{(k)}(x) e^{-inx} dx$, and the integral is bounded. ∎

The square wave has a jump and $|b_n| \sim 1/n$. The function $|x|$ has a corner and coefficients of order $1/n^2$. A smooth periodic function has coefficients decaying faster than any power. Sharp features need high frequencies. In quantum terms, a sharply localized wavefunction needs large momenta.

Theorem 11.7 (Bessel's inequality). *For any orthonormal set $\{e_n\}$ in an inner-product space and any f, $\sum_n |\langle e_n | f \rangle|^2 \leq \|f\|^2$.*

Proof. For a finite partial sum, $S_N = \sum_{n \leq N} \langle e_n | f \rangle e_n$ is the projection of f onto the span of $e_1, \ldots, e_N$. By Theorem 1.18, $\|f\|^2 = \|S_N\|^2 + \|f - S_N\|^2 \geq \|S_N\|^2 = \sum_{n \leq N} |\langle e_n | f \rangle|^2$. Let $N \to \infty$. ∎

Theorem 11.8 (Completeness and Parseval's identity). *The functions $\{e^{inx}/\sqrt{2\pi}\}_{n \in \mathbb{Z}}$ form a complete orthonormal set in $L^2(-\pi, \pi)$. Every $f \in L^2$ satisfies $\|f - \sum_{|n| \leq N} c_n e^{inx}\| \to 0$, and*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f|^2 dx = \sum_{n=-\infty}^{\infty} |c_n|^2 = \frac{a_0^2}{4} + \frac{1}{2} \sum_{n=1}^{\infty} (a_n^2 + b_n^2).$$

Completeness is the hard part. One proof shows that the Fourier partial sums of continuous functions converge in norm, using Fejér's averaging, and that continuous functions are dense in L^2. Given completeness, Bessel's inequality becomes an equality. Parseval's identity is the infinite-dimensional Pythagorean theorem.

Physically, the total energy of a wave is the sum of the energies in its modes, and in quantum mechanics the total probability is the sum of the probabilities of the outcomes.

Example 11.9 ($\zeta(4)$ from Parseval). For $f(x) = x^2$: $\frac{1}{\pi}\int_{-\pi}^{\pi} x^4\,dx = \frac{2\pi^4}{5}$, while the right side of Parseval's identity, multiplied by 2, is $\frac{a_0^2}{2} + \sum a_n^2 = \frac{2\pi^4}{9} + 16\sum \frac{1}{n^4}$. Hence $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$.

11.3 The Fourier Transform

A function on the whole line is not periodic, but it can be regarded as the limit of functions of period $2L$ as $L \to \infty$. With period $2L$, the complex series is $f = \sum c_n e^{ik_n x}$ with $k_n = n\pi/L$ and $c_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{-ik_n x}dx$. The allowed wavenumbers are spaced by $\Delta k = \pi/L$. Writing $c_n = \frac{\Delta k}{\sqrt{2\pi}}\hat f(k_n)$ with $\hat f(k) = \frac{1}{\sqrt{2\pi}}\int_{-L}^{L} f(x)e^{-ikx}dx$, the series becomes a Riemann sum $\frac{1}{\sqrt{2\pi}}\sum_n \hat f(k_n)e^{ik_n x}\Delta k$. As $L\to\infty$ the sum becomes an integral.

Definition 11.10 (Fourier transform). For $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, the *Fourier transform* and its inverse are

$$\hat f(k) = \mathcal{F}[f](k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)\,e^{-ikx}\,dx, \qquad f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \hat f(k)\,e^{ikx}\,dk.$$

We use the symmetric normalization standard in quantum mechanics, which puts $1/\sqrt{2\pi}$ on both transforms. Other books put $1/2\pi$ on one side only, or use $e^{-2\pi i\xi x}$. Always check the convention before using a table of transforms. The second formula, the *Fourier inversion theorem*, holds whenever f and $\hat f$ are both integrable, and at points of continuity of f.

Theorem 11.11 (The Gaussian transforms to a Gaussian). *For $a > 0$, $\mathcal{F}\left[e^{-ax^2}\right](k) = \dfrac{1}{\sqrt{2a}}\,e^{-k^2/(4a)}$.*

Proof. Complete the square: $-ax^2 - ikx = -a\left(x + \frac{ik}{2a}\right)^2 - \frac{k^2}{4a}$. Then

$$\hat f(k) = \frac{e^{-k^2/4a}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-a(x+ik/2a)^2}dx = \frac{e^{-k^2/4a}}{\sqrt{2\pi}}\sqrt{\frac{\pi}{a}}.$$

Shifting the contour of integration from the real line to $\operatorname{Im} x = k/2a$ does not change the integral. By Cauchy's theorem, the integral around the rectangle between the two lines vanishes, and the vertical sides contribute nothing as their distance from the origin tends to infinity. Alternatively, see the challenge exercise of Chapter 5. ■

The narrower the Gaussian in x (large a), the wider its transform in k. The product of the widths is fixed: the standard deviation of $|e^{-ax^2}|^2$ in x is $\frac{1}{2\sqrt{a}}$ and that of $|\hat{f}|^2$ in k is $\sqrt{a}$, with product $\frac{1}{2}$. This is the minimum possible (Theorem 11.20).

Example 11.12 (Box function). Let $f = 1$ for $|x| < a$ and 0 otherwise. Then

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-ikx} dx = \sqrt{\frac{2}{\pi}} \frac{\sin ka}{k}.$$

A sharp-edged pulse has a slowly decaying transform with oscillating side lobes. This is the single-slit diffraction pattern of optics, whose intensity is proportional to $\sin^2(ka)/k^2$.

Example 11.13 (Exponential and Lorentzian). For $f = e^{-\gamma|x|}$ with $\gamma > 0$, $\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \left[\frac{1}{\gamma + ik} + \frac{1}{\gamma - ik} \right] = \sqrt{\frac{2}{\pi}} \frac{\gamma}{\gamma^2 + k^2}$. Inverting, the transform of a Lorentzian is a two-sided exponential, which is Example 8.35 again. A state that decays exponentially in time has a Lorentzian distribution of energies, with width γ.

Theorem 11.14 (Properties of the Fourier transform). *For suitable f, g, real a, and $c \neq 0$:*

Property	*Function*	*Transform*
Shift	$f(x-a)$	$e^{-ika}\hat{f}(k)$
Modulation	$e^{ik_0 x} f(x)$	$\hat{f}(k-k_0)$
Scaling	$f(cx)$	$\frac{1}{\lvert c\rvert}\hat{f}(k/c)$
Derivative	$f'(x)$	$ik\,\hat{f}(k)$
Multiplication by x	$xf(x)$	$i\,\hat{f}'(k)$
Conjugate	$f^*(x)$	$\hat{f}(-k)^*$

Proof. Each property is a substitution or an integration by parts. For example, the transform of f' is $\frac{1}{\sqrt{2\pi}}\int f' e^{-ikx}dx = \frac{1}{\sqrt{2\pi}}\big[fe^{-ikx}\big] + \frac{ik}{\sqrt{2\pi}}\int f e^{-ikx}dx = ik\hat{f}$, since $f \to 0$ at $\pm\infty$. ∎

The derivative rule is the most important: *differentiation in x becomes multiplication by ik in k*. In momentum space, the operator $-i\hbar\, d/dx$ acts as multiplication by $\hbar k$. That is why plane waves e^{ikx} are momentum eigenstates and $\hat{f}(k)$ is the momentum-space wavefunction.

Theorem 11.15 (Plancherel). $\displaystyle\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |\hat{f}(k)|^2 dk$, *and more generally* $\langle f|g\rangle = \langle \hat{f}|\hat{g}\rangle$. *The Fourier transform is a unitary operator on* $L^2(\mathbb{R})$.

This is Parseval's identity for the continuum. With the symmetric normalization, no factors of 2π appear. A normalized wavefunction has a normalized transform, so $|\hat{\psi}(k)|^2$ is itself a probability density.

11.4 The Delta Function and Completeness

Substitute the definition of $\hat{f}$ into the inversion formula and exchange the integrals formally:

$$f(x) = \int_{-\infty}^{\infty} f(x')\Big[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x-x')}dk\Big]dx'.$$

For this to hold for every f, the bracket must be $\delta(x-x')$.

Theorem 11.16 (Fourier representation of δ). $\displaystyle\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x-x')}\,dk = \delta(x-x')$, *in the sense of distributions.*

Justification. Insert the convergence factor $e^{-\epsilon k^2}$. By Theorem 11.11,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x-x')-\epsilon k^2}dk = \frac{1}{\sqrt{4\pi\epsilon}}e^{-(x-x')^2/4\epsilon},$$

a normalized Gaussian, which tends to $\delta(x-x')$ as $\epsilon \to 0^+$ by (7.2). ∎

In Dirac notation with plane-wave states $\langle x|k\rangle = e^{ikx}/\sqrt{2\pi}$, this reads

$$\langle x|x'\rangle = \int \langle x|k\rangle\langle k|x'\rangle\, dk = \delta(x-x'), \qquad \int |k\rangle\langle k|\, dk = I,$$

which are the continuum versions of orthonormality and completeness. Plane waves are a complete set, but they are not normalizable: $\langle k|k'\rangle = \delta(k-k')$. The Fourier transform is simply the change of basis from $\{|x\rangle\}$ to $\{|k\rangle\}$: $\hat{f}(k) = \langle k|f\rangle$.

11.5 Convolution and the Heat Equation on the Line

Definition 11.17 (Convolution). $(f*g)(x) = \displaystyle\int_{-\infty}^{\infty} f(x-y)\,g(y)\,dy.$

Theorem 11.18 (Convolution theorem). $\mathcal{F}[f*g] = \sqrt{2\pi}\,\hat{f}\,\hat{g}$, *and* $\mathcal{F}[fg] = \frac{1}{\sqrt{2\pi}}\,\hat{f}*\hat{g}$.

Proof. $\frac{1}{\sqrt{2\pi}}\iint f(x-y)g(y)e^{-ikx}\,dy\,dx$. Substitute $x = u+y$; the integrand factors as $f(u)e^{-iku}\cdot g(y)e^{-iky}$, and the double integral is $\sqrt{2\pi}\hat{f}(k)\cdot\sqrt{2\pi}\hat{g}(k)/\sqrt{2\pi}$. ∎

Example 11.19 (Heat equation on the whole line). Solve $u_t = \kappa u_{xx}$ for $x\in\mathbb{R}$, $t>0$, with $u(x,0)=f(x)$.

Solution. Transform in x. By the derivative rule, $\hat{u}_t = -\kappa k^2\hat{u}$, an ordinary differential equation for each k, so $\hat{u}(k,t) = \hat{f}(k)e^{-\kappa k^2 t}$. By Theorem 11.11 with $a = 1/(4\kappa t)$, the function $e^{-\kappa k^2 t}$ is $\sqrt{2\pi}$ times the transform of $G(x,t) = \frac{1}{\sqrt{4\pi\kappa t}}e^{-x^2/(4\kappa t)}$. By the convolution theorem,

$$u(x,t) = (G*f)(x) = \frac{1}{\sqrt{4\pi\kappa t}}\int_{-\infty}^{\infty} e^{-(x-y)^2/(4\kappa t)}f(y)\,dy.$$

The *heat kernel* G spreads an initial point of heat into a Gaussian whose width grows like $\sqrt{t}$. This diffusive scaling is the signature of random walks.

11.6 The Uncertainty Principle for Fourier Transforms

For a function f with $\int|f|^2dx = 1$, regard $|f|^2$ as a probability density in x and $|\hat{f}|^2$ as one in k. Define the means $\langle x\rangle = \int x|f|^2dx$ and $\langle k\rangle = \int k|\hat{f}|^2dk$, and the

standard deviations $\sigma_x^2 = \int (x - \langle x\rangle)^2 |f|^2 dx$ and $\sigma_k^2 = \int (k - \langle k\rangle)^2 |\hat f|^2 dk$.

Theorem 11.20 (Fourier uncertainty principle). *For every normalized f for which these quantities are finite,*

$$\sigma_x \, \sigma_k \geq \frac{1}{2},$$

with equality if and only if f is a Gaussian, $f(x) = Ce^{-a(x-x_0)^2 + ik_0 x}$ with $\operatorname{Re} a > 0$.

Proof. By the modulation and shift rules, we may translate so that $\langle x\rangle = \langle k\rangle = 0$, which changes neither width. By Plancherel and the derivative rule, $\sigma_k^2 = \int |k\hat f|^2 dk = \int |f'|^2 dx$. Hence, by Cauchy–Schwarz,

$$\sigma_x^2 \sigma_k^2 = \int |xf|^2 dx \int |f'|^2 dx \geq \left| \int x f^* f' \, dx \right|^2 \geq \left(\operatorname{Re} \int x f^* f' dx \right)^2.$$

Now $\operatorname{Re}(f^* f') = \tfrac{1}{2}(|f|^2)'$, so integrating by parts, $\operatorname{Re} \int x f^* f' \, dx = \tfrac{1}{2} \int x (|f|^2)' dx = -\tfrac{1}{2} \int |f|^2 dx = -\tfrac{1}{2}$. So $\sigma_x^2 \sigma_k^2 \geq \tfrac{1}{4}$. Equality in Cauchy–Schwarz requires $f' = -cxf$ for a constant c, which gives a Gaussian. Equality in the second step requires c real and positive. ∎

LOOKING AHEAD. De Broglie's relation $p = \hbar k$ turns Theorem 11.20 into the *Heisenberg uncertainty principle* $\sigma_x \sigma_p \geq \hbar/2$. The position and momentum wavefunctions of a particle are Fourier transforms of each other, and no function is sharply peaked in both. This is not a statement about the clumsiness of measuring instruments. It is a theorem about functions. Chapter 15 proves the general version for any pair of observables, $\sigma_A \sigma_B \geq \tfrac{1}{2} |\langle [\hat A, \hat B] \rangle|$.

11.6.1 WAVE PACKETS AND GROUP VELOCITY

A *wave packet* is a superposition of plane waves $e^{i(kx - \omega t)}$ with wavenumbers near some k_0:

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int \phi(k) \, e^{i(kx - \omega(k)t)} \, dk.$$

The *dispersion relation* $\omega(k)$ depends on the medium. For light in vacuum $\omega = ck$. For a free quantum particle $\hbar\omega = E = \frac{\hbar^2 k^2}{2m}$. Expanding $\omega(k) \approx \omega_0 + v_g(k - k_0)$, with $v_g = \omega'(k_0)$,

$$\psi(x,t) \approx e^{i(k_0 v_g - \omega_0)t} \, \psi(x - v_g t, 0).$$

The envelope moves rigidly at the *group velocity* $v_g = \mathrm{d}\omega/\mathrm{d}k$, while individual crests move at the *phase velocity* $v_p = \omega/k$. For a free quantum particle, $v_g = \hbar k/m = p/m$, the classical velocity, while $v_p = \hbar k/2m$ is half of it. The particle is the packet, not the crest. The quadratic term in $\omega(k)$, neglected above, makes the packet spread. That spreading is computed exactly for a Gaussian in Chapter 13.

11.7 The Discrete Fourier Transform

For N samples $f_0, \dots, f_{N-1}$, the *discrete Fourier transform* is $F_k = \sum_{j=0}^{N-1} f_j\, \omega^{-jk}$, with $\omega = e^{2\pi i/N}$. The inverse is $f_j = \frac{1}{N}\sum_k F_k \omega^{jk}$, a consequence of $\sum_{k=0}^{N-1} \omega^{k(j-j')} = N\delta_{jj'}$ (the roots of unity sum to zero). The matrix $\frac{1}{\sqrt{N}}[\omega^{-jk}]$ is unitary. Computed naively, the DFT takes $O(N^2)$ operations. The *fast Fourier transform* of Cooley and Tukey splits it recursively into even and odd samples and takes $O(N \log N)$. The FFT underlies digital audio, image compression, and MRI, and its quantum analog, the quantum Fourier transform, is the engine of Shor's factoring algorithm.

11.8 Worked Problems

The problems below connect the smoothness of a function to the decay of its Fourier coefficients, use series and Parseval's identity to evaluate sums, compute transforms by duality and convolution, and examine two phenomena that matter in practice: the chirped pulse, which exceeds the uncertainty minimum, and aliasing in sampled data.

Problem 11.1 A smooth odd function, and two sums

Find the Fourier series of $f(x) = x(\pi^2 - x^2)$ on $[-\pi, \pi]$. Use it to show

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^3} = \frac{\pi^3}{32} \qquad \text{and} \qquad \sum_{n=1}^{\infty} \frac{1}{n^6} = \frac{\pi^6}{945}.$$

Solution. f is odd, so only sines appear. Integrating by parts three times (or differentiating under the integral sign),

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x(\pi^2 - x^2)\sin nx\, dx = \frac{12(-1)^{n+1}}{n^3}.$$

The periodic extension of f is continuous, and so is its derivative, because $f(\pm\pi) = 0$ and $f'(\pi) = f'(-\pi) = -2\pi^2$. Its second derivative $-6x$ jumps

at $\pm\pi$. By Theorem 11.6, a function whose first p derivatives are continuous (periodically) and whose next derivative has jumps has coefficients of order $1/n^{p+2}$. Here $p=1$, consistent with $1/n^3$. The series converges to f everywhere, uniformly.

At $x=\pi/2$, $\sin\frac{n\pi}{2}$ is 0 for even n and $(-1)^k$ for $n=2k+1$, and $(-1)^{n+1}=1$ for odd n. So

$$f\left(\tfrac{\pi}{2}\right)=\frac{\pi}{2}\cdot\frac{3\pi^2}{4}=\frac{3\pi^3}{8}=12\sum_{k\ge 0}\frac{(-1)^k}{(2k+1)^3},$$

which gives $\pi^3/32$. For the second sum, Parseval's identity (Theorem 11.8) in the form $\frac{1}{\pi}\int_{-\pi}^{\pi}f^2\,dx=\sum b_n^2$ gives

$$144\sum_{n\ge 1}\frac{1}{n^6}=\frac{2}{\pi}\int_0^{\pi}\left(\pi^4x^2-2\pi^2x^4+x^6\right)dx=\frac{2}{\pi}\cdot\pi^7\left(\frac{1}{3}-\frac{2}{5}+\frac{1}{7}\right)=\frac{16\pi^6}{105},$$

so $\sum 1/n^6=\frac{16\pi^6}{105\cdot 144}=\frac{\pi^6}{945}$. Each value $\zeta(2m)$ is a rational multiple of π^{2m}, and this method produces them one by one. Nothing comparable is known for odd arguments: the value of $\zeta(3)$ remains mysterious.

Problem 11.2 A cosine series for a sine

Find the Fourier cosine series of $\sin x$ on $[0,\pi]$, and use it to evaluate $\sum_{k\ge 1}\frac{1}{4k^2-1}$.

Solution. The cosine series is the Fourier series of the even extension, which is $|\sin x|$, the output of a full-wave rectifier. Its coefficients are $a_n=\frac{2}{\pi}\int_0^{\pi}\sin x\cos nx\,dx$. Using $2\sin x\cos nx=\sin(n+1)x-\sin(n-1)x$, for $n\neq 1$,

$$a_n=\frac{1}{\pi}\Big[-\frac{\cos(n+1)x}{n+1}+\frac{\cos(n-1)x}{n-1}\Big]_0^{\pi}=\begin{cases}-\dfrac{4}{\pi(n^2-1)}, & n \text{ even},\\[2mm] 0, & n \text{ odd},\end{cases}$$

and $a_1=\frac{1}{\pi}\int_0^{\pi}\sin 2x\,dx=0$, $a_0=\frac{4}{\pi}$. Writing $n=2k$,

$$\sin x=\frac{2}{\pi}-\frac{4}{\pi}\sum_{k\ge 1}\frac{\cos 2kx}{4k^2-1}\qquad(0\le x\le\pi).$$

At $x=0$, $0=\frac{2}{\pi}-\frac{4}{\pi}\sum\frac{1}{4k^2-1}$, so $\sum_{k\ge 1}\frac{1}{4k^2-1}=\frac{1}{2}$. This also follows from the telescoping $\frac{1}{4k^2-1}=\frac{1}{2}\left(\frac{1}{2k-1}-\frac{1}{2k+1}\right)$. A sine function expanded in cosines may seem perverse, but the even extension is continuous, so the coefficients decay like $1/n^2$. The odd extension of $\sin x$ on $[0,\pi]$ is $\sin x$ itself, whose series is a single term.

Problem 11.3 A transform by duality

Find the Fourier transform of $f(x) = \dfrac{\sin x}{x}$.

Solution. Direct integration is awkward, since f is not absolutely integrable. Use the box function instead. By Example 11.12, the box $B(x) = 1$ for $|x| < 1$ (and 0 otherwise) has $\hat{B}(k) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-ikx} dx = \sqrt{\frac{2}{\pi}}\, \frac{\sin k}{k}$. By the inversion theorem,

$$B(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sqrt{\frac{2}{\pi}}\, \frac{\sin k}{k}\, e^{ikx}\, dk, \quad \text{so} \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\sin k}{k}\, e^{ikx}\, dk = \sqrt{\frac{\pi}{2}}\, B(x).$$

Rename $k \leftrightarrow x$ and replace x by $-k$; since B is even,

$$\mathcal{F}\Big[\frac{\sin x}{x}\Big](k) = \sqrt{\frac{\pi}{2}}\, B(k) = \begin{cases} \sqrt{\pi/2}, & |k| < 1, \\ 0, & |k| > 1. \end{cases}$$

At $|k| = 1$ the inverse transform gives the average, $\frac{1}{2}\sqrt{\pi/2}$. A numerical evaluation of the defining integral, cut off at $|x| = 2000$, gives 1.25322 at $k = \frac{1}{2}$ (exact $\sqrt{\pi/2} = 1.25331$) and 2×10^{-4} at $k = \frac{3}{2}$.

This is *duality*: since $\mathcal{F}^2[f](x) = f(-x)$, every transform pair gives a second one with the roles reversed. The function $\frac{\sin x}{x}$ is band-limited: it contains no wavenumbers above 1. By the sampling theorem, such a function is determined by its values at the points $x = n\pi$, which is the mathematical basis of digital audio.

Problem 11.4 Heat flow from a hot interval

Solve $u_t = u_{xx}$ on $\mathbb{R}$ with $u(x,0) = 1$ for $|x| < 1$ and 0 otherwise.

Solution. By Example 11.19, the solution is the convolution of the initial data with the heat kernel $G_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-x^2/4t}$:

$$u(x,t) = \int_{-1}^{1} \frac{e^{-(x-y)^2/4t}}{\sqrt{4\pi t}}\, dy.$$

Substitute $s = (y - x)/(2\sqrt{t})$, so $dy = 2\sqrt{t}\, ds$ and the limits become $s_{\pm} = \frac{\pm 1 - x}{2\sqrt{t}}$:

$$u(x,t) = \frac{1}{\sqrt{\pi}} \int_{s_-}^{s_+} e^{-s^2} ds = \frac{1}{2}\Big[\operatorname{erf}\Big(\frac{1+x}{2\sqrt{t}}\Big) + \operatorname{erf}\Big(\frac{1-x}{2\sqrt{t}}\Big)\Big],$$

where $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-s^2} ds$ is the error function, an odd function with $\operatorname{erf}(\pm\infty) = \pm 1$.

As $t \to 0^+$, both arguments tend to $+\infty$ for $|x| < 1$, so $u \to 1$, and one tends to $-\infty$ for $|x| > 1$, so $u \to 0$. For large t, $\operatorname{erf}(z) \approx \frac{2z}{\sqrt{\pi}}$, so $u \approx \frac{1}{\sqrt{\pi t}}$: the total heat 2 spreads into a Gaussian of width $\sqrt{2t}$. And for every $t > 0$, $u(x,t) > 0$ for all x, however far away. The heat equation propagates disturbances at infinite speed, a sign that it is a diffusion law, not a wave equation. The free Schrödinger equation, which differs by a factor of i, shares this feature (Chapter 13).

Problem 11.5 A chirped Gaussian

Let $f(x) = e^{-(a-ib)x^2}$ with $a > 0$ and b real. Show that $|f|^2$ has the same width as for $b = 0$, but that $\sigma_x \sigma_k = \frac{1}{2}\sqrt{1 + b^2/a^2}$.

Solution. $|f|^2 = e^{-2ax^2}$ does not depend on b, so $\sigma_x^2 = \frac{1}{4a}$, as for the unchirped Gaussian. The phase e^{ibx^2} has a local wavenumber $\frac{d}{dx}(bx^2) = 2bx$ that varies with position, which is the meaning of *chirp*.

Theorem 11.11 holds for complex a with positive real part, by analytic continuation in a (both sides are analytic for $\operatorname{Re} a > 0$ and agree for real a). So $\hat{f}(k) = \frac{1}{\sqrt{2(a-ib)}} e^{-k^2/4(a-ib)}$, and

$$|\hat{f}(k)|^2 \propto \exp\Big(-\frac{k^2}{2}\operatorname{Re}\frac{1}{a-ib}\Big) = \exp\Big(-\frac{ak^2}{2(a^2+b^2)}\Big).$$

This Gaussian has $\sigma_k^2 = \frac{a^2+b^2}{a}$. Therefore

$$\sigma_x \sigma_k = \sqrt{\frac{1}{4a}\cdot\frac{a^2+b^2}{a}} = \frac{1}{2}\sqrt{1+\frac{b^2}{a^2}}.$$

Only the unchirped Gaussian attains the minimum $\frac{1}{2}$ of Theorem 11.20. A numerical check with $a = 1.3$, $b = 0.7$, using $\sigma_k^2 = \int |f'|^2 / \int |f|^2$, gives 0.567877, against the formula's 0.567878. In quantum mechanics, a free Gaussian packet develops exactly this chirp as it spreads (Chapter 13). The spreading is invisible in σ_k, which is constant, but appears in the growing σ_x.

Problem 11.6 Aliasing

A signal $\cos(2\pi \cdot 7t)$, of frequency 7 Hz, is sampled 8 times per second, at $t_j = j/8$. Show that the samples are identical to those of a 1 Hz signal. What sampling rate would distinguish the two?

Solution. At the sample times, $\cos(2\pi \cdot 7j/8) = \cos(2\pi j - 2\pi j/8) = \cos(2\pi j/8)$, since cos has period 2π and is even. So the samples coincide with those of $\cos(2\pi \cdot 1 \cdot t)$. In terms of the DFT, frequencies f and $f \pm N$ (in cycles

per record of N samples) are indistinguishable, because $e^{2\pi i(f+N)j/N} = e^{2\pi ifj/N}$. And the real cosine at 7 equals the one at $-7 \equiv 1$.

To distinguish all frequencies below $f_{\max}$, the sampling rate must exceed $2f_{\max}$, the *Nyquist rate*. For a 7 Hz signal, more than 14 samples per second. A rate of 8 Hz can represent unambiguously only frequencies below 4 Hz, and everything above folds back into that band. This is why audio is sampled at 44.1 kHz, just over twice the 20 kHz limit of human hearing, and why a filter removes higher frequencies before sampling. The wagon wheels that seem to turn backward in films are aliasing in time.

Problem 11.7 A periodic response

Find the 2π-periodic solution of $y'' + 2y = s(x)$, where s is the square wave of Example 11.3. Why is there exactly one?

Solution. The square wave has the series $s(x) = \frac{4}{\pi}\sum_{n\ \text{odd}} \frac{\sin nx}{n}$. Look for $y = \sum c_n \sin nx$. Then $y'' + 2y = \sum (2 - n^2)c_n \sin nx$, and matching coefficients gives

$$c_n = \frac{4}{\pi n(2-n^2)} \ (n \text{ odd}),$$

$$y(x) = \frac{4}{\pi}\sum_{n\ \text{odd}} \frac{\sin nx}{n(2-n^2)} = \frac{4}{\pi}\Big(\sin x - \frac{\sin 3x}{21} - \frac{\sin 5x}{115} - \cdots\Big).$$

The coefficients decay like $1/n^3$, two powers faster than those of s: integrating twice smooths the jumps of s into jumps of y''. A numerical check, differentiating the partial sum with 2000 terms at $x = 1$, gives $y'' + 2y = 1.00003$, against $s(1) = 1$.

The solution is unique because the homogeneous equation $y'' + 2y = 0$ has solutions $\cos\sqrt{2}x$ and $\sin\sqrt{2}x$, which are not 2π-periodic since $\sqrt{2}$ is not an integer. If the coefficient were n^2 for some odd n, say $y'' + 9y = s$, the $n = 3$ coefficient would require dividing by 0. The forcing contains a frequency the system resonates with, and no periodic solution exists: the response grows like $x\cos 3x$ (Example 10.8).

EXERCISES

A. COMPUTATION

11.1 Find the Fourier series of $f(x) = |x|$ on $[-\pi, \pi]$ and use it to evaluate $\sum_{n\ \text{odd}} 1/n^2$.

11.2 Find the Fourier series of $f(x) = e^x$ on $(-\pi, \pi)$ in complex form.

11.3 Find the Fourier series of $f(x) = \sin^2 x$ by inspection.

11.4 Find the Fourier series of $f(x) = 0$ on $(-\pi, 0)$, $f(x) = x$ on $(0, \pi)$. To what value does it converge at $x = \pi$?

11.5 Find the sine series and the cosine series of $f(x) = x$ on $[0, 1]$.

11.6 Find the complex Fourier coefficients of $f(x) = x$ on $(-\pi, \pi)$ and verify that $c_{-n} = c_n^*$.

11.7 Use Parseval's identity with the square wave to evaluate $\sum_{n \text{ odd}} 1/n^2$.

11.8 Use Parseval's identity with $f(x) = x$ to evaluate $\sum 1/n^2$ again.

11.9 Use the Fourier series of x^2 at $x = 0$ to evaluate $\sum (-1)^{n+1}/n^2$.

11.10 Find the Fourier transform of $f(x) = e^{-a|x|}$ directly, for $a > 0$.

11.11 Find the Fourier transform of $f(x) = xe^{-x^2/2}$ using Theorem 11.14.

11.12 Find the Fourier transform of $f(x) = e^{-(x-3)^2} \cos 5x$.

11.13 Find the Fourier transform of the triangle function $f(x) = 1 - |x|$ for $|x| \le 1$, 0 otherwise. Relate it to the box function of Example 11.12 by convolution.

11.14 Use Plancherel's theorem with the box function to evaluate $\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx$.

11.15 Compute σ_x and σ_k for $f(x) = (2a/\pi)^{1/4} e^{-ax^2}$ and verify that $\sigma_x \sigma_k = \frac{1}{2}$.

11.16 Compute $\sigma_x \sigma_k$ for $f(x) = \sqrt{\gamma}\, e^{-\gamma|x|}$ and confirm that it exceeds $\frac{1}{2}$.

11.17 Evaluate $\displaystyle\int_{-\infty}^{\infty} \delta(x-2) e^{ikx} dx$ and $\displaystyle\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x-3)} \cos x \, dx$ (as a distribution in k).

11.18 Compute the convolution of two Gaussians $e^{-x^2/(2\sigma_1^2)}$ and $e^{-x^2/(2\sigma_2^2)}$ using the convolution theorem. How do the variances combine?

11.19 Solve $u_t = u_{xx}$ on $\mathbb{R}$ with $u(x, 0) = e^{-x^2}$.

11.20 For the dispersion relation $\omega = \sqrt{gk}$ of deep-water waves, compute v_p and v_g. Which is larger?

11.21 For relativistic matter waves with $\hbar\omega = \sqrt{(\hbar k c)^2 + (mc^2)^2}$, show that $v_g v_p = c^2$ and that $v_g < c$.

11.22 Compute the DFT of $(1,1,1,1)$ and of $(1,0,-1,0)$ for $N=4$.

B. THEORY AND PROOF

11.23 Prove the orthogonality relations for $\{\cos nx, \sin nx\}$ on $[-\pi,\pi]$.

11.24 Show that the Fourier series of an even function contains only cosines, and that of an odd function only sines.

11.25 Prove that the Fourier coefficients of an integrable function tend to zero (the Riemann–Lebesgue lemma) for piecewise smooth f, by integrating by parts.

11.26 Prove the scaling and modulation properties in Theorem 11.14.

11.27 Prove Plancherel's theorem formally using Theorem 11.16.

11.28 Show that $\mathcal{F}^2[f](x) = f(-x)$ and $\mathcal{F}^4 = I$. Deduce that the only possible eigenvalues of $\mathcal{F}$ are $\pm 1, \pm i$.

11.29 Show that the Hermite functions $\psi_n(x) = H_n(x)e^{-x^2/2}$ are eigenfunctions of $\mathcal{F}$ with eigenvalue $(-i)^n$, using the generating function of the Hermite polynomials.

11.30 Prove that $\mathcal{F}[f*g] = \sqrt{2\pi}\,\hat f \hat g$ implies that convolution is commutative and associative.

11.31 Prove the Poisson summation formula $\sum_{n} f(n) = \sqrt{2\pi}\sum_{m}\hat f(2\pi m)$ for rapidly decaying f. (Expand the periodic function $\sum_n f(x+n)$ in a Fourier series.)

C. CHALLENGE

11.32 Compute the size of the Gibbs overshoot for the square wave: show that the maximum of the partial sum near $x=0^+$ tends to $\frac{2}{\pi}\int_0^{\pi}\frac{\sin t}{t}dt \approx 1.179$, an overshoot of about 9% of the jump of 2.

11.33 Use Poisson summation with a Gaussian to prove the theta-function identity $\sum_{n} e^{-\pi n^2 t} = \frac{1}{\sqrt{t}}\sum_{n} e^{-\pi n^2/t}$.

11.34 Show that a function and its Fourier transform cannot both vanish outside bounded intervals unless both are zero. (If f vanishes outside $[-a,a]$, show $\hat f$ extends to an entire function of k.)

11.35 Let $\psi_0(x) = (2\pi\sigma^2)^{-1/4}e^{-x^2/4\sigma^2 + ik_0 x}$. Show that $|\hat\psi_0(k)|^2$ is a Gaussian centered at k_0 with standard deviation $1/(2\sigma)$, and compute $\int k|\hat\psi_0|^2 dk$ and

$\int k^2 |\hat{\psi}_0|^2 dk$. These give the mean and spread of momentum for a Gaussian wave packet.



PART III

Quantum Mechanics

From the experiments that broke classical physics to the Schrödinger equation, the hydrogen atom, spin, and entanglement.

CHAPTER 12

The Breakdown of Classical Physics

LEARNING OBJECTIVES

- State the equipartition theorem and use it to derive the Rayleigh–Jeans law and the ultraviolet catastrophe.
- Count the electromagnetic modes in a cavity, and derive Planck's radiation law, the Stefan–Boltzmann law, and Wien's displacement law.
- Explain the photoelectric effect with Einstein's photon hypothesis, and analyze stopping potentials and work functions.
- Derive the Compton shift from relativistic energy and momentum conservation.
- Estimate the classical collapse time of a Rutherford atom, and derive the energy levels and radii of the Bohr model.
- Apply de Broglie's relation $\lambda = h/p$ and interpret Bohr quantization as a standing-wave condition.
- Describe the double-slit experiment with single particles, and the Stern–Gerlach experiment, and explain what each implies about the structure of quantum theory.

By 1900, classical physics seemed nearly complete. Newton's mechanics explained the motion of everything from projectiles to planets (Chapter 2). Maxwell's equations unified electricity, magnetism, and light (Example 6.25). Statistical mechanics explained heat as the random motion of atoms (Example 4.20). Lord Kelvin, in a celebrated lecture of 1900, spoke of only two "clouds" on the horizon. Within a quarter century, those clouds, and several more, had grown into a storm that swept away the classical picture of the microscopic world. This chapter examines the experiments that forced the change. In each

case we do the classical calculation, show precisely where it fails, and show what the new idea is and why it works. The mathematics is that of Parts I and II; the physics is new.

We use the constants

Planck's constant	$h = 6.626 \times 10^{-34}\,\mathrm{J\,s}, \quad \hbar = h/2\pi = 1.055 \times 10^{-34}\,\mathrm{J\,s}$
Speed of light	$c = 2.998 \times 10^{8}\,\mathrm{m/s}$
Boltzmann's constant	$k_B = 1.381 \times 10^{-23}\,\mathrm{J/K}$
Electron charge and mass	$e = 1.602 \times 10^{-19}\,\mathrm{C}, \quad m_e = 9.109 \times 10^{-31}\,\mathrm{kg}$
Useful combinations	$hc = 1240\,\mathrm{eV\,nm}, \quad m_e c^2 = 511.0\,\mathrm{keV}$

and the energy unit $1\ \mathrm{eV} = 1.602 \times 10^{-19}\ \mathrm{J}$, the energy gained by an electron falling through one volt.

12.1 Equipartition and Its Discontents

Classical statistical mechanics assigns probability $p \propto e^{-\beta E}$ to a state of energy E, with $\beta = 1/k_B T$, as derived from the maximum-entropy principle in Example 4.20. For a continuous system the sum becomes an integral over positions and momenta.

Theorem 12.1 (Equipartition). *If the energy of a classical system contains a term aq^2 quadratic in one coordinate or momentum q, then in thermal equilibrium the average of that term is $\frac{1}{2}k_B T$, independent of a.*

Proof. Integrating out the other variables, the average is

$$\langle aq^2 \rangle = \frac{\int_{-\infty}^{\infty} aq^2 e^{-\beta a q^2}\,dq}{\int_{-\infty}^{\infty} e^{-\beta a q^2}\,dq} = -\frac{\partial}{\partial \beta} \ln \int_{-\infty}^{\infty} e^{-\beta a q^2}\,dq = -\frac{\partial}{\partial \beta} \ln \sqrt{\frac{\pi}{\beta a}} = \frac{1}{2\beta},$$

using Theorem 5.12 and differentiation under the integral sign. ∎

A one-dimensional harmonic oscillator has energy $\frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$, with two quadratic terms, so its mean energy is $k_B T$, independent of its frequency. A monatomic gas has $\frac{3}{2}k_B T$ per atom, which correctly gives the heat capacity $\frac{3}{2}R$ per mole. But equipartition also predicts heat capacities for diatomic gases and solids that fail at low temperature. Some degrees of freedom are mysteriously "frozen out" when $k_B T$ is small, which is the first hint that energy cannot be divided arbitrarily finely.

12.2 BLACKBODY RADIATION

Every object emits electromagnetic radiation, with a spectrum that depends on its temperature: a stove burner glows dull red, the Sun yellow-white. An ideal absorber, a *blackbody*, has a universal spectrum that depends on T alone. It can be realized as a small hole in a cavity whose walls are in equilibrium at temperature T. The goal is to compute the *spectral energy density* $u(\nu, T)$, the energy per unit volume per unit frequency inside the cavity.

12.2.1 COUNTING MODES

Inside a cubical cavity of side L with conducting walls, the electric field satisfies the wave equation (Example 6.25), and separating variables (Section 10.6) gives standing waves with wave vectors

$$\mathbf{k} = \frac{\pi}{L}\,\langle n_x, n_y, n_z\rangle, \qquad n_x, n_y, n_z = 1, 2, 3, \dots,$$

and frequencies $\nu = \frac{c|\mathbf{k}|}{2\pi}$. Each allowed $\mathbf{k}$ supports two independent polarizations.

Theorem 12.2 (Mode density). *For a large cavity of volume $V = L^3$, the number of electromagnetic modes with frequency between ν and $\nu + d\nu$ is*

$$g(\nu)\,d\nu = \frac{8\pi V \nu^2}{c^3}\,d\nu.$$

Proof. The allowed $\mathbf{k}$ form a lattice in the positive octant of k-space with spacing π/L, so each point occupies volume $(\pi/L)^3$. The number of points with $|\mathbf{k}| \le k$ is approximately the volume of one octant of a ball divided by that cell volume: $\frac{1}{8}\cdot\frac{4}{3}\pi k^3 \big/ (\pi/L)^3 = \frac{Vk^3}{6\pi^2}$. Multiply by 2 for polarizations and substitute $k = 2\pi\nu/c$: $N(\nu) = \frac{8\pi V \nu^3}{3c^3}$. Differentiate with respect to ν. ∎

The number of modes grows like ν^2, without limit. That fact alone dooms the classical theory.

12.2.2 THE ULTRAVIOLET CATASTROPHE

Each mode is a harmonic oscillator, so classically, by equipartition, each has mean energy $k_B T$. Hence

$$u_{\mathrm{RJ}}(\nu, T) = \frac{g(\nu)}{V} k_B T = \frac{8\pi\nu^2}{c^3} k_B T \tag{12.1}$$

(the *Rayleigh–Jeans law*). It agrees with experiment at low frequencies, but it is absurd at high ones. The total energy density $\int_0^\infty u\,d\nu$ diverges, so a cavity in equilibrium would contain infinite energy, concentrated in the ultraviolet and beyond. This is the *ultraviolet catastrophe*. The derivation uses only Maxwell's equations and classical statistical mechanics, so one of them must be wrong.

12.2.3 PLANCK'S HYPOTHESIS

In 1900 Max Planck found a formula that fit the data perfectly. Within weeks he had found an interpretation, which he regarded as an act of desperation: an oscillator of frequency ν can hold only energies that are integer multiples of a quantum $h\nu$,

$$E_n = nh\nu, \qquad n = 0, 1, 2, \dots.$$

Theorem 12.3 (Planck's mean energy). *If an oscillator of frequency ν has energies $nh\nu$ with Boltzmann probabilities $p_n \propto e^{-\beta n h\nu}$, its mean energy is*

$$\langle E \rangle = \frac{h\nu}{e^{h\nu/k_B T} - 1}.$$

Proof. With $x = \beta h\nu$, the partition function is the geometric series $Z = \sum_{n\ge 0} e^{-nx} = \frac{1}{1-e^{-x}}$. Then $\langle E\rangle = -\,\partial \ln Z/\partial\beta = h\nu \frac{e^{-x}}{1-e^{-x}} = \frac{h\nu}{e^{x}-1}$. ∎

For $h\nu \ll k_B T$, $e^x - 1 \approx x$ and $\langle E\rangle \approx k_B T$: equipartition is recovered. For $h\nu \gg k_B T$, $\langle E\rangle \approx h\nu e^{-h\nu/k_B T} \to 0$. High-frequency modes cannot be excited because even one quantum costs far more than the typical thermal energy. They are frozen out, which resolves the catastrophe.

Theorem 12.4 (Planck's radiation law).

$$u(\nu, T) = \frac{8\pi h\nu^3}{c^3}\,\frac{1}{e^{h\nu/k_B T} - 1}.$$

Corollary 12.5 (Stefan–Boltzmann and Wien laws). *The total energy density is $u(T) = aT^4$ with $a = \dfrac{8\pi^5 k_B^4}{15 h^3 c^3}$, and the power radiated per unit area of a blackbody surface is σT^4 with $\sigma = \dfrac{ca}{4} = \dfrac{2\pi^5 k_B^4}{15 h^3 c^2} = 5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$. The spectral peak in frequency occurs at $h\nu_{\max} = 2.821\,k_B T$.*

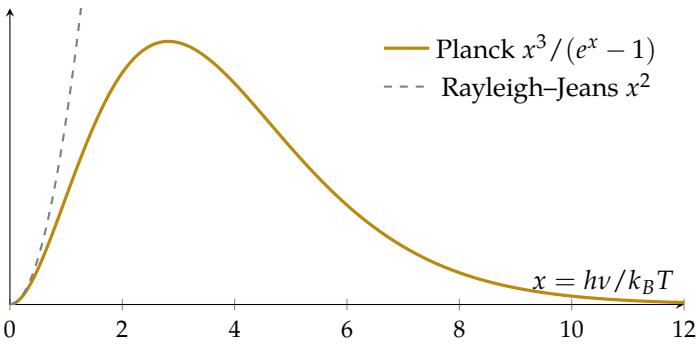


Figure 12.1: Planck's law in dimensionless form. The classical Rayleigh–Jeans curve agrees at low frequency and diverges at high frequency; Planck's curve peaks at $x \approx 2.82$ and falls off exponentially.

Proof. Substitute $x = h\nu/k_BT$: $u(T) = \frac{8\pi(k_BT)^4}{h^3c^3}\int_0^\infty \frac{x^3}{e^x-1}dx$. Expand $\frac{1}{e^x-1} = \sum_{n\geq 1} e^{-nx}$ and integrate term by term using Theorem 5.30: $\int_0^\infty x^3 e^{-nx}dx = \frac{6}{n^4}$, so the integral is $6\zeta(4) = 6\cdot\frac{\pi^4}{90} = \frac{\pi^4}{15}$, using the value of $\zeta(4)$ computed by Parseval's identity in Chapter 11. The factor $c/4$ comes from integrating the flux of isotropic radiation over a hemisphere. For the peak, setting the derivative of $x^3/(e^x-1)$ to zero gives $3(1-e^{-x}) = x$, whose positive root is $x \approx 2.821$. ∎

From the measured σ and the peak position, Planck extracted both h and k_B in 1900. His values were within a few percent of the modern ones, and k_B in turn gave the first accurate value of Avogadro's number.

Example 12.6 (The Sun and the universe). The solar spectrum is close to that of a blackbody at $T \approx 5800$ K, and the peak in wavelength (Wien's law, $\lambda_{\max}T = 2.898\times10^{-3}$ m K) is at 500 nm, in the green. The cosmic microwave background, the afterglow of the Big Bang, is the most perfect blackbody spectrum ever measured, with $T = 2.725$ K and a wavelength peak near 1.06 mm.

12.3 The Photoelectric Effect

When light shines on a clean metal surface, electrons are ejected. The experiments of Lenard (1902) and later Millikan established three facts.

(1) The maximum kinetic energy of the ejected electrons depends on the *frequency* of the light, not on its intensity.

(2) Below a *threshold frequency* ν_0, which depends on the metal, no electrons are emitted, however intense the light.

(3) Above threshold, emission begins with no measurable delay, even for very faint light.

Classically, light is a continuous wave whose energy density is proportional to its intensity. A brighter beam should eject faster electrons, any frequency should work if the light is bright enough, and faint light should need time to accumulate enough energy in one electron. All three predictions fail.

In 1905 Einstein proposed that light itself consists of quanta, later called *photons*, each of energy $E = h\nu$. An electron absorbs a single photon, all or nothing. It must spend at least the *work function* ϕ to escape the metal, and the rest becomes kinetic energy:

$$K_{\max} = h\nu - \phi. \tag{12.2}$$

Intensity sets the *number* of photons, and hence the number of electrons, not their energy. The threshold is $\nu_0 = \phi/h$. Emission is immediate because one photon delivers its whole energy at once. Experimentally, $K_{\max}$ is measured by the *stopping potential* V_s that just halts the fastest electrons, $eV_s = K_{\max}$. A plot of V_s against ν is a straight line of slope h/e for every metal. Millikan set out to refute Einstein's equation, and his careful measurements in 1916 confirmed it and gave h to within 0.5%.

Example 12.7 (Photoemission from sodium). Sodium has $\phi = 2.28$ eV. Find the threshold wavelength, and the maximum speed of electrons ejected by 400 nm light.

Solution. $\lambda_0 = hc/\phi = 1240\ \text{eV nm}/2.28\ \text{eV} = 544$ nm, in the green. For 400 nm light, $h\nu = 1240/400 = 3.10$ eV, so $K_{\max} = 0.82$ eV and $v = \sqrt{2K/m_e} = \sqrt{2(0.82)(1.602\times10^{-19})/(9.109\times10^{-31})} \approx 5.4\times10^{5}$ m/s.

The photon has momentum as well as energy. Since $E^2 = (pc)^2 + (mc^2)^2$ for any particle, and the photon is massless,

$$E = h\nu = pc \qquad \Longrightarrow \qquad p = \frac{h}{\lambda} = \hbar k,$$

where $k = 2\pi/\lambda$. Together with $E = \hbar\omega$, these are the *Planck–Einstein relations*.

12.4 Compton Scattering

In 1923 Arthur Compton scattered X-rays off nearly free electrons in graphite. He found that the scattered X-rays had a *longer* wavelength than the incident ones, by an amount that depended only on the scattering angle. A classical wave would set the electron oscillating at the incident frequency and re-radiate at that same frequency.

Theorem 12.8 (Compton formula). *A photon of wavelength λ that scatters through angle θ off an electron initially at rest emerges with wavelength*

$$\lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta).$$

The Compton wavelength *of the electron is $\lambda_C = h/m_e c = 2.426 \times 10^{-12}$ m.*

Proof. Treat the collision as one between particles. Let the photon have momenta $\mathbf{p}$ before and $\mathbf{p}'$ after, with $|\mathbf{p}| = h/\lambda$ and $|\mathbf{p}'| = h/\lambda'$. Let the electron recoil with momentum $\mathbf{q}$. Momentum conservation gives $\mathbf{q} = \mathbf{p} - \mathbf{p}'$, so

$$q^2 = p^2 + p'^2 - 2pp'\cos\theta.$$

Energy conservation, using the relativistic energy of the electron, gives $pc + m_e c^2 = p'c + \sqrt{q^2c^2 + m_e^2 c^4}$. Isolate the square root and square both sides:

$$(p - p')^2 c^2 + 2(p - p')m_e c^3 + m_e^2 c^4 = q^2 c^2 + m_e^2 c^4.$$

Substitute q^2 from the momentum equation. The p^2 and p'^2 terms cancel, leaving $-2pp'c^2 + 2(p - p')m_e c^3 = -2pp'c^2\cos\theta$, that is, $(p - p')m_e c = pp'(1 - \cos\theta)$. Divide by $pp'm_e c$ and use $\frac{1}{p'} - \frac{1}{p} = \frac{\lambda' - \lambda}{h}$. ∎

Example 12.9 (Size of the shift). For 0.0710 nm X-rays (molybdenum K_α) scattered at 90°, $\Delta\lambda = \lambda_C = 0.00243$ nm, a 3.4% shift, easily measured. For visible light at 500 nm, the same shift is five parts per million. That is why the effect went unnoticed until X-rays were available: $\Delta\lambda$ is significant only when λ is comparable to λ_C.

The Compton effect showed that photons carry momentum h/λ exactly as particles do, in individual collisions that conserve energy and momentum. Yet the same X-rays diffract from crystals like waves. Light is neither a classical wave nor a classical particle.

12.5 Atoms: Rutherford, Bohr, and the Spectrum of Hydrogen

12.5.1 THE SPECTRUM OF HYDROGEN

A hydrogen gas lamp emits light only at specific wavelengths. In 1885 Balmer found that the visible lines fit $\lambda = B\frac{n^2}{n^2-4}$ for $n = 3, 4, 5, 6$. Rydberg generalized this to

$$\frac{1}{\lambda} = R_\infty\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right), \qquad n_2 > n_1 \geq 1, \tag{12.3}$$

with $R_\infty = 1.0974 \times 10^7\ \mathrm{m}^{-1}$. The Lyman series ($n_1 = 1$) lies in the ultraviolet, the Balmer series ($n_1 = 2$) in the visible, and the Paschen series ($n_1 = 3$) in the infrared. Classical physics offered no explanation at all for why atoms should emit sharp lines, let alone for this formula.

12.5.2 THE RUTHERFORD ATOM AND ITS INSTABILITY

Rutherford's scattering experiments of 1909–1911 showed that an atom's positive charge and almost all its mass are concentrated in a nucleus about 10^{-15} m across, 10^5 times smaller than the atom. The natural picture is a miniature solar system, with electrons orbiting the nucleus under the inverse-square Coulomb force, exactly as in Theorem 2.28. But an accelerating charge radiates. By the Larmor formula, its radiated power is $P = \dfrac{e^2 a^2}{6\pi\varepsilon_0 c^3}$.

Example 12.10 (Classical collapse of hydrogen). Estimate how long a classical electron takes to spiral from $r_0 = 0.0529$ nm into the nucleus.

Solution. For a circular orbit, $a = \frac{e^2}{4\pi\varepsilon_0 m_e r^2}$ and the energy is $E = -\frac{e^2}{8\pi\varepsilon_0 r}$. Energy loss by radiation gives $\mathrm{d}E/\mathrm{d}t = -P$. Since $\mathrm{d}E/\mathrm{d}t = \frac{e^2}{8\pi\varepsilon_0 r^2}\dot{r}$ and $a = r_e c^2/r^2$, where $r_e = \frac{e^2}{4\pi\varepsilon_0 m_e c^2} = 2.82 \times 10^{-15}$ m is the *classical electron radius*, a short computation gives $r^2\dot{r} = -\frac{4}{3}r_e^2 c$. Integrating from r_0 to 0,

$$t = \frac{r_0^3}{4r_e^2 c} = \frac{(5.29\times10^{-11})^3}{4(2.82\times10^{-15})^2(3.00\times10^8)} \approx 1.6\times10^{-11}\ \mathrm{s}.$$

Classically, every atom in the universe should collapse in about sixteen picoseconds, emitting a continuous smear of radiation as it goes. Neither happens.

12.5.3 THE BOHR MODEL

In 1913 Niels Bohr kept the planetary picture but imposed two postulates that contradict classical physics.

(I) An electron can occupy only certain *stationary* orbits, in which it does not radiate. These are the circular orbits whose angular momentum is an integer multiple of $\hbar$: $L = m_e v r = n\hbar$, $n = 1, 2, 3, \dots$

(II) Radiation is emitted or absorbed only when the electron jumps between stationary orbits, as a single photon carrying the energy difference: $h\nu = E_{n_2} - E_{n_1}$.

Theorem 12.11 (Bohr energy levels). *Under Bohr's postulates, the orbits of a hydrogen-like atom with nuclear charge Ze have*

$$r_n = \frac{n^2}{Z} a_0, \qquad E_n = -\frac{Z^2}{n^2}\,\frac{\hbar^2}{2m_e a_0^2} = -\frac{Z^2}{n^2}\,(13.6\ \text{eV}),$$

$$a_0 = \frac{4\pi\varepsilon_0 \hbar^2}{m_e e^2} = 0.0529\ \text{nm}.$$

Proof. For a circular orbit, the Coulomb force supplies the centripetal force: $\frac{m_e v^2}{r} = \frac{Ze^2}{4\pi\varepsilon_0 r^2}$. With $v = \frac{n\hbar}{m_e r}$ from postulate (I), this gives $r = \frac{4\pi\varepsilon_0 n^2\hbar^2}{Z m_e e^2} = \frac{n^2 a_0}{Z}$. The energy is $E = \frac{1}{2}m_e v^2 - \frac{Ze^2}{4\pi\varepsilon_0 r} = -\frac{Ze^2}{8\pi\varepsilon_0 r}$, by the virial relation implied by the force equation. Substituting r_n gives the result, and $\frac{e^2}{8\pi\varepsilon_0 a_0} = \frac{\hbar^2}{2m_e a_0^2} = 13.6$ eV. ∎

Combining with postulate (II), $\frac{1}{\lambda} = \frac{E_{n_2} - E_{n_1}}{hc} = \frac{13.6\ \text{eV}}{hc}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$, which is Rydberg's formula with $R_\infty = \frac{13.6\ \text{eV}}{hc} = 1.097 \times 10^7\ \text{m}^{-1}$. This agreement, with a constant computed entirely from m_e, e, h, and ε_0, was a sensation. The model also explained the spectrum of ionized helium ($Z = 2$), which had been attributed to an unknown element.

Remark 12.12 (Correspondence principle). For large n, the frequency of the photon emitted in the jump $n \to n-1$ approaches the classical orbital frequency of the electron. Quantum predictions must agree with classical physics in the limit where the quantum numbers are large. Bohr used this *correspondence principle* as a guide, and it remains a valuable check on any quantum calculation.

The Bohr model is not quantum mechanics. It cannot explain the intensities of spectral lines, the spectra of atoms with more than one electron, or why the ground state has zero orbital angular momentum, which the full theory shows it

does. Its circular orbits are wrong, and the hydrogen ground state is spherically symmetric (Chapter 16). Remarkably, its energy levels are exactly right. The reason is the special symmetry of the inverse-square force, the same symmetry that conserves the Runge–Lenz vector classically.

12.6 Matter Waves

In his 1924 doctoral thesis, Louis de Broglie reversed Einstein's logic. If waves of light behave like particles with $p = h/\lambda$, perhaps particles of matter behave like waves with

$$\lambda = \frac{h}{p}, \qquad \text{equivalently} \qquad \mathbf{p} = \hbar\mathbf{k}, \quad E = \hbar\omega. \tag{12.4}$$

Bohr's quantization condition then has a simple meaning. An electron wave on a circular orbit of circumference $2\pi r$ must join smoothly onto itself, so $2\pi r = n\lambda = nh/m_e v$, which is $m_e v r = n\hbar$. The allowed orbits are those that hold a whole number of wavelengths: standing waves, like the modes of a string (Section 10.6).

Example 12.13 (Wavelengths of matter). (a) An electron accelerated through $V = 54$ V has $p = \sqrt{2m_e eV}$ and $\lambda = \frac{h}{\sqrt{2m_e eV}} = 0.167$ nm, comparable to the spacing of atoms in a crystal. (b) A 0.145 kg baseball at 40 m/s has $\lambda = 1.1 \times 10^{-34}$ m, twenty orders of magnitude smaller than a proton. Matter waves are invisible at everyday scales because h is so small.

In 1927 Davisson and Germer, scattering 54 eV electrons off a nickel crystal, found a strong reflection peak at $50°$. The Bragg condition for the nickel surface lattice, with spacing 0.215 nm, gives $\lambda = d\sin\theta = (0.215\,\text{nm})\sin 50° = 0.165$ nm, in agreement with de Broglie's 0.167 nm. G. P. Thomson independently observed electron diffraction rings through thin foils. Since then, diffraction has been demonstrated for neutrons, atoms, and molecules as large as C_{60} and beyond. Electron microscopes exploit the short wavelengths of fast electrons to resolve individual atoms.

12.7 The Double-Slit Experiment

Feynman called the double-slit experiment "a phenomenon which is impossible, absolutely impossible, to explain in any classical way, and which has in it the heart of quantum mechanics." We describe it with electrons, for which it has

been performed exactly as described.

A source sends electrons toward a barrier with two narrow slits separated by d. Behind the barrier, a screen a distance $D \gg d$ away records where each electron lands, as a single localized flash. The experiment is run in three ways.

Only one slit open. The flashes pile up into a single broad hump behind the open slit, with probability density $P_1(y)$ or $P_2(y)$.

Both slits open. The flashes build up an *interference pattern* of bright and dark fringes. At the dark fringes almost no electrons arrive, even though electrons do arrive there when either slit alone is open. The pattern is

$$P_{12}(y) \neq P_1(y) + P_2(y).$$

Opening a second path can *reduce* the probability of arriving at a point. The fringe spacing is $\Delta y = \lambda D/d$ with $\lambda = h/p$, as for light.

One electron at a time. When the source is so weak that only one electron is in the apparatus at any moment, each electron still lands at a single point, apparently at random. But after thousands of electrons, the interference pattern emerges. Each electron interferes with itself. The experiment was performed this way by Tonomura and collaborators in 1989.

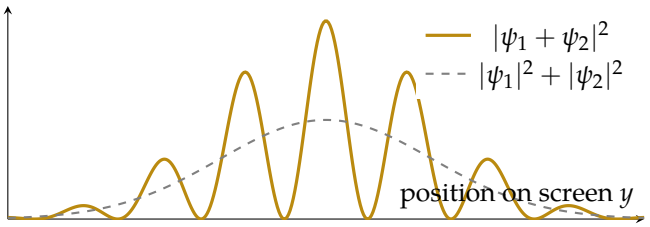


Figure 12.2: The two-slit pattern (gold) compared with the sum of the one-slit patterns (dashed). The difference is the interference term $2\,\mathrm{Re}(\psi_1^*\psi_2)$.

The pattern is explained, perfectly, by a rule that no classical theory contains. Associate with each way the event can happen a *complex amplitude*: $\psi_1(y)$ for passage through slit 1 and $\psi_2(y)$ for slit 2. The amplitude for the event is the sum of the amplitudes for the alternatives, and the probability is the squared modulus of the total:

$$P_{12} = |\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + 2\,\mathrm{Re}(\psi_1^*\psi_2) = P_1 + P_2 + 2\sqrt{P_1 P_2}\cos\varphi,$$

where φ is the relative phase of the two amplitudes, which varies with y. This is exactly the phasor addition of Example 8.10. The interference term can be negative, producing dark fringes.

Watching the slits. Suppose a light source is placed near the slits so that each electron, as it passes, scatters a photon that reveals which slit it went through. Then the interference pattern disappears, and $P = P_1 + P_2$. If the light is dimmed so that some electrons pass undetected, those electrons show interference and the detected ones do not. If the wavelength of the light is lengthened to disturb the electrons less, the interference returns at exactly the point where the photons can no longer resolve which slit was used.

The general principle, which the formalism of Chapter 15 makes precise, is this. *When alternatives are indistinguishable in principle, add amplitudes; when they are distinguishable, even in principle, add probabilities.* Interference is lost not because the measurement is clumsy, but because the electron has become correlated with the photon. The two alternatives now lead to orthogonal states of the photon, and the cross term vanishes, as we will see in Chapter 18.

12.8 The Stern–Gerlach Experiment

A current loop, or a spinning charge, has a magnetic dipole moment $\boldsymbol{\mu}$, with energy $U = -\boldsymbol{\mu}\cdot\mathbf{B}$ in a magnetic field. In a uniform field the dipole feels a torque but no net force. In a nonuniform field there is a force $\mathbf{F} = -\nabla U = \nabla(\boldsymbol{\mu}\cdot\mathbf{B})$. If $\mathbf{B}$ points mainly along z and varies with z, then $F_z \approx \mu_z\,\partial B_z/\partial z$. The deflection measures μ_z.

In 1922 Otto Stern and Walther Gerlach sent a beam of silver atoms, evaporated from an oven, through such a field. The atoms' magnetic moments come from a single unpaired electron and are oriented at random as they leave the oven. Classically, $\mu_z = |\boldsymbol{\mu}|\cos\alpha$ should take every value between $-|\boldsymbol{\mu}|$ and $+|\boldsymbol{\mu}|$, and the beam should smear into a continuous band. Instead it split into exactly two spots. The component μ_z takes only two values, $\pm\mu_B$, where $\mu_B = \frac{e\hbar}{2m_e} = 9.27\times 10^{-24}\,\text{J/T}$ is the *Bohr magneton*. The electron has an intrinsic angular momentum, its *spin*, with $S_z = \pm\hbar/2$ only.

12.8.1 Sequential Measurements

The experiment becomes a laboratory for the logic of quantum mechanics when Stern–Gerlach devices are placed in sequence. Write SGz for a device that splits the beam along z, and $z+$ and $z-$ for its two output beams.

(1) *SGz, keep $z+$, then SGz again*: all atoms emerge $z+$. A measurement repeated immediately gives the same result.

(2) *SGz, keep $z+$, then SGx*: half emerge $x+$ and half $x-$. A state with definite

S_z has no definite S_x.

(3) *SGz, keep z+; SGx, keep x+; then SGz*: half emerge $z+$ and half $z-$. The $z-$ component, removed by the first filter, has *reappeared*. Measuring S_x destroyed the information about S_z.

Experiment (3) cannot be explained by atoms carrying hidden values of both S_z and S_x that the devices merely reveal. If the first filter selected atoms with $S_z = +\frac{\hbar}{2}$, a later filter on S_x could not restore atoms with $S_z = -\frac{\hbar}{2}$. The formalism of Chapter 9 describes these results exactly. States are unit vectors in $\mathbb{C}^2$; $z\pm$ are the eigenvectors $|\uparrow\rangle, |\downarrow\rangle$ of σ_z; $x\pm$ are the eigenvectors $|\pm\rangle$ of σ_x; and the probability of passing a filter is the squared inner product. Since $|\langle\pm|\uparrow\rangle|^2 = \frac{1}{2}$, experiments (2) and (3) follow at once. Because σ_x and σ_z do not commute, they have no common eigenvectors (Theorem 9.30). Measuring one disturbs the other.

12.9 Toward a New Theory

By 1925 the experimental facts demanded a theory with these features:

- energy and angular momentum of bound systems are quantized, though not always (free particles have continuous energies);
- radiation and matter both show wave and particle properties, linked by $E = \hbar\omega$ and $\mathbf{p} = \hbar\mathbf{k}$;
- outcomes of individual measurements are random, but their probabilities are computed by squaring the modulus of complex amplitudes that add when alternatives are indistinguishable;
- measurement of one quantity can disturb another, and some pairs of quantities cannot be sharp simultaneously.

Two equivalent theories appeared within a year. Heisenberg, Born, and Jordan built *matrix mechanics* in 1925, and Schrödinger built *wave mechanics* in 1926. Dirac and von Neumann showed that both are representations of one structure, linear algebra in Hilbert space, which we built in Chapter 9. The next chapter begins with Schrödinger's equation for the de Broglie wave.

12.10 Worked Problems

The problems below are quantitative. Each takes one of the experiments of this chapter and extracts a number that can be checked against measurement, or estimates what classical

physics predicted and how badly it failed. Constants are taken from Appendix A.

Problem 12.1 Photons in the cosmic background

Find the number density of photons in blackbody radiation at temperature T, and their mean energy. Evaluate both for the cosmic microwave background, $T = 2.725$ K.

Solution. A mode of frequency ν holds on average $\langle E\rangle/h\nu = 1/(e^{h\nu/k_BT}-1)$ photons, by Theorem 12.3. With the mode density of Theorem 12.2, the photon number density is

$$n = \int_0^\infty \frac{8\pi\nu^2}{c^3}\,\frac{d\nu}{e^{h\nu/k_BT}-1} = 8\pi\Big(\frac{k_BT}{hc}\Big)^3\int_0^\infty \frac{x^2\,dx}{e^x-1}.$$

Expanding as in Corollary 12.5, $\int_0^\infty \frac{x^2}{e^x-1}dx = \sum_{n\ge1}\frac{2}{n^3} = 2\zeta(3) \approx 2.404$, where $\zeta(3)\approx 1.2021$ is Apéry's constant. So

$$n = 16\pi\zeta(3)\Big(\frac{k_BT}{hc}\Big)^3 \approx 2.03\times10^7\,T^3\ \text{m}^{-3} \quad (T \text{ in kelvin}).$$

The mean energy per photon is $u/n = \frac{\pi^4/15}{2\zeta(3)}k_BT \approx 2.70\,k_BT$.

For the CMB, $n \approx 2.03\times10^7\times(2.725)^3 \approx 4.11\times10^8\ \text{m}^{-3}$, or about 411 photons per cubic centimeter, each with mean energy $2.70\times(2.35\times10^{-4}\ \text{eV}) \approx 6.3\times10^{-4}$ eV. There are roughly 10^9 CMB photons for every proton in the universe. That ratio, fixed since the first seconds after the Big Bang, is one of the basic numbers of cosmology.

Problem 12.2 The recoil direction in Compton scattering

A photon of energy E scatters from a free electron at rest through angle θ. Show that the electron recoils at angle ϕ below the incident direction, where

$$\cot\phi = \Big(1+\frac{E}{m_ec^2}\Big)\tan\frac{\theta}{2}.$$

Evaluate ϕ and the scattered wavelength for $\lambda = 0.0500$ nm and $\theta = 90^\circ$.

Solution. Let the photon arrive along x with momentum $p = E/c$, and leave with momentum p' at angle θ above the axis. Momentum conservation gives the electron's momentum components: $p_ex = p - p'\cos\theta$ and $p_ey = -p'\sin\theta$. So

$$\tan\phi = \frac{p'\sin\theta}{p-p'\cos\theta} = \frac{\sin\theta}{p/p'-\cos\theta}.$$

By the Compton formula (Theorem 12.8) in the form $\frac{p}{p'} = \frac{\lambda'}{\lambda} = 1 + \frac{E}{m_e c^2}(1 - \cos\theta)$,

$$\frac{p}{p'} - \cos\theta = (1 - \cos\theta)\left(1 + \frac{E}{m_e c^2}\right).$$

Using $\frac{1-\cos\theta}{\sin\theta} = \tan\frac{\theta}{2}$, $\cot\phi = \left(1 + \frac{E}{m_e c^2}\right)\tan\frac{\theta}{2}$.

For $\lambda = 0.0500$ nm, $E = hc/\lambda = 24.80$ keV and $E/m_e c^2 = 0.0485$. At $\theta = 90^\circ$ the wavelength increases by exactly one Compton wavelength, $\lambda_C = h/m_e c = 2.43$ pm, to $\lambda' = 0.0524$ nm. And $\cot\phi = 1.0485$, so $\phi = 43.6^\circ$. As $E/m_e c^2 \to 0$, $\phi \to 90^\circ - \frac{\theta}{2}$, the result for elastic scattering from an infinitely heavy target. The electron always recoils forward ($\phi < 90^\circ$), and is knocked straight ahead when the photon bounces straight back.

Problem 12.3 Discovering deuterium

The Rydberg constant for a nucleus of finite mass M is $R_M = R_\infty \frac{\mu}{m_e}$, where $\mu = \frac{m_e M}{m_e + M}$ is the reduced mass (Theorem 12.11). Find the difference between the Hα wavelengths ($n = 3 \to 2$) of hydrogen and deuterium.

Solution. For a given transition $\lambda \propto 1/R_M \propto 1/\mu$, so $\frac{\lambda_H - \lambda_D}{\lambda_H} = 1 - \frac{\mu_H}{\mu_D}$. Since $m_e \ll M$, expand $\mu = m_e(1 - \frac{m_e}{M} + \cdots)$:

$$\frac{\Delta\lambda}{\lambda} \approx \frac{m_e}{m_p} - \frac{m_e}{m_d} \approx \frac{m_e}{2m_p} = \frac{1}{3672} = 2.72 \times 10^{-4},$$

using $m_d \approx 2m_p$. With $\lambda_H = 656.47$ nm (in vacuum, from $\frac{1}{\lambda} = R_H(\frac{1}{4} - \frac{1}{9})$), $\Delta\lambda \approx 0.179$ nm. Deuterium's line is at the shorter wavelength, since its reduced mass is closer to m_e and its levels are slightly deeper. A full calculation with the measured masses gives 0.1786 nm.

In 1931 Harold Urey found faint companion lines displaced by exactly this amount in the spectrum of hydrogen from evaporated liquid hydrogen, enriched in the heavy isotope. That was the discovery of deuterium, which won him the 1934 Nobel Prize in Chemistry. The Bohr model's reduced-mass correction, a refinement that seems academic, was precise enough to identify a new isotope.

Problem 12.4 Positronium

Positronium is a bound state of an electron and a positron, which have equal masses. Find its Bohr radius, ground-state energy, and the wavelength of its Lyman-α line ($n = 2 \to 1$).

Solution. The reduced mass is $\mu = \frac{m_e m_e}{2m_e} = \frac{m_e}{2}$. In the Bohr model, lengths scale as $1/\mu$ and energies as μ, so

$$a_{\mathrm{Ps}} = 2a_0 = 0.106\ \mathrm{nm}, \qquad E_1 = \frac{-13.6\ \mathrm{eV}}{2} = -6.80\ \mathrm{eV}.$$

The Lyman-α energy is $6.80\left(1 - \frac{1}{4}\right) = 5.10$ eV, so $\lambda = \frac{hc}{5.10\ \mathrm{eV}} = 243$ nm, twice the hydrogen value of 121.5 nm.

Here a_{Ps} is the mean separation of the two particles, each of which orbits the center of mass at distance a_0. Positronium is unstable: the electron and positron annihilate into gamma rays, in 0.125 ns or 142 ns depending on whether their spins are antiparallel or parallel (Chapter 17). Its spectrum, measured to parts per billion, is one of the most stringent tests of quantum electrodynamics, because it involves no nuclear structure at all.

Problem 12.5 The missing delay in the photoelectric effect

Light of intensity $I = 1.0 \times 10^{-2}\ \mathrm{W/m^2}$ (a dim lamp) falls on potassium, with work function 2.30 eV. Suppose that, as classical wave theory implies, an electron gathers energy continuously from the wave over an area comparable to an atom, $A \approx 1 \times 10^{-19}\ \mathrm{m^2}$. Estimate the time before the first electron could be ejected. Compare with experiment.

Solution. The power falling on the area is $IA = 1.0 \times 10^{-21}$ W. To accumulate $W = 2.30\ \mathrm{eV} = 3.7 \times 10^{-19}$ J takes

$$t \approx \frac{W}{IA} = \frac{3.7 \times 10^{-19}\ \mathrm{J}}{1.0 \times 10^{-21}\ \mathrm{W}} \approx 370\ \mathrm{s},$$

about six minutes, and longer still if some of the absorbed energy is lost in the meantime. Experiments find photoelectrons within 10^{-9} s of switching on the light, even at intensities far below this one. Lower intensity reduces the *number* of photoelectrons, but the first appear immediately.

The photon picture explains this at once: the energy arrives in lumps $h\nu$, and an electron that absorbs one photon receives the whole amount instantly. At low intensity, photons arrive rarely, but each one that arrives can eject an electron. The estimate shows that the failure of the wave picture is not a subtle effect of a few percent. It is wrong by eleven orders of magnitude.

Problem 12.6 The Davisson–Germer experiment

In 1927 Davisson and Germer directed 54 eV electrons at a nickel crystal whose surface atoms form rows spaced $d = 0.215$ nm apart. Find the de Broglie wavelength and the angle of the first diffraction maximum, $d \sin\theta = \lambda$.

Solution. The electron is nonrelativistic ($54\text{ eV} \ll m_e c^2 = 511\text{ keV}$), so $p = \sqrt{2m_e K}$ and

$$\lambda = \frac{h}{\sqrt{2m_e K}} = \frac{6.626 \times 10^{-34}}{\sqrt{2(9.109 \times 10^{-31})(54 \times 1.602 \times 10^{-19})}}\ \text{m} = 0.167\ \text{nm}.$$

A convenient form is $\lambda \approx \frac{1.226}{\sqrt{K/\text{eV}}}$ nm. Then $\sin\theta = \frac{0.167}{0.215} = 0.776$, so $\theta = 50.9^\circ$. Davisson and Germer observed a sharp peak in the scattered intensity at 50°.

The two key numbers, $\lambda \approx 0.17$ nm and $d \approx 0.2$ nm, are comparable. That is why the effect was visible: a wavelength much smaller than the lattice spacing would give maxima at tiny angles, and one much larger would give none ($\sin\theta > 1$). The experiment was in part an accident: the crystal formed after the nickel target was heated to repair damage from a broken vacuum tube. The same principle now underlies electron microscopy and low-energy electron diffraction (LEED), the standard tool for studying crystal surfaces.

Problem 12.7 Zero-point energy from the uncertainty principle

For a harmonic oscillator, $E = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$. Using $\Delta x\,\Delta p \geq \frac{\hbar}{2}$, find a lower bound for the energy.

Solution. For a state centered at the origin, $\langle x\rangle = \langle p\rangle = 0$ by symmetry, so $\langle x^2\rangle = (\Delta x)^2$ and $\langle p^2\rangle = (\Delta p)^2$. Then

$$\langle E\rangle = \frac{(\Delta p)^2}{2m} + \frac{1}{2}m\omega^2(\Delta x)^2 \geq \frac{\hbar^2}{8m(\Delta x)^2} + \frac{1}{2}m\omega^2(\Delta x)^2.$$

Minimize over $s = (\Delta x)^2$: the derivative $-\frac{\hbar^2}{8ms^2} + \frac{1}{2}m\omega^2$ vanishes at $s = \frac{\hbar}{2m\omega}$, where the bound equals $\frac{\hbar\omega}{4} + \frac{\hbar\omega}{4} = \frac{1}{2}\hbar\omega$. (Alternatively, AM–GM gives $\frac{\hbar^2}{8ms} + \frac{m\omega^2 s}{2} \geq 2\sqrt{\frac{\hbar^2\omega^2}{16}} = \frac{\hbar\omega}{2}$.)

So $\langle E\rangle \geq \frac{1}{2}\hbar\omega$. In Chapter 14 the exact ground-state energy will turn out to be $\frac{1}{2}\hbar\omega$: this estimate is not just of the right order but exact. The reason is that the ground state is a Gaussian, which saturates the uncertainty relation (Theorem 11.20). The oscillator can never be at rest at the bottom of its well. This *zero-point energy* keeps helium liquid at absolute zero under ordinary pressure.

Problem 12.8 Frozen vibrations

Classically, a diatomic molecule has 3 translational, 2 rotational, and 1 vibrational degree of freedom, the last contributing k_B (kinetic plus potential) to the heat capacity. So equipartition predicts $C_V = \frac{7}{2}k_B$ per molecule. For hydrogen the

measured value at 300 K is $\frac{5}{2}k_B$. Using Planck's mean energy for the vibration, with $h\nu/k_B = 6000$ K, explain the discrepancy.

Solution. Differentiate $\langle E\rangle = \frac{h\nu}{e^{h\nu/k_BT}-1}$ with respect to T. With $x = h\nu/k_BT$,

$$C_{\text{vib}} = \frac{\mathrm{d}\langle E\rangle}{\mathrm{d}T} = k_B\,\frac{x^2e^x}{(e^x-1)^2}.$$

This tends to k_B as $x \to 0$ (high temperature, equipartition) and to $k_Bx^2e^{-x}$ as $x \to \infty$. At 300 K, $x = 20$, so $C_{\text{vib}} \approx k_B \cdot 400e^{-20} \approx 8\times10^{-7}k_B$. The vibration contributes nothing measurable, and $C_V = \frac{3}{2}k_B + k_B = \frac{5}{2}k_B$, as observed. At 3000 K, $x = 2$ and $C_{\text{vib}} \approx 0.72k_B$: the vibration is thawing, and C_V rises toward $\frac{7}{2}k_B$.

This was one of the "clouds" over classical physics identified by Kelvin in 1900. Equipartition is not wrong, but it applies only to modes whose quantum $h\nu$ is small compared with k_BT. The rotations of hydrogen freeze out in the same way below about 100 K, by the quantization of angular momentum studied in Chapter 17.

EXERCISES

A. COMPUTATION

12.1 Compute the energy, in eV, and the momentum of a photon of wavelength 650 nm (red light) and of 0.10 nm (X-ray).

12.2 A 100 W light bulb emits about 5% of its power as visible light near 550 nm. How many visible photons does it emit per second?

12.3 Using Wien's law, find the peak wavelength for a human body at 310 K. In what part of the spectrum is it?

12.4 The Sun has radius 6.96×10^8 m and surface temperature 5800 K. Use the Stefan–Boltzmann law to find its luminosity, and the solar power per unit area at the Earth, 1.50×10^{11} m away.

12.5 Verify that the root of $3(1-e^{-x}) = x$ is $x \approx 2.821$, and find the root of the corresponding equation $5(1-e^{-x}) = x$ for the wavelength peak.

12.6 Show that Planck's law reduces to the Rayleigh–Jeans law for $h\nu \ll k_BT$ and to Wien's approximation $u \propto \nu^3e^{-h\nu/k_BT}$ for $h\nu \gg k_BT$.

12.7 The work function of potassium is 2.30 eV. Find the threshold wavelength and the stopping potential for 300 nm light.

12.8 In a photoelectric experiment, the stopping potentials are 1.48 V at $\lambda = 366$ nm and 0.62 V at $\lambda = 492$ nm. Find h/e and the work function.

12.9 X-rays of wavelength 0.0500 nm scatter from free electrons at 120°. Find the scattered wavelength and the kinetic energy of the recoil electron.

12.10 What is the maximum kinetic energy a free electron can receive from a 1.00 MeV gamma ray in Compton scattering?

12.11 Find the wavelengths of the first three Balmer lines and the series limit.

12.12 Find the ionization energy of He^+ and of Li^{2+} from the Bohr model.

12.13 Find the speed of the electron in the $n = 1$ Bohr orbit, and show that $v_1/c = \alpha$, the *fine-structure constant* $\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \approx \frac{1}{137}$.

12.14 Find the orbital period of the electron in the $n = 1$ Bohr orbit and the number of orbits it completes in the classical collapse time of Example 12.10.

12.15 Muonic hydrogen replaces the electron by a muon, 207 times heavier. Find its Bohr radius and ground-state energy. (Use the reduced mass.)

12.16 Find the de Broglie wavelength of a thermal neutron at 300 K ($K = \frac{3}{2}k_BT$), of a 1 keV electron, and of a nitrogen molecule at room temperature.

12.17 An electron microscope uses 100 keV electrons. Find their wavelength, (a) nonrelativistically and (b) relativistically.

12.18 In a double-slit experiment with 50 eV electrons, slit separation 1.0 μm, and screen distance 1.0 m, find the fringe spacing.

12.19 In Fig. 12.2, suppose $|\psi_1| = |\psi_2|$ everywhere. What are the maximum and minimum of $P_{12}/(P_1 + P_2)$?

12.20 Find the force on a silver atom with $\mu_z = \mu_B$ in a field gradient of 10 T/m, and its deflection after a 0.10 m magnet if the atoms move at 500 m/s. The silver atom has mass 1.79×10^{-25} kg.

12.21 A beam passes SGz (keep $z+$), then an SG device along the direction at angle θ from z in the xz-plane (keep $+$). Given that the probability is $\cos^2(\theta/2)$, what fraction emerges? Evaluate at $\theta = 60^\circ$ and $\theta = 180^\circ$.

B. THEORY AND PROOF

12.22 Derive the Stefan–Boltzmann factor $c/4$ by integrating the energy flux of isotropic radiation through a small hole over a hemisphere.

12.23 Show that the classical partition function of a harmonic oscillator, $\frac{1}{h}\iint e^{-\beta E(x,p)}\,dx\,dp$, gives mean energy k_BT. Show that Planck's sum gives the same result as $h \to 0$.

12.24 Prove that a free electron cannot absorb a photon completely. (Use energy and momentum conservation.) Why, then, can a bound electron absorb one in the photoelectric effect?

12.25 Show that for large n the frequency of the photon emitted in the transition $n \to n-1$ in hydrogen approaches the classical orbital frequency $v_n/(2\pi r_n)$ (the correspondence principle).

12.26 Show that Bohr's quantization $L = n\hbar$ for a particle in a circular orbit is equivalent to $\oint p\,dq = nh$ around the orbit (the Wilson–Sommerfeld rule). Apply the rule to a harmonic oscillator to find $E_n = nh\nu$, and compare with Planck.

12.27 Show that the de Broglie group velocity $\mathrm{d}\omega/\mathrm{d}k$, with $E = \hbar\omega$ and $p = \hbar k$, equals the particle velocity both nonrelativistically and relativistically.

C. CHALLENGE

12.28 (Einstein's heat capacity of solids.) Model a solid as $3N$ independent oscillators of frequency ν with Planck's mean energy. Find the heat capacity $C(T)$ and show that it tends to $3Nk_B$ (the Dulong–Petit law) at high T and vanishes exponentially as $T \to 0$.

12.29 Show that in the Bohr model the magnetic moment of the electron in orbit n is $\mu = n\mu_B$, by computing the current loop's moment IA.

12.30 Using $\Delta x\,\Delta p \gtrsim \hbar$, estimate the minimum energy of an electron confined within $r \sim a$ of a proton by minimizing $E(a) = \frac{\hbar^2}{2m_e a^2} - \frac{e^2}{4\pi\varepsilon_0 a}$. Compare with the Bohr result.

12.31 The Wien displacement law was known before Planck. Show, using only dimensional analysis and the assumption that $u(\nu, T)$ depends on ν, T, c, k_B and one unknown constant with dimensions of action, that $u = \nu^3 F(\nu/T)$ for some function F, and deduce $\lambda_{\max}T = \text{constant}$.

CHAPTER 13

The Wavefunction and the Schrödinger Equation

LEARNING OBJECTIVES

- Motivate the Schrödinger equation from de Broglie waves, and explain why the wavefunction must be complex.
- State Born's statistical interpretation, normalize wavefunctions, and compute probabilities.
- Derive the continuity equation for probability and prove that normalization is preserved in time.
- Compute expectation values of position, momentum, and energy using operators, in position and momentum space.
- Prove Ehrenfest's theorem and explain the classical limit.
- State the position–momentum uncertainty principle and use it to make physical estimates.
- Solve exactly for the spreading of a free Gaussian wave packet.
- Separate variables to obtain stationary states, and expand general solutions in energy eigenstates.

In Chapter 12 we saw that electrons interfere like waves, are detected like particles, and obey a rule in which complex amplitudes add and probabilities are their squared moduli. In 1926 Erwin Schrödinger wrote down the equation that governs how these amplitudes evolve. This chapter introduces that equation, the meaning of its solutions, and the mathematical properties that make the theory consistent. We work mostly with a single particle moving in one dimension, where every idea appears in its simplest form. The extension to three dimensions changes the notation, not the ideas.

13.1 The Schrödinger Equation

A free particle with definite momentum $p = \hbar k$ and energy $E = \hbar\omega = \frac{p^2}{2m}$ should be described by the de Broglie plane wave

$$\Psi(x,t) = Ae^{i(kx-\omega t)}.$$

Differentiating,

$$-i\hbar\frac{\partial\Psi}{\partial x} = \hbar k\,\Psi = p\Psi, \qquad i\hbar\frac{\partial\Psi}{\partial t} = \hbar\omega\,\Psi = E\Psi.$$

So on plane waves, momentum acts as the operator $\hat{p} = -i\hbar\,\partial_x$ and energy as $i\hbar\,\partial_t$. The classical relation $E = \frac{p^2}{2m} + V$ then suggests an equation that holds for every superposition of plane waves, and that includes a potential.

Definition 13.1 (Schrödinger equation). The wavefunction $\Psi(x,t)$ of a particle of mass m in a potential $V(x)$ obeys

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V(x)\Psi \equiv \hat{H}\Psi. \tag{13.1}$$

In three dimensions, $i\hbar\,\partial_t\Psi = -\frac{\hbar^2}{2m}\nabla^2\Psi + V(\mathbf{r})\Psi$. The operator $\hat{H} = \frac{\hat{p}^2}{2m} + V$ is the *Hamiltonian*.

This is a motivation, not a derivation. The Schrödinger equation is a postulate, justified by its agreement with experiment, and that agreement has been total within its domain of nonrelativistic physics. Three structural features deserve comment.

It is linear. If Ψ_1 and Ψ_2 are solutions, so is $c_1\Psi_1 + c_2\Psi_2$. The *superposition principle* of quantum mechanics is the statement that solutions form a vector space (Example 9.2). Interference follows.

It is first order in time. The present wavefunction $\Psi(x,0)$ determines the whole future, by the uniqueness theorem (Theorem 10.1) extended to this setting. The dynamics of the wavefunction is deterministic. The randomness of quantum mechanics enters only through measurement.

It contains i, and Ψ must be complex. Compare the heat equation $\partial_t u = k\,\partial_x^2 u$, which is also first order in time. Its solutions decay and spread irreversibly, and the total $\int u^2$ decreases. The factor of i is what makes the Schrödinger equation describe waves that propagate without loss, and conserve $\int|\Psi|^2$ (Theorem 13.5). A real-valued Ψ cannot satisfy the equation except in the trivial case: if V is real

and Ψ real, the left side of (13.1) is imaginary and the right side real, so both must vanish.

13.2 The Statistical Interpretation

What does Ψ mean? Schrödinger first hoped it described a smeared-out electron, with charge density proportional to $|\Psi|^2$. But an electron is always detected whole, at one point, as the double-slit experiment shows. In 1926 Max Born proposed the interpretation that has been accepted ever since.

Definition 13.2 (Born's rule). $|\Psi(x,t)|^2$ is the probability density for finding the particle at position x at time t:

$$P(a \le x \le b \text{ at time } t) = \int_a^b |\Psi(x,t)|^2\,dx.$$

Consequently Ψ must be *normalized*: $\int_{-\infty}^{\infty} |\Psi(x,t)|^2 dx = 1$.

Since the Schrödinger equation is linear, multiplying a solution by a constant gives a solution. So any Ψ with $0 < \int |\Psi|^2 < \infty$ can be normalized by rescaling. Functions with infinite norm, such as plane waves, cannot represent physical states on their own. Multiplying Ψ by a constant *phase* $e^{i\alpha}$ changes nothing observable: the *global phase* is unphysical. *Relative* phases between components of a superposition are physical, and they produce interference.

Example 13.3 (Normalizing a wavefunction). At $t = 0$ a particle has $\Psi(x,0) = Ax(a-x)$ for $0 \le x \le a$ and $\Psi = 0$ elsewhere. Normalize it, and find the probability that the particle is in the left quarter of the interval.

Solution. $1 = |A|^2 \int_0^a x^2(a-x)^2 dx = |A|^2 \frac{a^5}{30}$, so $A = \sqrt{30/a^5}$. Then $P(0 \le x \le a/4) = \frac{30}{a^5}\int_0^{a/4} x^2(a-x)^2 dx = 30\left(\frac{1}{192} - \frac{1}{512} + \frac{1}{5120}\right) = \frac{53}{512} \approx 0.104$.

13.2.1 What Happens at a Measurement

If a measurement finds the particle at x_0, an immediately repeated measurement finds it at (or extremely near) x_0 again. So right after the measurement, the wavefunction must be sharply peaked at x_0, whatever it was before. In the standard *Copenhagen interpretation*, measurement causes the wavefunction to *collapse* to a

state consistent with the result. Between measurements, Ψ evolves smoothly by the Schrödinger equation. At a measurement it jumps, randomly, with probabilities given by Born's rule.

Whether collapse is a physical process, an update of knowledge, or an approximate description of the interaction of the system with a macroscopic apparatus is a question of interpretation. It is still debated, and the debate includes the many-worlds, pilot-wave, and spontaneous-collapse alternatives. All standard interpretations make the same predictions for every experiment performed to date. In this book we use the operational rules, which are not in dispute, and return to what they imply about correlations in Chapter 18.

13.3 Conservation of Probability

Born's rule is consistent only if a normalized wavefunction stays normalized as it evolves. That is not automatic. It is a property of the Schrödinger equation that must be proved.

Theorem 13.4 (Continuity equation for probability). *If Ψ satisfies* (13.1) *with real V, then $\rho = |\Psi|^2$ and the* probability current

$$J(x,t) = \frac{\hbar}{2mi}\left(\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x}\right) = \frac{\hbar}{m}\,\mathrm{Im}\left(\Psi^* \frac{\partial \Psi}{\partial x}\right)$$

satisfy $\dfrac{\partial \rho}{\partial t} + \dfrac{\partial J}{\partial x} = 0$. *In three dimensions,* $\mathbf{J} = \frac{\hbar}{m}\,\mathrm{Im}(\Psi^* \nabla \Psi)$ *and* $\partial_t \rho + \boldsymbol{\nabla} \cdot \mathbf{J} = 0$.

Proof. From (13.1) and its complex conjugate (using $V^* = V$),

$$\frac{\partial \Psi}{\partial t} = \frac{i\hbar}{2m}\Psi'' - \frac{i}{\hbar}V\Psi, \qquad \frac{\partial \Psi^*}{\partial t} = -\frac{i\hbar}{2m}\Psi^{*\prime\prime} + \frac{i}{\hbar}V\Psi^*.$$

Then $\partial_t(\Psi^*\Psi) = \Psi^* \partial_t \Psi + \Psi\, \partial_t \Psi^* = \frac{i\hbar}{2m}(\Psi^*\Psi'' - \Psi\Psi^{*\prime\prime})$. The potential terms cancel. The bracket is the derivative of $\Psi^*\Psi' - \Psi\Psi^{*\prime}$, so $\partial_t \rho = -\partial_x\left[\frac{\hbar}{2mi}(\Psi^*\Psi' - \Psi\Psi^{*\prime})\right] = -\partial_x J$. ∎

Theorem 13.5 (Normalization is preserved). *If $\Psi \to 0$ as $|x| \to \infty$ fast enough that $J \to 0$, then* $\mathrm{d}/\mathrm{d}t \int_{-\infty}^{\infty} |\Psi|^2 dx = 0$.

Proof. $\frac{\mathrm{d}}{\mathrm{d}t}\int \rho\, dx = -\int \partial_x J\, dx = -\big[J\big]_{-\infty}^{\infty} = 0$. ∎

This is the local conservation law of Theorem 7.24, with probability as the conserved quantity. Probability does not merely stay constant in total. It flows continuously from place to place, with current J. Equivalently, time evolution is unitary (Theorem 9.31).

Example 13.6 (Current of a plane wave). For $\Psi = Ae^{i(kx-\omega t)}$, $\Psi^*\Psi' = ik|A|^2$, so $J = \frac{\hbar k}{m}|A|^2 = v\rho$. The current is the density times the classical velocity $v = p/m$, exactly as for a classical fluid. For any wavefunction that is real up to a constant phase, $J = 0$. Stationary bound states (Section 13.8) carry no current.

13.4 Expectation Values and Operators

Because $|\Psi|^2$ is a probability density, the average result of measuring position on many identically prepared particles is

$$\langle x\rangle = \int_{-\infty}^{\infty} x\,|\Psi(x,t)|^2\,dx, \qquad \langle f(x)\rangle = \int f(x)|\Psi|^2 dx.$$

The *expectation value* $\langle x\rangle$ is not the most probable position, and not the position of any one particle. It is the mean over an ensemble. The *uncertainty* is the standard deviation, $\sigma_x = \sqrt{\langle x^2\rangle - \langle x\rangle^2}$.

What is the average momentum? There is no "momentum density" in x to integrate against. Instead, compute how fast $\langle x\rangle$ moves.

Theorem 13.7 (Velocity of the mean). $m\dfrac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} = \displaystyle\int_{-\infty}^{\infty} \Psi^*\Big(-i\hbar\frac{\partial}{\partial x}\Big)\Psi\,dx.$

Proof. Using the continuity equation and integrating by parts,

$$\frac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} = \int x\,\partial_t\rho\,dx = -\int x\,\partial_x J\,dx = \int J\,dx = \frac{\hbar}{2mi}\int(\Psi^*\Psi' - \Psi\Psi^{*\prime})\,dx.$$

Integrate the second term by parts once more: $-\int \Psi\Psi^{*\prime}dx = \int \Psi^*\Psi' dx$. So $\mathrm{d}\langle x\rangle/\mathrm{d}t = \frac{\hbar}{mi}\int \Psi^*\Psi' dx$, which is the claim. ∎

Classically $p = m\ \mathrm{d}x/\mathrm{d}t$, which motivates the definition

$$\langle p\rangle = \int \Psi^*\Big(-i\hbar\frac{\partial}{\partial x}\Big)\Psi\,dx.$$

Position is represented by the operator $\hat{x}$, multiplication by x, and momentum by $\hat{p} = -i\hbar\,\partial_x$. The expectation value of any *observable* $Q(x,p)$ is obtained by substituting the operators and sandwiching:

$$\langle Q\rangle = \int \Psi^*\,\hat{Q}\,\Psi\,dx = \langle \Psi|\hat{Q}|\Psi\rangle. \tag{13.2}$$

For example, the kinetic energy is $\langle T\rangle = -\frac{\hbar^2}{2m}\int \Psi^*\Psi''dx$, and the total energy is $\langle H\rangle = \int \Psi^*\hat{H}\Psi\,dx$. Because $\hat{x}$ and $\hat{p}$ are Hermitian (Theorem 9.35), these expectation values are real.

Example 13.8 (A Gaussian state). Let $\Psi(x,0) = (2a/\pi)^{1/4}e^{-ax^2+ik_0x}$ with $a>0$. Find $\langle x\rangle$, $\langle x^2\rangle$, $\langle p\rangle$, $\langle p^2\rangle$, and $\sigma_x\sigma_p$.

Solution. $|\Psi|^2 = \sqrt{2a/\pi}\,e^{-2ax^2}$ is a normalized Gaussian with $\langle x\rangle = 0$ and $\langle x^2\rangle = \frac{1}{4a}$, by Example 5.34. Next, $\hat{p}\Psi = -i\hbar(-2ax+ik_0)\Psi = \hbar(k_0 + 2iax)\Psi$, so $\langle p\rangle = \hbar k_0 + 2ia\hbar\langle x\rangle = \hbar k_0$. Applying $\hat{p}$ again, $\hat{p}^2\Psi = \hbar^2\big[(k_0+2iax)^2 + 2a\big]\Psi$. Taking the expectation, the odd terms vanish and $\langle x^2\rangle = \frac{1}{4a}$, so

$$\langle p^2\rangle = \hbar^2\big(k_0^2 - 4a^2\cdot\tfrac{1}{4a} + 2a\big) = \hbar^2(k_0^2+a).$$

Therefore $\sigma_p = \hbar\sqrt{a}$ and $\sigma_x\sigma_p = \frac{1}{2\sqrt{a}}\cdot\hbar\sqrt{a} = \frac{\hbar}{2}$, the minimum allowed by Theorem 13.11. The factor e^{ik_0x} gives the packet its mean momentum $\hbar k_0$ without changing its shape in x.

13.4.1 MOMENTUM SPACE

By the derivative rule for Fourier transforms (Theorem 11.14), $\hat{p}$ acts on the transform as multiplication by $\hbar k$. This suggests that the Fourier transform is the wavefunction of the particle in momentum space.

Definition 13.9 (Momentum-space wavefunction).

$$\Phi(p,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} e^{-ipx/\hbar}\,\Psi(x,t)\,dx,$$
$$\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} e^{ipx/\hbar}\,\Phi(p,t)\,dp.$$

This is the Fourier transform of Chapter 11 with $k = p/\hbar$, rescaled so that, by Plancherel's theorem, $\int|\Phi|^2dp = \int|\Psi|^2dx = 1$. The momentum distribution is $|\Phi(p,t)|^2$, and $\langle p\rangle = \int p|\Phi|^2dp$, which agrees with the operator definition. In momentum space $\hat{p}$ is multiplication by p, and $\hat{x} = i\hbar\,\partial_p$. Neither representation

is more fundamental. $\Psi(x)$ and $\Phi(p)$ are the components of one state vector $|\Psi\rangle$ in two different bases (Theorem 11.16).

13.5 Ehrenfest's Theorem and the Classical Limit

Theorem 13.10 (Ehrenfest). *For a particle in a potential $V(x)$,*

$$\frac{d\langle x\rangle}{dt} = \frac{\langle p\rangle}{m}, \qquad \frac{d\langle p\rangle}{dt} = \left\langle -\frac{dV}{dx}\right\rangle.$$

Proof. The first equation is Theorem 13.7. For the second, differentiate under the integral and use (13.1) to replace $\partial_t\Psi$ and $\partial_t\Psi^*$:

$$\frac{d\langle p\rangle}{dt} = -i\hbar\int \left(\partial_t\Psi^*\,\Psi' + \Psi^*\,\partial_t\Psi'\right)dx.$$

The kinetic terms cancel after integration by parts. The potential terms give $-\int\left[V\Psi^*\Psi' - \Psi^*(V\Psi)'\right]dx = -\int \Psi^* V'\Psi\,dx$. A slicker proof uses the commutator identity of Theorem 15.5: $d/dt\,\langle\hat{p}\rangle = \frac{i}{\hbar}\langle[\hat{H},\hat{p}]\rangle$ and $[V,\hat{p}] = i\hbar V'$. ∎

Expectation values obey Newton's law in the form $m\;d^2\langle x\rangle/dt^2 = \langle F(x)\rangle$. This is not quite classical mechanics, because $\langle F(x)\rangle \neq F(\langle x\rangle)$ in general. They are equal when F is linear in x: for a free particle, a uniform field, or a harmonic oscillator, the center of a wave packet follows the classical trajectory *exactly*. For other potentials they are approximately equal when the packet is narrow compared with the scale on which F varies. That is the *classical limit*. A macroscopic object has a packet so narrow, relative to every relevant length scale, that quantum corrections are utterly negligible.

13.6 The Uncertainty Principle

Theorem 13.11 (Heisenberg uncertainty principle). *For every normalized wavefunction,*

$$\sigma_x\,\sigma_p \geq \frac{\hbar}{2},$$

with equality exactly for Gaussian wave packets.

Proof. At a fixed time, $\Phi(p)$ is the Fourier transform of $\Psi(x)$ with $k = p/\hbar$, so $\sigma_p = \hbar\sigma_k$. Apply Theorem 11.20. ∎

The principle is a statement about the state, not about the measuring apparatus. No state exists in which both position and momentum are sharp. A measurement of position on an ensemble prepared in state Ψ gives results spread by σ_x; a measurement of momentum on an identically prepared ensemble gives results spread by σ_p; and the product of the two spreads is at least $\hbar/2$. In Chapter 15 the principle is generalized to every pair of observables and traced to the commutator $[\hat{x}, \hat{p}] = i\hbar$.

Example 13.12 (Why atoms have the size they do). A hydrogen electron confined within radius r has $\sigma_p \gtrsim \hbar/r$, so its kinetic energy is at least about $\frac{\hbar^2}{2m_e r^2}$. Its potential energy is about $-\frac{e^2}{4\pi\varepsilon_0 r}$. Minimizing $E(r) = \frac{\hbar^2}{2m_e r^2} - \frac{e^2}{4\pi\varepsilon_0 r}$ gives $r = a_0$ and $E = -13.6$ eV. The estimate is exact, which is partly luck. The uncertainty principle stabilizes the atom: squeezing the electron closer to the nucleus would lower its potential energy, but would raise its kinetic energy more. This answers the classical collapse paradox of Example 12.10.

Example 13.13 (Zero-point energy). A harmonic oscillator has $E = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$. In any state, $\langle E\rangle \geq \frac{\sigma_p^2}{2m} + \frac{1}{2}m\omega^2\sigma_x^2 \geq \frac{\hbar^2}{8m\sigma_x^2} + \frac{1}{2}m\omega^2\sigma_x^2$. Minimizing over σ_x gives $\langle E\rangle \geq \frac{1}{2}\hbar\omega$. The oscillator cannot be at rest at the bottom of its well. Its lowest energy, the *zero-point energy* $\frac{1}{2}\hbar\omega$, is exactly what we find by solving the Schrödinger equation in Chapter 14. Zero-point motion is why helium remains liquid at absolute zero under atmospheric pressure.

13.7 The Free Gaussian Wave Packet

For a free particle ($V = 0$), the Schrödinger equation can be solved for any initial condition by Fourier transformation. Each plane wave e^{ikx} evolves as $e^{i(kx-\omega(k)t)}$ with $\hbar\omega = \frac{\hbar^2 k^2}{2m}$, so

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \phi(k)\, e^{i(kx - \hbar k^2 t/2m)}\, dk, \qquad \phi(k) = \frac{1}{\sqrt{2\pi}}\int \Psi(x,0) e^{-ikx} dx. \tag{13.3}$$

Theorem 13.14 (Spreading of a Gaussian). *The free particle with* $\Psi(x,0) = (2\pi\sigma_0^2)^{-1/4} e^{-x^2/(4\sigma_0^2) + ik_0 x}$ *has* $|\Psi(x,t)|^2$ *a normalized Gaussian centered at* $v_g t$,

with $v_g = \hbar k_0/m$, and with standard deviation

$$\sigma(t) = \sigma_0\sqrt{1 + \left(\frac{\hbar t}{2m\sigma_0^2}\right)^2}.$$

Proof. By Theorem 11.11 with the modulation rule, $\phi(k) = (2\sigma_0^2/\pi)^{1/4} e^{-\sigma_0^2(k-k_0)^2}$. Substituting into (13.3), the exponent is quadratic in k:

$$-\sigma_0^2(k-k_0)^2 + ikx - \frac{i\hbar t}{2m}k^2 = -\left(\sigma_0^2 + \frac{i\hbar t}{2m}\right)k^2 + (2\sigma_0^2 k_0 + ix)k - \sigma_0^2 k_0^2.$$

Let $\alpha = \sigma_0^2 + \frac{i\hbar t}{2m}$, which has positive real part. The Gaussian integral $\int e^{-\alpha k^2 + \beta k}dk = \sqrt{\pi/\alpha}\, e^{\beta^2/4\alpha}$ (Chapter 5, extended to complex α with $\operatorname{Re}\alpha > 0$ by analytic continuation) gives

$$|\Psi(x,t)|^2 \propto \left|\exp\left(\frac{(2\sigma_0^2 k_0 + ix)^2}{4\alpha}\right)\right|^2.$$

Write $\frac{1}{\alpha} = \frac{\alpha^*}{|\alpha|^2}$ and expand. After simplification, the exponent of $|\Psi|^2$ is $-\frac{(x-\hbar k_0 t/m)^2\,\sigma_0^2}{2|\alpha|^2}$. This is a Gaussian centered at $v_g t$ with variance $\frac{|\alpha|^2}{\sigma_0^2} = \sigma_0^2 + \frac{\hbar^2 t^2}{4m^2\sigma_0^2}$. Normalization is guaranteed by Theorem 13.5. ∎

The momentum distribution $|\phi(k)|^2$ does not change, since a free particle's momentum is conserved. But the different momentum components travel at different speeds, so the position distribution spreads. For $t \gg 2m\sigma_0^2/\hbar$, $\sigma(t) \approx \frac{\hbar t}{2m\sigma_0} = \sigma_v t$, where $\sigma_v = \sigma_p/m$ is the spread in velocity. The packet then spreads like a cloud of classical particles with that spread of velocities. The narrower the initial packet, the faster it spreads, because a narrow packet contains a wide range of momenta.

Example 13.15 (Electrons and dust). The time scale for doubling the width, roughly, is $\tau = 2m\sigma_0^2/\hbar$. (a) An electron localized to $\sigma_0 = 0.1$ nm has $\tau = \frac{2(9.11\times10^{-31})(10^{-10})^2}{1.055\times10^{-34}} \approx 1.7\times10^{-16}$ s. Atomic-scale electron packets spread essentially instantly. (b) A dust grain of mass 10^{-15} kg localized to 1 μm has $\tau \approx 1.9\times10^{7}$ s, about seven months. (c) For a 1 g ball localized to 1 nm, $\tau \approx 1.9\times10^{13}$ s, about six hundred thousand years.

13.8 Stationary States

When V does not depend on t, the Schrödinger equation can be separated (Section 10.6). Try $\Psi(x,t) = \psi(x)\,T(t)$. Substituting into (13.1) and dividing by ψT,

$$i\hbar\frac{T'}{T} = \frac{1}{\psi}\Big(-\frac{\hbar^2}{2m}\psi'' + V\psi\Big) = E,$$

a constant, because the left side depends only on t and the middle only on x. The time equation gives $T = e^{-iEt/\hbar}$. The space equation is the *time-independent Schrödinger equation*.

Definition 13.16 (Time-independent Schrödinger equation).

$$-\frac{\hbar^2}{2m}\psi''(x) + V(x)\psi(x) = E\psi(x), \qquad \text{that is,} \qquad \hat{H}\psi = E\psi. \tag{13.4}$$

A normalizable solution ψ is an *energy eigenfunction*, E is its *energy eigenvalue*, and $\Psi = \psi(x)e^{-iEt/\hbar}$ is a *stationary state*.

Theorem 13.17 (Properties of stationary states). *If $\Psi = \psi(x)e^{-iEt/\hbar}$ with $\hat{H}\psi = E\psi$, then:*

(a) *$|\Psi(x,t)|^2 = |\psi(x)|^2$ does not change in time, and neither does the expectation value of any observable that does not depend explicitly on t;*

(b) *the energy is sharp: $\langle H\rangle = E$ and $\sigma_H = 0$;*

(c) *E is real, and if ψ is normalizable then $E > V_{\min}$, the minimum of the potential.*

Proof. (a) The phase factor cancels in $\Psi^*\hat{Q}\Psi$. (b) $\langle H\rangle = \int \psi^*\hat{H}\psi = E$ and $\langle H^2\rangle = \int \psi^*\hat{H}^2\psi = E^2$, so $\sigma_H^2 = 0$. (c) E is real because $\hat{H}$ is Hermitian (Theorem 9.25). For the bound,

$$E = \langle H\rangle = \frac{\hbar^2}{2m}\int |\psi'|^2dx + \int V|\psi|^2dx > V_{\min}\int |\psi|^2dx = V_{\min},$$

after integrating the kinetic term by parts. The inequality is strict because $\psi' \equiv 0$ is incompatible with normalizability. ∎

Despite the name, nothing is stationary about the wavefunction itself; its phase rotates at angular frequency $E/\hbar$. But that rotation is a global phase and is un-

observable, so every measurable property is constant in time. A stationary state is a standing wave, the quantum analog of a normal mode (Example 10.11).

13.8.1 REQUIREMENTS ON ψ

For the solutions of (13.4) to be physically meaningful, ψ must be:

- *continuous*, since a jump in ψ would make ψ'' contain the derivative of a delta function, which nothing in the equation could balance;
- *differentiable*, with ψ' continuous wherever V is finite. Integrating (13.4) across a point shows that ψ' jumps only where V is infinite or contains a delta function (Chapter 14);
- *normalizable*, for a bound state. Unnormalizable solutions such as plane waves describe scattering, and are used only in superpositions.

The last condition is what quantizes the energy. For most values of E, every solution of (13.4) grows exponentially at large $|x|$, as in Theorem 10.12. Only for special values of E does a solution decay on both sides. This is the eigenvalue problem of Sturm–Liouville theory (Theorem 10.14), with the Hamiltonian as the operator.

13.8.2 GENERAL SOLUTIONS

Suppose $\hat{H}$ has normalized eigenfunctions $\psi_1, \psi_2, \dots$ with energies $E_1, E_2, \dots$. By the Sturm–Liouville theorem they are orthonormal and complete, so any initial wavefunction can be expanded, and each term then evolves with its own phase.

Theorem 13.18 (Expansion in stationary states). *If* $\Psi(x,0) = \sum_n c_n \psi_n(x)$, *then*

$$\Psi(x,t) = \sum_n c_n\, \psi_n(x)\, e^{-iE_n t/\hbar}, \qquad c_n = \int \psi_n^*(x)\, \Psi(x,0)\, dx.$$

Moreover $\sum_n |c_n|^2 = 1$, *and* $\langle H \rangle = \sum_n |c_n|^2 E_n$ *is constant in time.*

Proof. Each term solves (13.1) and the sum matches the initial condition, so the sum is the solution by uniqueness. The formula for c_n is Theorem 9.14. Normalization of Ψ and Parseval's identity give $\sum |c_n|^2 = 1$. Finally $\langle H \rangle = \sum_{m,n} c_m^* c_n E_n \langle \psi_m | \psi_n \rangle e^{i(E_m - E_n)t/\hbar} = \sum_n |c_n|^2 E_n$. ∎

This theorem is the recipe for solving any time-dependent problem with a time-

independent potential: find the stationary states, expand the initial state, and attach the phases. It also has a probabilistic meaning that anticipates the general measurement postulate of Chapter 15. A measurement of energy yields one of the eigenvalues E_n, with probability $|c_n|^2$, and leaves the system in the state ψ_n.

Example 13.19 (A sloshing superposition). Let $\Psi(x,0) = \frac{1}{\sqrt{2}}(\psi_1 + \psi_2)$, with ψ_1, ψ_2 real and orthonormal. Then

$$|\Psi(x,t)|^2 = \tfrac{1}{2}\left[\psi_1^2 + \psi_2^2 + 2\psi_1\psi_2\cos(\omega_{21}t)\right], \qquad \omega_{21} = \frac{E_2 - E_1}{\hbar}.$$

The probability density oscillates at the *Bohr frequency* ω_{21}. So does $\langle x\rangle$, unless $\int x\psi_1\psi_2\,dx = 0$. A charged particle in such a superposition has an oscillating dipole moment, and it radiates precisely at the frequency $\omega_{21} = (E_2 - E_1)/\hbar$. This is the origin of Bohr's second postulate (Chapter 12). Only superpositions of different energies show time dependence. A pure stationary state neither moves nor radiates, which is why atoms in their ground states are stable.

LOOKING AHEAD. Theorem 13.18 reduces the dynamics of quantum mechanics to an eigenvalue problem. Chapter 14 solves that problem for the most important one-dimensional potentials: the infinite well, the harmonic oscillator, the delta function, the finite well, and barriers. Chapter 15 then recasts everything in the abstract language of Chapter 9, and Chapter 16 solves the hydrogen atom.

13.9 Worked Problems

The problems below compute expectation values and uncertainties for a wavefunction with a kink, examine a state with infinite position uncertainty, trace the flow of probability in a spreading packet, and establish three structural facts: when Ehrenfest's theorem is exact, how wavefunctions transform under a change of reference frame, and how fast a state can change.

Problem 13.1 A triangular wavefunction

Let $\Psi(x,0) = A(a - |x|)$ for $|x| \le a$ and 0 otherwise. Normalize it, and compute σ_x, σ_p, and $\sigma_x\sigma_p$.

Solution. By symmetry, $1 = 2A^2\int_0^a (a-x)^2 dx = \frac{2A^2a^3}{3}$, so $A^2 = \frac{3}{2a^3}$. The

probability density is even, so $\langle x\rangle = 0$, and

$$\langle x^2\rangle = 2A^2\int_0^a x^2(a-x)^2\,dx = 2A^2\cdot\frac{a^5}{30} = \frac{a^2}{10}, \qquad \sigma_x = \frac{a}{\sqrt{10}}.$$

Ψ is real, so $\langle p\rangle = 0$ (the current vanishes). For $\langle p^2\rangle$, note that Ψ'' contains delta functions at $x = 0$ and $x = \pm a$, where Ψ' jumps. The safe route is $\langle p^2\rangle = \langle\hat{p}\Psi|\hat{p}\Psi\rangle = \hbar^2\int|\Psi'|^2dx$, which uses only the first derivative, since $\hat{p}$ is Hermitian. Here $|\Psi'| = A$ on $(-a, a)$, so

$$\langle p^2\rangle = \hbar^2A^2\cdot 2a = \frac{3\hbar^2}{a^2}, \qquad \sigma_p = \frac{\sqrt{3}\,\hbar}{a}, \qquad \sigma_x\sigma_p = \sqrt{\frac{3}{10}}\,\hbar \approx 0.548\,\hbar.$$

This is only 10% above the minimum $\hbar/2$. The triangle is close in shape to a Gaussian, and the uncertainty product is not sensitive to the kinks. Computing $\langle p^2\rangle$ as $-\hbar^2\int\Psi\Psi''\,dx$ while ignoring the delta functions in Ψ'' would give 0, an impossible result.

Problem 13.2 The momentum distribution of the triangle

For the state of the previous problem, find the momentum-space wavefunction $\phi(k) = \frac{1}{\sqrt{2\pi}}\int\Psi(x)e^{-ikx}dx$, where $p = \hbar k$, and check $\int|\phi|^2dk = 1$. How does $|\phi(k)|^2$ decay, and why?

Solution. Ψ is even, so $\phi(k) = \frac{2A}{\sqrt{2\pi}}\int_0^a(a-x)\cos kx\,dx$. Integrating by parts, $\int_0^a(a-x)\cos kx\,dx = \frac{1-\cos ka}{k^2}$, so

$$\phi(k) = \frac{2A}{\sqrt{2\pi}}\cdot\frac{2\sin^2(ka/2)}{k^2} = \sqrt{\frac{3}{\pi a^3}}\,\frac{2\sin^2(ka/2)}{k^2}.$$

This is the square of the transform of a box, as it must be: the triangle is the convolution of two boxes of half-width $a/2$ (Theorem 11.18). Normalization follows from Plancherel's theorem, and a direct check with $a = 1$ gives $\int|\phi|^2dk = 1$.

$|\phi(k)|^2 \sim k^{-4}$ for large k. That is just fast enough for $\langle k^2\rangle = \int k^2|\phi|^2dk$ to converge (the integrand decays like k^{-2}), consistent with the finite σ_p found above, but $\langle k^4\rangle$ diverges. The slow decay is caused by the kinks of Ψ: a function with a jump in its first derivative has a transform decaying like k^{-2}. Sharp features in position require high momenta. A wavefunction with a jump in Ψ itself would have $|\phi|^2 \sim k^{-2}$ and infinite kinetic energy, which is why physical wavefunctions must be continuous.

Problem 13.3 Normalizable, with infinite σ_x

Show that $\Psi(x) = \sqrt{\frac{a}{\pi}}\,\frac{1}{\sqrt{x^2+a^2}}$ is normalized, that $\sigma_x = \infty$, and that σ_p is finite.

Solution. $\int |\Psi|^2 dx = \frac{a}{\pi}\int \frac{dx}{x^2+a^2} = \frac{a}{\pi}\cdot\frac{\pi}{a} = 1$. The position distribution is the Lorentzian (Cauchy) distribution. But $\langle x^2\rangle = \frac{a}{\pi}\int \frac{x^2}{x^2+a^2}dx$ diverges, since the integrand tends to 1. So $\sigma_x = \infty$. Even $\langle x\rangle$ is defined only as a principal value, since $\int \frac{|x|}{x^2+a^2}dx$ diverges.

The momentum spread is finite: $\Psi' = -\sqrt{\frac{a}{\pi}}\,\frac{x}{(x^2+a^2)^{3/2}}$, so

$$\langle p^2\rangle = \hbar^2\int |\Psi'|^2 dx = \frac{\hbar^2 a}{\pi}\int_{-\infty}^{\infty}\frac{x^2\,dx}{(x^2+a^2)^3} = \frac{\hbar^2 a}{\pi}\cdot\frac{\pi}{8a^3} = \frac{\hbar^2}{8a^2}.$$

The uncertainty principle is satisfied trivially, as $\infty\cdot\sigma_p \geq \hbar/2$. The example warns that σ_x is a poor measure of spread for distributions with heavy tails. The Lorentzian is perfectly localized in the sense that half of its probability lies in $|x| < a$, yet its standard deviation is infinite. The *entropic* uncertainty relations, which use the entropies of $|\Psi|^2$ and $|\phi|^2$, handle such cases better.

Problem 13.4 How probability flows in a spreading packet

For the free Gaussian of Theorem 13.14 with $k_0 = 0$, show that the probability current is $J = \rho v$, with velocity field

$$v(x,t) = x\,\frac{\dot\sigma(t)}{\sigma(t)}, \qquad \sigma(t) = \sigma_0\sqrt{1+\left(\frac{\hbar t}{2m\sigma_0^2}\right)^2}.$$

Interpret the result.

Solution. Write $\Psi = |\Psi|e^{iS/\hbar}$ with real phase S. Then by Theorem 13.4, $J = \frac{\hbar}{m}\operatorname{Im}(\Psi^*\Psi') = \frac{|\Psi|^2}{m}S'$. So $v = J/\rho = S'/m$: the velocity field is the gradient of the phase, divided by the mass. For the Gaussian,

$$\Psi \propto \exp\Big(-\frac{x^2}{4\sigma_0^2(1+i\tau)}\Big), \qquad \tau = \frac{\hbar t}{2m\sigma_0^2}.$$

Since $\frac{1}{1+i\tau} = \frac{1-i\tau}{1+\tau^2}$, the phase is $\frac{S}{\hbar} = \frac{x^2\tau}{4\sigma_0^2(1+\tau^2)}$, plus a function of t only. Hence

$$v = \frac{S'}{m} = \frac{\hbar}{m}\cdot\frac{x\tau}{2\sigma_0^2(1+\tau^2)}.$$

On the other hand, $\sigma = \sigma_0\sqrt{1+\tau^2}$ gives $\frac{\dot\sigma}{\sigma} = \frac{\tau\dot\tau}{1+\tau^2}$ with $\dot\tau = \frac{\hbar}{2m\sigma_0^2}$, so $x\frac{\dot\sigma}{\sigma} = \frac{\hbar x\tau}{2m\sigma_0^2(1+\tau^2)}$, the same expression. A numerical check at $\hbar = m = \sigma_0 = 1$, $t = 1.7$, $x = 0.9$ gives 0.222061 for both.

The flow is a uniform *dilation*: each part of the distribution moves outward with speed proportional to its distance from the center, like the galaxies in an expanding universe, and the shape stays Gaussian. At $t = 0$, $v = 0$ everywhere, though the momentum spread is already $\hbar/(2\sigma_0)$. The phase gradient that drives the expansion builds up as faster components outrun slower ones. The quadratic phase $\propto x^2$ is the chirp of the chirped-Gaussian problem in Chapter 11.

Problem 13.5 When Ehrenfest's theorem is exact

(a) Show that if $V(x)$ is a polynomial of degree at most 2, then $\langle x\rangle(t)$ obeys Newton's law $m\ddot{\langle x\rangle} = -V'(\langle x\rangle)$ exactly.

(b) For $V = \lambda x^4$ and a state with $\langle x\rangle = x_0$ and $\langle (x-x_0)^2\rangle = \sigma^2$ whose distribution is symmetric about x_0, find the difference between $-\langle V'(x)\rangle$ and $-V'(\langle x\rangle)$.

Solution. (a) By Theorem 13.10, $m\ddot{\langle x\rangle} = -\langle V'(x)\rangle$. If $V = \alpha + \beta x + \gamma x^2$, then $V' = \beta + 2\gamma x$ is linear, and the expectation of a linear function is that function of the expectation: $\langle V'(x)\rangle = \beta + 2\gamma\langle x\rangle = V'(\langle x\rangle)$. So the center of the packet moves exactly like a classical particle, for any state. This covers the free particle, a uniform field, and the harmonic oscillator.

(b) $V' = 4\lambda x^3$. Write $x = x_0 + u$ with $\langle u\rangle = 0$ and $\langle u^2\rangle = \sigma^2$, and $\langle u^3\rangle = 0$ by symmetry:

$$\langle x^3\rangle = x_0^3 + 3x_0^2\langle u\rangle + 3x_0\langle u^2\rangle + \langle u^3\rangle = x_0^3 + 3x_0\sigma^2.$$

So $-\langle V'\rangle = -V'(x_0) - 12\lambda x_0\sigma^2$. The packet feels a stronger restoring force than a classical particle at its center, because the force grows faster than linearly and the packet samples regions of larger $|x|$. The correction is negligible when $\sigma^2 \ll x_0^2$, that is, when the packet is narrow compared with the scale on which the force varies. That is the precise content of the classical limit. Since σ itself evolves, the equations for $\langle x\rangle$ and σ are coupled, and quantum dynamics in an anharmonic potential cannot be reduced to classical dynamics of the center alone.

Problem 13.6 Galilean boosts

Let $\Psi(x,t)$ solve the free Schrödinger equation. Show that

$$\Psi_v(x,t) = \Psi(x - vt,\, t)\, e^{i(mvx - \frac{1}{2}mv^2t)/\hbar}$$

is also a solution, and interpret it. How do $\langle x\rangle$, $\langle p\rangle$, and $\langle E\rangle$ change?

Solution. Let $\xi = x - vt$ and $\theta = (mvx - \frac{1}{2}mv^2t)/\hbar$, so $\Psi_v = \Psi(\xi, t)e^{i\theta}$, with $\partial_x\theta = \frac{mv}{\hbar}$ and $\partial_t\theta = -\frac{mv^2}{2\hbar}$. Then

$$i\hbar\partial_t\Psi_v = e^{i\theta}\Big[i\hbar\big(\Psi_t - v\Psi_\xi\big) + \tfrac{1}{2}mv^2\Psi\Big],$$
$$-\frac{\hbar^2}{2m}\partial_x^2\Psi_v = e^{i\theta}\Big[-\frac{\hbar^2}{2m}\Psi_{\xi\xi} - i\hbar v\Psi_\xi + \tfrac{1}{2}mv^2\Psi\Big],$$

using $\partial_x^2(\Psi e^{i\theta}) = e^{i\theta}\big(\Psi_{\xi\xi} + \frac{2imv}{\hbar}\Psi_\xi - \frac{m^2v^2}{\hbar^2}\Psi\big)$. The difference of the two lines is $e^{i\theta}\big(i\hbar\Psi_t + \frac{\hbar^2}{2m}\Psi_{\xi\xi}\big) = 0$, since Ψ solves the free equation. So Ψ_v is a solution.

Ψ_v describes the same physical state seen from a frame moving at velocity $-v$. Its density is $|\Psi(x - vt, t)|^2$, the original density translated at velocity v, so $\langle x\rangle_v = \langle x\rangle + vt$. The factor $e^{imvx/\hbar}$ shifts every momentum by mv (Theorem 11.14), so $\langle p\rangle_v = \langle p\rangle + mv$, and $\langle E\rangle_v = \frac{\langle (p+mv)^2\rangle}{2m} = \langle E\rangle + v\langle p\rangle + \frac{1}{2}mv^2$, exactly the classical transformation. The phase $e^{-imv^2t/2\hbar}$ is not optional: without it, Ψ_v would fail the equation. The wavefunction is not a scalar under a change of frame. It acquires a position- and time-dependent phase, which makes sense, since the de Broglie wavelength h/p depends on the frame.

Problem 13.7 How fast can a state change?

Let $\Psi(0) = \sum c_n\psi_n$ be a superposition of stationary states with energies E_n, and define the *survival probability* $P(t) = |\langle\Psi(0)|\Psi(t)\rangle|^2$.

(a) Show that $P(t) = \big|\sum_n |c_n|^2 e^{-iE_nt/\hbar}\big|^2$, and that $P(t) = 1 - \frac{\sigma_E^2t^2}{\hbar^2} + O(t^4)$, where σ_E is the energy uncertainty.

(b) For $\Psi(0) = \frac{1}{\sqrt{2}}(\psi_1 + \psi_2)$, find $P(t)$ exactly, and the time at which $\Psi(t)$ first becomes orthogonal to $\Psi(0)$.

Solution. (a) By Theorem 13.18, $\Psi(t) = \sum c_n\psi_n e^{-iE_nt/\hbar}$, and orthonormality gives $\langle\Psi(0)|\Psi(t)\rangle = \sum |c_n|^2 e^{-iE_nt/\hbar}$. Write $p_n = |c_n|^2$, a probability distribution over energies, and expand the exponential:

$$\langle\Psi(0)|\Psi(t)\rangle = 1 - \frac{it}{\hbar}\langle E\rangle - \frac{t^2}{2\hbar^2}\langle E^2\rangle + O(t^3).$$

Its squared modulus is $1 - \frac{t^2}{\hbar^2}\big(\langle E^2\rangle - \langle E\rangle^2\big) + O(t^4) = 1 - \frac{\sigma_E^2t^2}{\hbar^2} + O(t^4)$; the cross terms of order t^3 cancel because P is even in t. So a state begins to change on the time scale $\hbar/\sigma_E$. A stationary state, with $\sigma_E = 0$, does not change at all.

(b) With $p_1 = p_2 = \tfrac{1}{2}$,

$$P(t) = \left|\tfrac{1}{2}e^{-iE_1 t/\hbar} + \tfrac{1}{2}e^{-iE_2 t/\hbar}\right|^2 = \frac{1}{2}\Big(1 + \cos\frac{(E_2 - E_1)t}{\hbar}\Big) = \cos^2\frac{(E_2 - E_1)t}{2\hbar}.$$

Here $\sigma_E = \frac{E_2 - E_1}{2}$, so $P(t) = \cos^2(\sigma_E t/\hbar)$. The state first becomes orthogonal to its initial form at $t_\perp = \frac{\pi\hbar}{2\sigma_E} = \frac{\pi\hbar}{E_2 - E_1}$.

The *Mandelstam–Tamm inequality* states that $P(t) \geq \cos^2(\sigma_E t/\hbar)$ for every state, for $0 \leq t \leq \pi\hbar/2\sigma_E$. So no state can become orthogonal to itself sooner than $\frac{\pi\hbar}{2\sigma_E}$, and the equal two-state superposition is the fastest possible. This is the precise form of the energy–time uncertainty relation $\Delta E\,\Delta t \gtrsim \hbar$: here Δt is not the uncertainty of a time measurement, since time is a parameter and not an observable, but the time a state needs to change appreciably. It sets a fundamental limit on the speed of quantum computation.

EXERCISES

A. COMPUTATION

13.1 Normalize $\Psi(x,0) = Ae^{-\lambda|x|}$, with $\lambda > 0$, and compute $\langle x\rangle$, $\langle x^2\rangle$, and σ_x.

13.2 Normalize $\Psi(x,0) = A/(x^2 + a^2)$ and find the probability that $|x| < a$.

13.3 For $\Psi(x,0) = \sqrt{30/a^5}\,x(a-x)$ on $[0,a]$ (Example 13.3), compute $\langle x\rangle$, $\langle x^2\rangle$, σ_x, $\langle p\rangle$, $\langle p^2\rangle$, and $\sigma_x\sigma_p$. Verify the uncertainty principle.

13.4 For the state of Exercise 13.1, compute $\langle p^2\rangle$ and $\sigma_x\sigma_p$. (Careful: Ψ' has a jump at 0.)

13.5 Compute the probability current for $\Psi = Ae^{ikx} + Be^{-ikx}$, and interpret.

13.6 Show that $J = 0$ for $\Psi = f(x)e^{i\alpha}$ with f real. Compute J for $\Psi = (1 + ix)e^{-x^2}$.

13.7 Compute the momentum-space wavefunction $\Phi(p)$ for $\Psi(x) = \sqrt{\lambda}\,e^{-\lambda|x|}$, and verify $\int |\Phi|^2 dp = 1$.

13.8 For the Gaussian of Example 13.8, compute $\Phi(p)$ directly and confirm $\langle p\rangle = \hbar k_0$ and $\sigma_p = \hbar\sqrt{a}$.

13.9 A proton is confined to a nucleus of radius 5×10^{-15} m. Estimate its minimum kinetic energy in MeV.

13.10 An electron is confined to a box 0.10 nm wide. Estimate its minimum kinetic energy in eV.

13.11 A free electron is described at $t=0$ by a Gaussian with $\sigma_0 = 1.0$ nm. Find $\sigma(t)$ after 1 ps and after 1 ns.

13.12 Find the time for a Gaussian packet of a 1 g ball with $\sigma_0 = 10^{-10}$ m to double its width.

13.13 For the free Gaussian, verify that $\sigma(t)\sigma_p \geq \hbar/2$ for all t, with equality only at $t=0$.

13.14 A particle is in the state $\Psi(x,0) = \frac{1}{\sqrt{3}}\psi_1 + \sqrt{\frac{2}{3}}\,\psi_2$, where ψ_n are orthonormal stationary states. Find $\Psi(x,t)$, the probabilities of measuring E_1 and E_2, and $\langle H\rangle$.

13.15 In the state of the previous exercise, with ψ_1, ψ_2 real, show that $\langle x\rangle$ oscillates and find its angular frequency.

13.16 Show that $\Psi(x,t) = Ae^{-a[(mx^2/\hbar)+it]}$ satisfies the Schrödinger equation for a suitable potential $V(x)$. Find V and the energy.

13.17 Verify that $\psi(x) = \sqrt{2/L}\sin(n\pi x/L)$ satisfies (13.4) with $V=0$ on $(0,L)$, and find E.

B. THEORY AND PROOF

13.18 Prove that if $\Psi(x,t)$ solves (13.1) with real V, then $\Psi^*(x,-t)$ also does (time-reversal symmetry).

13.19 Prove that if $V(x)$ is even, then the energy eigenfunctions can be chosen to be even or odd.

13.20 Show that adding a constant V_0 to the potential multiplies Ψ by $e^{-iV_0 t/\hbar}$ and has no effect on any expectation value. Why, then, are potential differences observable?

13.21 Prove that $\langle p\rangle$ is real by showing $\hat{p}$ is Hermitian (Theorem 9.35).

13.22 Show that for a stationary state ψ that is real, $\langle p\rangle = 0$.

13.23 Derive Ehrenfest's theorem in three dimensions: $m\,\mathrm{d}\langle \mathbf{r}\rangle/\mathrm{d}t = \langle \mathbf{p}\rangle$ and $\mathrm{d}\langle \mathbf{p}\rangle/\mathrm{d}t = -\langle \nabla V\rangle$.

13.24 Prove the *virial theorem* for stationary states: $2\langle T\rangle = \langle x\,V'(x)\rangle$. (Compute $\mathrm{d}/\mathrm{d}t\,\langle xp\rangle$ and use stationarity.) Deduce $\langle T\rangle = \langle V\rangle$ for the harmonic oscillator.

13.25 Show that the time-independent Schrödinger equation has no normalizable solution with $E < V_{\min}$, using Theorem 13.17.

13.26 Show that a complex potential $V = V_r - i\Gamma$, with constant $\Gamma > 0$, makes $\int |\Psi|^2$ decay as $e^{-2\Gamma t/\hbar}$. What could this describe?

C. CHALLENGE

13.27 (Spreading in momentum space.) Show that for a free particle, $\Phi(p,t) = \Phi(p,0)e^{-ip^2t/2m\hbar}$, so $|\Phi|^2$ is constant. Then show that $\mathrm{d}^2\langle x^2\rangle/\mathrm{d}t^2 = 2\langle p^2\rangle/m^2$, and deduce that $\sigma_x(t)^2$ is a quadratic function of t for *every* free wave packet, not only Gaussians.

13.28 Show that the most general state with $\sigma_x\sigma_p = \hbar/2$ is a Gaussian with real width parameter, and that a Gaussian with complex width parameter $a = a_r + ia_i$ has $\sigma_x\sigma_p = \frac{\hbar}{2}\sqrt{1 + a_i^2/a_r^2}$. Interpret this using Theorem 13.14.

13.29 The *Wigner function* $W(x,p) = \frac{1}{\pi\hbar}\int \Psi^*(x+y)\Psi(x-y)e^{2ipy/\hbar}dy$ is a quasi-probability distribution in phase space. Compute it for the Gaussian of Example 13.8 and show that it is a positive Gaussian in (x,p). Show that its marginals $\int W\,dp$ and $\int W\,dx$ are $|\Psi|^2$ and $|\Phi|^2$.

13.30 Estimate the ground-state energy of a particle of mass m in the potential $V = \lambda x^4$ using the uncertainty principle. Compare with the variational estimate obtained by minimizing $\langle H\rangle$ over Gaussians e^{-ax^2}.

CHAPTER 14

Stationary States in One Dimension

LEARNING OBJECTIVES

- Solve the infinite square well completely, and expand arbitrary states in its eigenfunctions.
- Solve the harmonic oscillator algebraically with ladder operators, and analytically with Hermite polynomials.
- Compute matrix elements and expectation values of powers of x and p using ladder operators.
- Distinguish bound states from scattering states, and normalize scattering states.
- Solve the delta-function potential for its bound state and its reflection and transmission coefficients.
- Find the bound states of the finite square well graphically, and analyze its limits.
- Compute reflection and transmission at steps and barriers, explain tunneling, and apply the WKB estimate.

Theorem 13.18 reduced quantum dynamics to the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\psi'' + V(x)\psi = E\psi. \tag{14.1}$$

This chapter solves it for the potentials that every physicist must know. They are chosen because they can be solved exactly, and because each one models something real. The square well models electrons in a quantum dot or a long molecule. The harmonic oscillator models every system near equilibrium. The delta function models a short-range impurity. Barriers model tunneling through

junctions and the radioactive decay of nuclei. Along the way the general features of quantum bound states emerge: discrete energies, a ground state with no nodes, alternating parity, and penetration into classically forbidden regions.

14.1 The Infinite Square Well

Let $V = 0$ for $0 < x < a$ and $V = \infty$ elsewhere. The particle cannot be outside the well, so $\psi = 0$ there, and continuity requires $\psi(0) = \psi(a) = 0$. Inside, (14.1) is $\psi'' = -k^2\psi$ with $k = \sqrt{2mE}/\hbar$. This is precisely the Sturm–Liouville problem of Example 10.15.

Theorem 14.1 (Infinite square well). *The stationary states are*

$$\psi_n(x) = \sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a}, \qquad E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}, \qquad n = 1, 2, 3, \dots$$

They are orthonormal, $\int_0^a \psi_m\psi_n\,dx = \delta_{mn}$, and complete in $L^2(0,a)$.

Proof. The general solution inside is $A\sin kx + B\cos kx$. The condition $\psi(0) = 0$ forces $B = 0$, and $\psi(a) = 0$ forces $ka = n\pi$. Here $n = 0$ gives $\psi \equiv 0$, and negative n gives nothing new. Normalization gives $|A|^2\int_0^a \sin^2(n\pi x/a)\,dx = |A|^2 a/2 = 1$. Orthogonality and completeness are Theorem 10.14 and Theorem 11.8 for the Fourier sine series. ∎

The energies grow as n^2, so the levels spread apart. The ground state ψ_1 has no nodes, ψ_2 has one, and in general ψ_n has $n-1$. The states alternate between even and odd about the center of the well. The ground-state energy E_1 is not zero, which is the uncertainty principle again: confinement to width a costs kinetic energy of order $\hbar^2/ma^2$.

Example 14.2 (Expanding a state). At $t = 0$ a particle in the well is in the state $\Psi(x,0) = \sqrt{30/a^5}\,x(a-x)$ of Example 13.3. Find $\Psi(x,t)$, the probability of each energy, and $\langle H\rangle$.

Solution. By Theorem 13.18,

$$c_n = \sqrt{\frac{2}{a}}\sqrt{\frac{30}{a^5}}\int_0^a x(a-x)\sin\frac{n\pi x}{a}\,dx = \frac{4\sqrt{15}}{(n\pi)^3}\bigl(1-(-1)^n\bigr),$$

which is $\frac{8\sqrt{15}}{(n\pi)^3}$ for odd n and 0 for even n. The even states are absent because

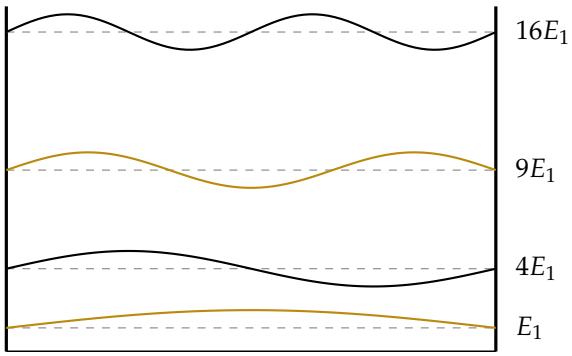


Figure 14.1: The first four stationary states of the infinite square well, each drawn at the height of its energy.

Ψ is symmetric about the center. So $\Psi(x,t) = \sum_{n \text{ odd}} \frac{8\sqrt{15}}{(n\pi)^3}\psi_n(x)e^{-iE_nt/\hbar}$. The probability of measuring E_1 is $|c_1|^2 = \frac{960}{\pi^6} \approx 0.9986$. The initial state is almost pure ground state, which is no surprise, since a parabola looks very like a sine arch. As a check on completeness, $\sum |c_n|^2 = \frac{960}{\pi^6}\sum_{n \text{ odd}} \frac{1}{n^6} = \frac{960}{\pi^6} \cdot \frac{\pi^6}{960} = 1$. Finally

$$\langle H \rangle = \sum |c_n|^2 E_n = \frac{960}{\pi^6} \cdot \frac{\pi^2\hbar^2}{2ma^2} \sum_{n \text{ odd}} \frac{1}{n^4} = \frac{480\hbar^2}{\pi^4 ma^2} \cdot \frac{\pi^4}{96} = \frac{5\hbar^2}{ma^2},$$

which agrees with the direct computation $\langle H \rangle = \frac{\langle p^2 \rangle}{2m} = \frac{10\hbar^2}{2ma^2}$ from Exercise 13.3.

Example 14.3 (Quantum revivals). Every E_n is an integer multiple of E_1, so after the time $T = 2\pi\hbar/E_1 = \frac{4ma^2}{\pi\hbar}$ every phase factor $e^{-iE_nT/\hbar} = e^{-2\pi i n^2}$ equals 1. Any initial state, however complicated, returns exactly to itself. Such *revivals* have been observed in atomic wave packets and in light in waveguides.

14.2 THE HARMONIC OSCILLATOR: ALGEBRAIC METHOD

The potential $V = \frac{1}{2}m\omega^2x^2$ governs every system near a stable equilibrium (Chapter 10). The Hamiltonian is

$$\hat{H} = \frac{1}{2m}(\hat{p}^2 + (m\omega\hat{x})^2).$$

If $\hat{p}$ and $\hat{x}$ were numbers, we would factor $u^2+v^2=(iu+v)(-iu+v)$. They are operators that do not commute, and the attempt to factor anyway is one of the most fruitful ideas in quantum mechanics.

Definition 14.4 (Ladder operators).

$$\hat{a}=\frac{1}{\sqrt{2\hbar m\omega}}\big(m\omega\hat{x}+i\hat{p}\big),\qquad \hat{a}^\dagger=\frac{1}{\sqrt{2\hbar m\omega}}\big(m\omega\hat{x}-i\hat{p}\big).$$

Since $\hat{x}$ and $\hat{p}$ are Hermitian, $\hat{a}^\dagger$ is indeed the adjoint of $\hat{a}$. Inverting,

$$\hat{x}=\sqrt{\frac{\hbar}{2m\omega}}\,(\hat{a}+\hat{a}^\dagger),\qquad \hat{p}=i\sqrt{\frac{\hbar m\omega}{2}}\,(\hat{a}^\dagger-\hat{a}). \tag{14.2}$$

Theorem 14.5 (Ladder algebra).

$$[\hat{x},\hat{p}]=i\hbar,\qquad [\hat{a},\hat{a}^\dagger]=1,\qquad \hat{H}=\hbar\omega\Big(\hat{a}^\dagger\hat{a}+\frac{1}{2}\Big),$$
$$[\hat{H},\hat{a}^\dagger]=\hbar\omega\,\hat{a}^\dagger,\qquad [\hat{H},\hat{a}]=-\hbar\omega\,\hat{a}.$$

Proof. For any function f, $[\hat{x},\hat{p}]f=x(-i\hbar f')-(-i\hbar)(xf)'=i\hbar f$. Then

$$\hat{a}\hat{a}^\dagger=\frac{1}{2\hbar m\omega}\big(m^2\omega^2\hat{x}^2+\hat{p}^2-im\omega[\hat{x},\hat{p}]\big)=\frac{\hat{H}}{\hbar\omega}+\frac{1}{2},$$

and similarly $\hat{a}^\dagger\hat{a}=\frac{\hat{H}}{\hbar\omega}-\frac{1}{2}$. Subtracting gives $[\hat{a},\hat{a}^\dagger]=1$, and the second expression gives $\hat{H}$. Finally, $[\hat{H},\hat{a}^\dagger]=\hbar\omega[\hat{a}^\dagger\hat{a},\hat{a}^\dagger]=\hbar\omega\,\hat{a}^\dagger[\hat{a},\hat{a}^\dagger]=\hbar\omega\hat{a}^\dagger$. ∎

Theorem 14.6 (Spectrum of the oscillator). *The energies of the harmonic oscillator are $E_n=\big(n+\frac{1}{2}\big)\hbar\omega$, $n=0,1,2,\dots$, with normalized eigenstates*

$$\psi_0(x)=\Big(\frac{m\omega}{\pi\hbar}\Big)^{1/4}e^{-m\omega x^2/2\hbar},\qquad \psi_n=\frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n\psi_0.$$

Moreover $\hat{a}^\dagger\psi_n=\sqrt{n+1}\,\psi_{n+1}$ and $\hat{a}\psi_n=\sqrt{n}\,\psi_{n-1}$.

Proof. Raising and lowering. If $\hat{H}\psi=E\psi$, then by Theorem 14.5 $\hat{H}(\hat{a}^\dagger\psi)=\hat{a}^\dagger\hat{H}\psi+\hbar\omega\hat{a}^\dagger\psi=(E+\hbar\omega)\hat{a}^\dagger\psi$. So $\hat{a}^\dagger$ raises the energy by one quantum $\hbar\omega$, and likewise $\hat{a}$ lowers it by $\hbar\omega$, unless the result is zero.

The bottom rung. Since $\langle\hat{H}\rangle=\hbar\omega(\|\hat{a}\psi\|^2+\frac{1}{2})\geq\frac{1}{2}\hbar\omega$, energies are bounded below. Repeated lowering must therefore end at a state with $\hat{a}\psi_0=0$, which is

the first-order equation

$$\left(m\omega x + \hbar\frac{\mathrm{d}}{\mathrm{d}x}\right)\psi_0 = 0 \quad\Longrightarrow\quad \psi_0 = Ae^{-m\omega x^2/2\hbar}.$$

Its energy is $\hat{H}\psi_0 = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})\psi_0 = \frac{1}{2}\hbar\omega\,\psi_0$. Normalize with Theorem 5.12.

The ladder. $\hat{a}^\dagger$ applied repeatedly to ψ_0 gives states of energy $(n+\frac{1}{2})\hbar\omega$. Every eigenstate lies on this ladder: lowering any eigenstate must eventually reach ψ_0, the unique solution of $\hat{a}\psi = 0$, and so the energy of the original state is $(n+\frac{1}{2})\hbar\omega$ for some n.

Normalization. If ψ_n is normalized, then $\|\hat{a}^\dagger\psi_n\|^2 = \langle\psi_n|\hat{a}\hat{a}^\dagger|\psi_n\rangle = \langle\psi_n|(\hat{a}^\dagger\hat{a} + 1)|\psi_n\rangle = n+1$, using $\hat{a}^\dagger\hat{a}\psi_n = n\psi_n$. So $\hat{a}^\dagger\psi_n = \sqrt{n+1}\,\psi_{n+1}$, and induction gives the formula for ψ_n. The lowering relation follows similarly. ■

The operator $\hat{N} = \hat{a}^\dagger\hat{a}$ is the *number operator*, with $\hat{N}\psi_n = n\psi_n$. In the quantum theory of light, $\hat{a}^\dagger$ *creates* a photon in a mode and $\hat{a}$ *annihilates* one, and n counts photons. The energy levels are equally spaced, so all transitions between adjacent levels emit or absorb the same quantum $\hbar\omega$. Planck's hypothesis of Theorem 12.4 is recovered, apart from the zero-point energy $\frac{1}{2}\hbar\omega$.

14.2.1 MATRIX ELEMENTS BY ALGEBRA

Ladder operators turn integrals into counting. Write $\langle m|\hat{Q}|n\rangle = \int \psi_m^*\hat{Q}\psi_n\,dx$ and $x_0 = \sqrt{\hbar/m\omega}$. From (14.2),

$$\langle m|\hat{x}|n\rangle = \frac{x_0}{\sqrt{2}}\left(\sqrt{n}\,\delta_{m,n-1} + \sqrt{n+1}\,\delta_{m,n+1}\right). \tag{14.3}$$

So $\hat{x}$ connects only neighboring levels. This is the *selection rule* for dipole radiation from an oscillator: only transitions with $\Delta n = \pm 1$ occur.

Example 14.7 (Expectation values in ψ_n). $\langle x\rangle = \langle p\rangle = 0$ in every stationary state. For $\langle x^2\rangle$, expand $(\hat{a}+\hat{a}^\dagger)^2 = \hat{a}^2 + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}^{\dagger 2}$. Only the middle terms have diagonal matrix elements, and they give $(n+1)+n$. So

$$\langle x^2\rangle = \frac{\hbar}{2m\omega}(2n+1), \qquad \langle p^2\rangle = \frac{\hbar m\omega}{2}(2n+1), \qquad \sigma_x\sigma_p = \left(n+\frac{1}{2}\right)\hbar.$$

Then $\langle V\rangle = \frac{1}{2}m\omega^2\langle x^2\rangle = \frac{1}{2}E_n = \langle T\rangle$, as the virial theorem demands (Exercise 13.24). The ground state saturates the uncertainty bound.

Example 14.8 (Fourth moment). Compute $\langle n|\hat{x}^4|n\rangle$.

Solution. $\hat{x}^4 = \left(\frac{x_0^2}{2}\right)^2(\hat{a}+\hat{a}^\dagger)^4$. Of the sixteen orderings of four factors, the ones with nonzero diagonal elements have two $\hat{a}$'s and two $\hat{a}^\dagger$'s. There are six of them, and their values in ψ_n are

$$\hat{a}\hat{a}\hat{a}^\dagger\hat{a}^\dagger : (n+1)(n+2), \qquad \hat{a}\hat{a}^\dagger\hat{a}\hat{a}^\dagger : (n+1)^2, \qquad \hat{a}\hat{a}^\dagger\hat{a}^\dagger\hat{a} : n(n+1),$$
$$\hat{a}^\dagger\hat{a}\hat{a}\hat{a}^\dagger : n(n+1), \qquad \hat{a}^\dagger\hat{a}\hat{a}^\dagger\hat{a} : n^2, \qquad \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} : n(n-1).$$

The sum is $6n^2+6n+3$, so $\langle n|\hat{x}^4|n\rangle = \left(\frac{\hbar}{2m\omega}\right)^2(6n^2+6n+3)$. For $n=0$ this is $3\left(\frac{\hbar}{2m\omega}\right)^2$, which agrees with the Gaussian moment $\langle x^4\rangle = 3\sigma^4$ of Theorem 5.31.

14.3 The Harmonic Oscillator: Analytic Method

The algebraic method is elegant, but it relies on a trick that works for few potentials. The analytic method, which solves (14.1) by power series, is the one that generalizes, and it is the method we use for hydrogen.

Introduce the dimensionless variable $\xi = x/x_0$, with $x_0 = \sqrt{\hbar/m\omega}$, and the dimensionless energy $K = 2E/\hbar\omega$. Then (14.1) becomes

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}\xi^2} = (\xi^2 - K)\psi. \tag{14.4}$$

For large $|\xi|$, K is negligible, and $\psi'' \approx \xi^2\psi$ has approximate solutions $e^{\pm\xi^2/2}$. Only the decaying one is normalizable. So write

$$\psi(\xi) = h(\xi)e^{-\xi^2/2}.$$

Substituting gives $h'' - 2\xi h' + (K-1)h = 0$, which is Hermite's equation (10.1) with $2\nu = K-1$. By Theorem 10.12, a non-terminating series makes $h \sim e^{\xi^2}$ and $\psi \sim e^{\xi^2/2}$, which is not normalizable. So the series must terminate, which requires $\nu = n$, a nonnegative integer. Then $K = 2n+1$, that is, $E_n = (n+\frac{1}{2})\hbar\omega$, and $h = H_n$.

Theorem 14.9 (Oscillator eigenfunctions).

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\frac{1}{\sqrt{2^n n!}}\,H_n(\xi)\,e^{-\xi^2/2}, \qquad \xi = \sqrt{\frac{m\omega}{\hbar}}\,x.$$

The normalization follows from the orthogonality relation for Hermite polynomials (Chapter 10). The two methods agree, as they must. Comparing Theorem 14.6 with this formula shows that $\hat{a}^\dagger$ acting on $H_n e^{-\xi^2/2}$ produces $H_{n+1} e^{-\xi^2/2}$, which is the Hermite recurrence relation.

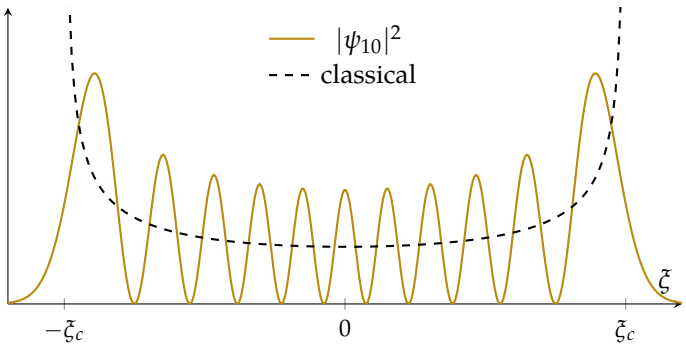


Figure 14.2: The probability density of the $n = 10$ oscillator state compared with the classical density $1/(\pi\sqrt{\xi_c^2 - \xi^2})$ for the same energy, where $\xi_c = \sqrt{21}$ is the classical turning point. The quantum density oscillates about the classical one and leaks slightly beyond the turning points.

Figure 14.2 illustrates the correspondence principle. For large n, the quantum probability density, averaged over its rapid oscillations, approaches the classical probability of finding the oscillating particle at each point. That classical density is largest near the turning points, where the particle moves slowly. The quantum density also extends slightly into the classically forbidden region $|\xi| > \xi_c$, where $E < V$.

14.4 The Free Particle and Scattering States

For $V = 0$ everywhere, (14.1) has solutions $e^{\pm ikx}$ with $E = \hbar^2 k^2/2m$ for every $k \in \mathbb{R}$. The spectrum is *continuous*, and no eigenfunction is normalizable. A free particle has no stationary states in the strict sense. Physical states are wave packets, superpositions $\int \phi(k) e^{ikx} dk$, which we analyzed in Chapter 13.

The unnormalizable solutions are still indispensable. They are *scattering states*, normalized to a delta function, $\int \psi_k^* \psi_{k'} dx = \delta(k - k')$ for $\psi_k = e^{ikx}/\sqrt{2\pi}$. A plane wave describes a steady beam of particles, with probability current $J = \hbar k/m$ per unit density.

Definition 14.10 (Bound and scattering states). Let $V(x) \to V_\pm$ as $x \to \pm\infty$. A solution of (14.1) with $E < \min(V_+, V_-)$ is a *bound state*: it decays exponentially at both ends, can be normalized, and occurs only at discrete energies. A solution with $E > \min(V_+, V_-)$ is a *scattering state*: it oscillates at one or both ends, and every such E is allowed.

In classical mechanics the same distinction holds: a particle with $E < V$ at both ends is trapped between turning points, while one with more energy escapes. The quantum difference is that energies of trapped particles are quantized, and that a particle can penetrate regions where $E < V$, and even pass through them.

14.5 The Delta-Function Potential

Let $V(x) = -\alpha\,\delta(x)$ with $\alpha > 0$, an idealized narrow attractive well.

Lemma 14.11 (Jump condition). *For $V = -\alpha\delta(x)$, ψ is continuous at 0 and ψ' jumps by $\Delta\psi' = \psi'(0^+) - \psi'(0^-) = -\frac{2m\alpha}{\hbar^2}\psi(0)$.*

Proof. Integrate (14.1) from $-\epsilon$ to ϵ: $-\frac{\hbar^2}{2m}[\psi'(\epsilon) - \psi'(-\epsilon)] - \alpha\psi(0) = E\int_{-\epsilon}^{\epsilon}\psi\,dx$. Let $\epsilon \to 0$; the right side vanishes. ∎

Theorem 14.12 (Bound state of the delta well). *The attractive delta potential has exactly one bound state:*

$$\psi(x) = \sqrt{\kappa}\,e^{-\kappa|x|}, \qquad \kappa = \frac{m\alpha}{\hbar^2}, \qquad E = -\frac{m\alpha^2}{2\hbar^2}.$$

Proof. For $E < 0$ write $\kappa = \sqrt{-2mE}/\hbar$. Away from 0, $\psi'' = \kappa^2\psi$, and normalizability selects $e^{-\kappa x}$ for $x > 0$ and $e^{\kappa x}$ for $x < 0$. Continuity gives $\psi = Ae^{-\kappa|x|}$, so $\Delta\psi' = -2\kappa A$. By Lemma 14.11, $-2\kappa A = -\frac{2m\alpha}{\hbar^2}A$, so $\kappa = m\alpha/\hbar^2$. Only one value is possible. Normalize: $|A|^2 \cdot \frac{2}{2\kappa} = 1$. ∎

Theorem 14.13 (Scattering by the delta well). *For a particle incident from the left with energy $E > 0$, the reflection and transmission probabilities are*

$$R = \frac{1}{1 + 2\hbar^2 E/m\alpha^2}, \qquad T = \frac{1}{1 + m\alpha^2/2\hbar^2 E}, \qquad R + T = 1.$$

Proof. With $k = \sqrt{2mE}/\hbar$, take $\psi = e^{ikx} + re^{-ikx}$ for $x < 0$ (incident plus reflected) and $\psi = te^{ikx}$ for $x > 0$ (transmitted only). Continuity gives $1 + r = t$. The jump condition gives $ikt - ik(1 - r) = -\frac{2m\alpha}{\hbar^2}t$. With $\beta = \frac{m\alpha}{\hbar^2 k}$, these give $r = \frac{i\beta}{1-i\beta}$ and $t = \frac{1}{1-i\beta}$. Since the incident, reflected, and transmitted waves have the same k, the probability currents are in the ratio $1 : |r|^2 : |t|^2$. So $R = |r|^2 = \frac{\beta^2}{1+\beta^2}$ and $T = |t|^2 = \frac{1}{1+\beta^2}$. Substitute $\beta^2 = \frac{m\alpha^2}{2\hbar^2 E}$. ∎

A classical particle passes over an attractive well without reflection. A quantum particle sometimes bounces off, and does so mostly at low energy. The same formulas hold for a repulsive delta barrier ($\alpha < 0$), since only α^2 appears. A classical particle with $E > 0$ would sail over a barrier of zero width; a quantum particle is sometimes reflected, and a particle with any $E > 0$ can be transmitted.

14.6 The Finite Square Well

Let $V = -V_0$ for $|x| < a$ and $V = 0$ for $|x| > a$, with $V_0 > 0$. For bound states, $-V_0 < E < 0$. Define

$$k = \frac{\sqrt{2m(E + V_0)}}{\hbar} \text{ (inside)}, \qquad \kappa = \frac{\sqrt{-2mE}}{\hbar} \text{ (outside)},$$
$$z = ka, \qquad z_0 = \frac{a}{\hbar}\sqrt{2mV_0}.$$

Then $k^2 + \kappa^2 = 2mV_0/\hbar^2$, so $(\kappa a)^2 = z_0^2 - z^2$. The single parameter z_0 measures the strength of the well.

Since V is even, the solutions can be chosen even or odd (Exercise 13.19). For even solutions, $\psi = \cos kx$ inside and $Ae^{-\kappa|x|}$ outside. Continuity of ψ and ψ' at $x = a$ gives $\cos ka = Ae^{-\kappa a}$ and $-k \sin ka = -\kappa Ae^{-\kappa a}$. Dividing, $\kappa = k \tan ka$. Odd solutions, with $\sin kx$ inside, give $\kappa = -k \cot ka$ in the same way.

Theorem 14.14 (Bound states of the finite well). *The bound-state energies are $E = -V_0 + \frac{\hbar^2 z^2}{2ma^2}$, where z runs over the solutions in $(0, z_0)$ of*

$$\tan z = \sqrt{\frac{z_0^2}{z^2} - 1} \quad (\textit{even}), \qquad -\cot z = \sqrt{\frac{z_0^2}{z^2} - 1} \quad (\textit{odd}).$$

There is always at least one bound state, which is even, and the total number of bound states is $N = \lceil 2z_0/\pi \rceil$.

Proof. Divide $\kappa a = ka \tan ka$ by z to get the even equation, and likewise for odd. The right side decreases from $+\infty$ at $z = 0^+$ to 0 at $z = z_0$. The function $\tan z$ rises from 0 to ∞ on each interval $[j\pi, j\pi + \frac{\pi}{2})$ and is negative elsewhere. So there is one even solution in each such interval that begins below z_0, and the first interval $[0, \pi/2)$ always qualifies. Similarly $-\cot z$ is positive on $(j\pi + \frac{\pi}{2}, (j+1)\pi)$, which gives one odd solution in each such interval starting below z_0. Together, there is one solution per interval of length $\frac{\pi}{2}$ that starts below z_0. ∎

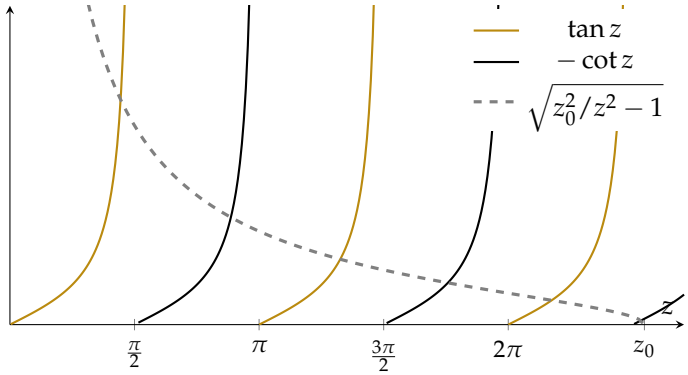


Figure 14.3: Graphical solution of the finite-well equations for $z_0 = 8$. Each crossing of the dashed curve gives a bound state, alternately even (gold) and odd (black); here there are $\lceil 16/\pi \rceil = 6$.

Limits. For a deep, wide well ($z_0 \to \infty$), the crossings approach $z = \frac{n\pi}{2}$, so $E + V_0 \approx \frac{n^2\pi^2\hbar^2}{2m(2a)^2}$: the infinite-well levels for a well of width $2a$. For a shallow, narrow well ($z_0 \to 0$), only one state survives, with $z \approx z_0$ and $\kappa a \approx z_0^2$. Its binding energy is $\frac{\hbar^2\kappa^2}{2m} \approx \frac{m(2aV_0)^2}{2\hbar^2}$, which is the delta-well result of Theorem 14.12 with $\alpha = 2aV_0$, the area of the well.

Penetration. Outside the well the bound-state wavefunction is $Ae^{-\kappa|x|}$, which is nonzero in the classically forbidden region. The *penetration depth* $1/\kappa$ grows as the binding energy decreases. A weakly bound particle spends much of its time where classically it cannot be.

Example 14.15 (Counting bound states). An electron is in a well of width $2a = 1.0$ nm and depth $V_0 = 5.0$ eV $= 8.01 \times 10^{-19}$ J. Then

$$z_0 = \frac{a}{\hbar}\sqrt{2m_eV_0} = \frac{(0.5 \times 10^{-9})\sqrt{2(9.11 \times 10^{-31})(8.01 \times 10^{-19})}}{1.055 \times 10^{-34}} \approx 5.73,$$

so there are $\lceil 2(5.73)/\pi \rceil = 4$ bound states.

14.7 Steps, Barriers, and Tunneling

14.7.1 THE POTENTIAL STEP

Let $V = 0$ for $x < 0$ and $V = V_0 > 0$ for $x > 0$, with a particle incident from the left.

Theorem 14.16 (Step potential). *For $E > V_0$, with $k = \sqrt{2mE}/\hbar$ and $q = \sqrt{2m(E - V_0)}/\hbar$,*

$$R = \left(\frac{k-q}{k+q}\right)^2, \qquad T = \frac{4kq}{(k+q)^2}.$$

For $E < V_0$, $R = 1$, but the wavefunction penetrates the step as $e^{-\kappa x}$, with $\kappa = \sqrt{2m(V_0 - E)}/\hbar$.

Proof. Take $\psi = e^{ikx} + re^{-ikx}$ for $x < 0$ and te^{iqx} for $x > 0$. Continuity of ψ and ψ' gives $1 + r = t$ and $k(1 - r) = qt$, so $r = \frac{k-q}{k+q}$ and $t = \frac{2k}{k+q}$. The transmitted current is $\frac{\hbar q}{m}|t|^2$, while the incident current is $\frac{\hbar k}{m}$. So $T = \frac{q}{k}|t|^2$, not $|t|^2$. For $E < V_0$, replace q by $i\kappa$; then $|r| = 1$. ∎

Note that $T \neq |t|^2$ when the speeds on the two sides differ. Transmission is a ratio of currents, not of amplitudes. A classical particle with $E > V_0$ would always continue over the step. A quantum particle is partly reflected even by a downward step, $V_0 < 0$, which is impossible classically.

14.7.2 TUNNELING THROUGH A BARRIER

Let $V = V_0$ for $0 < x < L$ and $V = 0$ elsewhere. A classical particle with $E < V_0$ is always reflected.

Theorem 14.17 (Rectangular barrier). *For $0 < E < V_0$, with $\kappa = \sqrt{2m(V_0 - E)}/\hbar$,*

$$T = \left[1 + \frac{V_0^2 \sinh^2(\kappa L)}{4E(V_0 - E)}\right]^{-1}.$$

For a thick barrier ($\kappa L \gg 1$), $T \approx \dfrac{16E(V_0 - E)}{V_0^2}\, e^{-2\kappa L}$.

Proof. Take $\psi = e^{ikx} + re^{-ikx}$ for $x < 0$, $Ce^{\kappa x} + De^{-\kappa x}$ inside, and te^{ikx} for $x > L$. Matching ψ and ψ' at $x = 0$ and $x = L$ gives four linear equations in r, C, D, t.

Eliminating C and D yields

$$t = \frac{e^{-ikL}}{\cosh \kappa L + i\frac{\kappa^2 - k^2}{2k\kappa}\sinh \kappa L}.$$

Then $T = |t|^2 = \left[\cosh^2 \kappa L + \left(\frac{\kappa^2-k^2}{2k\kappa}\right)^2 \sinh^2 \kappa L\right]^{-1}$. Use $\cosh^2 = 1 + \sinh^2$ and $1 + \left(\frac{\kappa^2-k^2}{2k\kappa}\right)^2 = \frac{(\kappa^2+k^2)^2}{4k^2\kappa^2} = \frac{V_0^2}{4E(V_0-E)}$. For large κL, $\sinh \kappa L \approx \frac{1}{2}e^{\kappa L}$. ■

The transmission probability is nonzero, though exponentially small: this is *quantum tunneling*. The exponential sensitivity to κL is the essential feature. Doubling the barrier width squares the already small transmission.

Example 14.18 (Electron tunneling). An electron with $E = 1.0$ eV meets a barrier of height 5.0 eV and width 0.50 nm. Then $\kappa = \sqrt{2m_e(4.0\text{ eV})}/\hbar = 1.02 \times 10^{10}\ \text{m}^{-1}$, so $\kappa L = 5.12$, and

$$T \approx \frac{16(1)(4)}{25} e^{-10.24} \approx 9.2 \times 10^{-5}.$$

Widening the barrier to 0.60 nm reduces T by a further factor $e^{-2\kappa(0.1\text{ nm})} \approx 0.13$.

This exponential sensitivity is exploited by the *scanning tunneling microscope*. A sharp tip is held a fraction of a nanometer above a conducting surface, and the tunneling current is measured. Because the current changes by about an order of magnitude for each 0.1 nm change in the gap, scanning the tip and holding the current fixed maps the surface with resolution finer than the size of a single atom.

14.7.3 THE WKB APPROXIMATION

For a barrier of general shape, approximate it by many thin slices, each a rectangular barrier. Multiplying their transmission factors and keeping only the exponentials gives

$$T \approx \exp\left(-\frac{2}{\hbar}\int_{x_1}^{x_2} \sqrt{2m\big(V(x) - E\big)}\, dx\right), \tag{14.5}$$

where x_1, x_2 are the classical turning points, at which $V = E$. This is the *WKB* tunneling formula, after Wentzel, Kramers, and Brillouin. It is accurate when the exponent is large and V varies slowly on the scale of the local penetration depth.

Example 14.19 (Alpha decay). A nucleus emits an alpha particle of energy E (a few MeV) by tunneling through the Coulomb barrier $V(r) = \frac{2Ze^2}{4\pi\varepsilon_0 r}$, which rises to about 25 MeV at the nuclear radius r_1. In the approximation $r_1 \ll r_2$, where $r_2 = \frac{2Ze^2}{4\pi\varepsilon_0 E}$ is the outer turning point, the WKB integral evaluates to

$$T \approx e^{-2G}, \qquad G \approx \frac{\pi}{2}\,\frac{2Ze^2}{4\pi\varepsilon_0\hbar}\sqrt{\frac{2m_\alpha}{E}}$$

(Exercise 14.34). The lifetime is inversely proportional to T, so $\ln\tau$ is linear in $Z/\sqrt{E}$: the *Geiger–Nuttall law*. Because G is large and enters an exponential, a modest change in E changes the lifetime enormously. Polonium-212 ($E = 8.95$ MeV) lives 0.3 μs, while uranium-238 ($E = 4.27$ MeV) lives 4.5 billion years: a factor of 10^{23} from a factor of two in energy. This explanation, given by Gamow and by Gurney and Condon in 1928, was one of the first triumphs of quantum mechanics in nuclear physics.

LOOKING AHEAD. Tunneling is everywhere once one looks for it. It makes nuclear fusion in the Sun possible at its core temperature of 1.5×10^7 K, where protons classically lack the energy to touch. It operates the tunnel diode, flash memory, and the Josephson junctions of superconducting qubits. It sets limits on how small transistors can be made. The formal theory, using the transfer matrices of the barrier proof, extends to periodic potentials and produces the energy *bands* of electrons in crystals, which explain why some materials conduct and others insulate.

14.8 Worked Problems

The problems below solve several potentials exactly, by completing the square, by matching at a delta function, or by recognizing a special solution, and test two approximations against exact results: the deep-well limit of the finite well and the semiclassical quantization of a linear potential.

Problem 14.1 An oscillator in a uniform field

A particle of charge q in the oscillator potential is placed in a uniform electric field $\mathcal{E}$, so $V = \frac{1}{2}m\omega^2x^2 - q\mathcal{E}x$. Find the energies and eigenfunctions exactly, the induced dipole moment in each state, and the polarizability.

Solution. Complete the square. With $x_0 = \frac{q\mathcal{E}}{m\omega^2}$,

$$V = \frac{1}{2}m\omega^2(x - x_0)^2 - \frac{q^2\mathcal{E}^2}{2m\omega^2}.$$

This is the original oscillator, shifted by x_0 and lowered by a constant. Since translation does not change the kinetic energy, the eigenfunctions are $\psi_n(x - x_0)$ and the energies are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{q^2\mathcal{E}^2}{2m\omega^2}.$$

Every level drops by the same amount, so the spectrum of transition frequencies is unchanged. In every state $\langle x\rangle = x_0$, so the induced dipole moment is $d = q\langle x\rangle = \frac{q^2}{m\omega^2}\mathcal{E}$, the same for all n, and the polarizability is $\alpha = d/\mathcal{E} = q^2/m\omega^2$, identical to the classical value. Consistently, $E_n = E_n^{(0)} - \frac{1}{2}\alpha\mathcal{E}^2$.

This exact result is a test for perturbation theory. First-order perturbation theory gives zero shift, since $\langle n|\hat{x}|n\rangle = 0$, and second-order theory must reproduce $-\frac{q^2\mathcal{E}^2}{2m\omega^2}$ exactly. It does, as a sum over the two neighbors $n \pm 1$ that $\hat{x}$ connects (Example 15.13).

Problem 14.2 The deep finite well

For a finite well of width $2a$ with $z_0 \gg 1$, show that the ground state has $z \approx \frac{\pi}{2}\frac{z_0}{z_0+1}$, and interpret the result as an effective width. Compare with the numerical root for $z_0 = 10$.

Solution. The even condition of Theorem 14.14 is $z\tan z = \sqrt{z_0^2 - z^2}$. For $z_0 \gg 1$ the ground-state root lies just below $\frac{\pi}{2}$, where $\tan z$ is large. Put $z = \frac{\pi}{2} - \epsilon$ with ϵ small. Then $\tan z = \cot\epsilon \approx \frac{1}{\epsilon}$, and the right side is $\approx z_0$ since $z \ll z_0$. So $\frac{z}{\epsilon} \approx z_0$, $\epsilon \approx \frac{z}{z_0}$, and

$$z \approx \frac{\pi}{2} - \frac{z}{z_0} \quad\Longrightarrow\quad z \approx \frac{\pi}{2}\,\frac{z_0}{z_0+1}.$$

For $z_0 = 10$ this gives 1.42800, against the numerical root 1.42755. For $z_0 = 30$, 1.520125 against 1.520104.

The energy above the bottom is $\frac{\hbar^2 z^2}{2ma^2} = \frac{\pi^2\hbar^2}{2m\left(2a(1+1/z_0)\right)^2}$. This is the ground-state energy of an infinite well of width $2a(1 + 1/z_0)$. Since $\kappa \approx z_0/a$ for a deep level, the extra width is $2a/z_0 \approx 2/\kappa$: the wavefunction leaks one penetration depth $1/\kappa$ into each wall, and the particle behaves as if the box were that much wider. The level is lower than in an infinite well of width $2a$, as it must be, since the finite well is a weaker confinement.

Problem 14.3 A delta well near a wall

A particle moves on $x > 0$ with an impenetrable wall at $x = 0$ and an attractive delta potential $V = -\alpha\delta(x - a)$. Show that a bound state exists if and only if $a > \frac{\hbar^2}{2m\alpha}$. How does its energy compare with that of the delta well without a wall?

Solution. For $E = -\frac{\hbar^2\kappa^2}{2m} < 0$, the solution vanishing at 0 and decaying at infinity is

$$\psi = \begin{cases} \sinh \kappa x, & 0 < x < a, \\ \sinh(\kappa a)\, e^{-\kappa(x-a)}, & x > a, \end{cases}$$

continuous at a. Integrating the Schrödinger equation across $x = a$ (Lemma 14.11) gives $\psi'(a^+) - \psi'(a^-) = -\frac{2m\alpha}{\hbar^2}\psi(a)$, that is,

$$-\kappa \sinh \kappa a - \kappa \cosh \kappa a = -g \sinh \kappa a, \qquad g = \frac{2m\alpha}{\hbar^2},$$

or $\kappa\big(1 + \coth \kappa a\big) = g$. The left side increases from $\frac{1}{a}$ (as $\kappa \to 0^+$, since $\kappa \coth \kappa a \to \frac{1}{a}$) to ∞, so a solution with $\kappa > 0$ exists if and only if $g > \frac{1}{a}$, that is, $a > \frac{\hbar^2}{2m\alpha}$.

Without the wall, the single bound state has $\kappa = g/2$ (Theorem 14.12). With the wall, $1 + \coth \kappa a > 2$, so $\kappa < g/2$: the state is less tightly bound. For example, with $g = 3$ and $a = 1$, $\kappa = 1.411$ against 1.5. The wall forces $\psi(0) = 0$, which costs kinetic energy. As $a \to \infty$, $\coth \kappa a \to 1$ and the free result returns. As a decreases to $\frac{1}{g}$, the state is squeezed out of existence. In three dimensions this is the situation of an s-wave in a short-range attractive potential, where the radial function must vanish at the origin, and the reason that a weak enough three-dimensional well has no bound state at all.

Problem 14.4 Semiclassical quantization of a V-shaped well

For $V = F|x|$, the Bohr–Sommerfeld rule with the turning-point correction, $\oint p\,dx = \big(n + \frac{1}{2}\big)h$, approximates the energies. Find them, and compare with the exact energies in units where $\hbar = 1$, $2m = 1$, $F = 1$, which are $E = |a_k'|$ (even states) and $|a_k|$ (odd states), where a_k and a_k' are the zeros of the Airy function Ai and of its derivative.

Solution. At energy E the turning points are $\pm x_0$ with $x_0 = E/F$. The integral over a full cycle is twice the integral from $-x_0$ to x_0:

$$\oint p\,dx = 2 \cdot 2 \int_0^{x_0} \sqrt{2m(E - Fx)}\,dx = 4\sqrt{2m} \cdot \frac{2E^{3/2}}{3F} = \frac{8\sqrt{2m}\,E^{3/2}}{3F}.$$

Setting this equal to $2\pi\hbar\left(n+\tfrac{1}{2}\right)$,

$$E_n = \left[\frac{3\pi\hbar F}{4\sqrt{2m}}\left(n+\frac{1}{2}\right)\right]^{2/3}.$$

In the stated units, $E_n = \left(\frac{3\pi}{4}(n+\frac{1}{2})\right)^{2/3}$. For $x>0$ the exact solution is $\psi = \mathrm{Ai}(x-E)$, and the even and odd conditions $\psi'(0)=0$ and $\psi(0)=0$ give the Airy zeros. The comparison:

n	0	1	2	3	4	5
Exact	1.0188	2.3381	3.2482	4.0879	4.8201	5.5206
Semiclassical	1.1155	2.3203	3.2616	4.0818	4.8263	5.5172

The error is 9% for the ground state and falls below 0.1% by $n=5$. The semiclassical rule is exact in the limit of large quantum numbers, the correspondence principle in quantitative form, and the $\frac{1}{2}$ is essential: without it, the ground state would be at $E=0$. The levels grow like $n^{2/3}$, between the oscillator's n^1 and the box's n^2, because the linear potential is softer than a quadratic one. The half-problem, $x>0$ with a hard floor at $x=0$ (only the odd states), describes a neutron bouncing on a horizontal mirror in gravity. Its quantized heights were observed at the Institut Laue–Langevin in 2002.

Problem 14.5 The thermal oscillator

An oscillator in equilibrium at temperature T occupies ψ_n with probability $p_n \propto e^{-E_n/k_BT}$. Find $\langle E\rangle$ and $\langle x^2\rangle$, and their limits for $k_BT \gg \hbar\omega$ and $k_BT \ll \hbar\omega$.

Solution. With $\beta = 1/k_BT$, the partition function is $Z = \sum_n e^{-\beta(n+\frac{1}{2})\hbar\omega} = \dfrac{e^{-\beta\hbar\omega/2}}{1-e^{-\beta\hbar\omega}} = \dfrac{1}{2\sinh(\beta\hbar\omega/2)}$. Then

$$\langle E\rangle = -\frac{\partial \ln Z}{\partial\beta} = \frac{\hbar\omega}{2}\coth\frac{\beta\hbar\omega}{2} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\beta\hbar\omega}-1}.$$

This is Planck's mean energy (Theorem 12.3) plus the zero-point energy that Planck's model lacked. For $\langle x^2\rangle$, use $\langle x^2\rangle_n = \frac{\hbar}{2m\omega}(2n+1) = \frac{E_n}{m\omega^2}$ from Example 14.7. Averaging over n,

$$\langle x^2\rangle_T = \frac{\langle E\rangle}{m\omega^2} = \frac{\hbar}{2m\omega}\coth\frac{\hbar\omega}{2k_BT}.$$

For $k_BT \gg \hbar\omega$, $\coth u \approx \frac{1}{u}$ gives $\langle E\rangle \approx k_BT$ and $\langle x^2\rangle \approx \frac{k_BT}{m\omega^2}$, the classical equipartition values. For $k_BT \ll \hbar\omega$, $\coth u \to 1$ and $\langle x^2\rangle \to \frac{\hbar}{2m\omega}$: the zero-point

fluctuations remain at absolute zero. The crossover temperature is $\hbar\omega/k_B$. For a 1 GHz mechanical resonator it is 48 mK, which dilution refrigerators can reach. Experiments since 2010 have cooled such resonators to their quantum ground state.

Problem 14.6 Momentum in the infinite well

For the stationary state $\psi_n = \sqrt{2/a}\sin(n\pi x/a)$ on $(0,a)$, find the momentum distribution $|\phi_n(k)|^2$, with $p = \hbar k$. Where is it peaked for $n = 1$, and for large n?

Solution. Write $\sin\frac{n\pi x}{a} = \frac{1}{2i}\left(e^{in\pi x/a} - e^{-in\pi x/a}\right)$ and integrate each exponential over $(0,a)$. After simplification,

$$|\phi_n(k)|^2 = \frac{1}{2\pi}\left|\int_0^a \psi_n e^{-ikx}\,dx\right|^2 = \frac{4\pi n^2 a}{\left((n\pi)^2 - (ka)^2\right)^2} \times \begin{cases} \cos^2(ka/2), & n \text{ odd}, \\ \sin^2(ka/2), & n \text{ even}. \end{cases}$$

The apparent poles at $ka = \pm n\pi$ are cancelled by zeros of the numerator, so $|\phi_n|^2$ is smooth. (A numerical check of the integral at several k and n agrees with the formula to 10^{-9}.)

For $n = 1$ the distribution is peaked at $k = 0$, not at the "classical" values $\pm\pi/a$. The ground state is a single half-wavelength, too short to have a definite wavelength. For large n the distribution has two sharp peaks near $k = \pm n\pi/a$: for $n = 10$ the maximum is at $ka = 31.22$, close to $10\pi = 31.42$. The particle moves back and forth with speed $p/m = n\pi\hbar/(ma)$, as a classical particle of energy E_n would. Note that ψ_n is an eigenstate of $\hat{p}^2$ inside the well but not of $\hat{p}$, and the momentum distribution is continuous: a measurement of p can give any value, though values near $\pm\sqrt{2mE_n}$ dominate when n is large.

Problem 14.7 A reflectionless well

Show that $V(x) = -\dfrac{\hbar^2}{ma^2}\operatorname{sech}^2\dfrac{x}{a}$ has the bound state $\psi = \operatorname{sech}\dfrac{x}{a}$ with $E = -\dfrac{\hbar^2}{2ma^2}$. Is it the ground state?

Solution. Let $u = x/a$ and write $s = \operatorname{sech} u$, $t = \tanh u$, with $s' = -st$ and $t' = s^2$ (derivatives in u). Then

$$\psi'' = \frac{1}{a^2}(-st)' = \frac{1}{a^2}\left(st^2 - s^3\right) = \frac{1}{a^2}\left(s(1-s^2) - s^3\right) = \frac{1}{a^2}\left(s - 2s^3\right).$$

So $-\frac{\hbar^2}{2m}\psi'' + V\psi = -\frac{\hbar^2}{2ma^2}(s - 2s^3) - \frac{\hbar^2}{ma^2}s^3 = -\frac{\hbar^2}{2ma^2}s$, which is $E\psi$ with $E = -\frac{\hbar^2}{2ma^2}$. The function $\operatorname{sech} u$ decays like $2e^{-|u|}$, so it is normalizable.

It is the ground state, because it has no nodes. (The ground state of a one-dimensional potential is the unique nodeless bound state, by the oscillation theorem of Sturm–Liouville theory, Theorem 10.14.) In fact it is the only bound state: this member of the *Pöschl–Teller* family, $V = -\frac{\hbar^2}{2ma^2}\lambda(\lambda+1)\,\mathrm{sech}^2\frac{x}{a}$ with $\lambda = 1$, has exactly λ bound states. And for every positive energy, it transmits with probability exactly 1. No wave is reflected, at any energy, in sharp contrast with the square well (Theorem 14.14) and the delta well (Theorem 14.13). Reflectionless potentials of this family describe the solitons of the Korteweg–de Vries equation, water waves that pass through each other unchanged.

EXERCISES

A. COMPUTATION

14.1 An electron is in an infinite well of width 0.20 nm. Find the energies of the first three levels in eV, and the wavelength of the photon emitted in the $3 \to 1$ transition.

14.2 For the infinite well, compute $\langle x\rangle$, $\langle x^2\rangle$, $\langle p\rangle$, $\langle p^2\rangle$, and $\sigma_x\sigma_p$ in the state ψ_n. For which n is the product smallest?

14.3 A particle in the infinite well starts in $\Psi(x,0) = \frac{1}{\sqrt{2}}(\psi_1+\psi_2)$. Find $\langle x\rangle(t)$ and its angular frequency and amplitude.

14.4 A particle in the infinite well starts in $\Psi(x,0) = A$ (constant) on $(0,a)$. Find c_n and the probability of measuring E_1. Why does $\langle H\rangle$ diverge?

14.5 The well suddenly expands from width a to $2a$ while the particle is in the ground state, so that the wavefunction has no time to change. Find the probability that the particle is in the new ground state.

14.6 Find the revival time for an electron in an infinite well of width 1.0 nm.

14.7 Verify $[\hat{a}, \hat{a}^\dagger] = 1$ by acting on a test function.

14.8 Compute $\psi_1(x)$ for the oscillator by applying $\hat{a}^\dagger$ to ψ_0, and verify that it is normalized.

14.9 Compute $\psi_2(x)$ for the oscillator, and verify that it is orthogonal to ψ_0.

14.10 Find $\langle p^2\rangle$ in ψ_n using ladder operators.

14.11 Find $\langle 0|\hat{x}^3|1\rangle$ and $\langle 1|\hat{x}^2|3\rangle$ for the oscillator.

14.12 An oscillator is in the state $\Psi(x,0) = \frac{1}{\sqrt{2}}(\psi_0 + \psi_1)$. Find $\langle x\rangle(t)$ and $\langle p\rangle(t)$, and verify Ehrenfest's theorem.

14.13 Find the probability that a particle in the oscillator ground state is found outside the classically allowed region.

14.14 The vibration of the CO molecule has $\hbar\omega = 0.266$ eV. Find the wavelength of the emitted photon, the zero-point energy, and the classical amplitude of zero-point motion, using the reduced mass 1.14×10^{-26} kg.

14.15 Find the bound-state energy and penetration depth for an electron in a delta well with $\alpha = 1.0$ eV nm.

14.16 For the delta well of the previous exercise, find R and T for an electron with $E = 0.10$ eV and $E = 10$ eV.

14.17 Find $\langle x^2\rangle$ and σ_p for the delta-well bound state, and check the uncertainty principle.

14.18 A finite well has $z_0 = 4$. Find the number of bound states and, numerically, the values of z.

14.19 A proton is in a nuclear well of depth 35 MeV and width $2a = 4.0$ fm. How many bound states are there?

14.20 For the potential step with $E = 2V_0$, find R and T, and verify $R + T = 1$.

14.21 An electron with $E = 2.0$ eV meets a 3.0 eV barrier of width 0.30 nm. Find T exactly and in the thick-barrier approximation.

14.22 Use the WKB formula to find the tunneling probability through the triangular barrier $V = V_0 - Fx$ for $x > 0$ (and $V = 0$ for $x < 0$) at energy $E < V_0$. (This is field emission of electrons from a metal.)

14.23 Find the transmission resonances of the finite square well for $E > 0$: show that $T = 1$ when $2qa = n\pi$, where $q = \sqrt{2m(E + V_0)}/\hbar$.

B. THEORY AND PROOF

14.24 Prove that the bound states of a one-dimensional potential are nondegenerate. (If ψ_1, ψ_2 have the same E, show that their Wronskian is constant and vanishes at infinity.)

14.25 Prove that a bound state of a one-dimensional potential can be chosen real.

14.26 Show that $\hat{a}$ has eigenvectors (*coherent states*) $|\alpha\rangle = e^{-|\alpha|^2/2}\sum_n \frac{\alpha^n}{\sqrt{n!}}|n\rangle$ for every $\alpha \in \mathbb{C}$, and that $|\alpha\rangle$ evolves into $|\alpha e^{-i\omega t}\rangle$ up to a phase. Show that $\langle x\rangle$ oscillates classically.

14.27 Prove that $\sigma_x\sigma_p = \hbar/2$ in every coherent state.

14.28 Show that in the oscillator, $\hat{a}(t) = \hat{a}e^{-i\omega t}$ in the Heisenberg picture, and deduce $\hat{x}(t) = \hat{x}\cos\omega t + \frac{\hat{p}}{m\omega}\sin\omega t$.

14.29 Derive the odd-state condition $\kappa = -k\cot ka$ for the finite square well.

14.30 Show that a delta-function barrier $V = \alpha\delta(x)$ with $\alpha > 0$ has no bound states.

14.31 Show that for the step with $E > V_0$, $R + T = 1$.

14.32 Verify that probability current is the same on both sides of a barrier: $J_{\text{left}} = J_{\text{right}}$, and deduce $R + T = 1$.

14.33 For the double delta well $V = -\alpha[\delta(x-a) + \delta(x+a)]$, find the equations for the even and odd bound states and show that there are one or two bound states depending on $m\alpha a/\hbar^2$.

C. CHALLENGE

14.34 (Gamow factor.) Carry out the WKB integral for the Coulomb barrier, $\int_{r_1}^{r_2}\sqrt{\frac{r_2}{r}-1}\,dr$, exactly, and show it tends to $\frac{\pi}{2}r_2$ when $r_1 \ll r_2$. Hence verify the formula for G in Example 14.19.

14.35 (Kronig–Penney.) A periodic array of delta functions $V = \alpha\sum_n \delta(x - na)$ has solutions $\psi(x+a) = e^{iKa}\psi(x)$ (Bloch's theorem). Show that the allowed energies satisfy $\cos Ka = \cos ka + \frac{m\alpha}{\hbar^2 k}\sin ka$, and explain why this produces energy bands separated by gaps.

14.36 Show that the half-oscillator, with $V = \frac{1}{2}m\omega^2x^2$ for $x > 0$ and $V = \infty$ for $x < 0$, has energies $(2n + \frac{3}{2})\hbar\omega$, the odd states of the full oscillator.

14.37 Find the stationary states of the infinite well $0 < x < a$ to first order in a small perturbation $V' = \epsilon$ for $0 < x < a/2$, by computing $\langle\psi_n|V'|\psi_n\rangle$. (This is first-order perturbation theory.)

CHAPTER 15

The Formalism: Hilbert Space, Operators, and Uncertainty

LEARNING OBJECTIVES

- State the postulates of quantum mechanics in the language of Hilbert spaces and Dirac notation.
- Apply the general measurement rule, including degenerate eigenvalues and continuous spectra.
- Prove the generalized uncertainty principle for any two observables, and derive the energy–time relation.
- Compute the time dependence of expectation values, identify conserved quantities, and use the Heisenberg picture.
- Explain how symmetries generate conservation laws, with momentum as the generator of translations.
- Prove the variational principle and use it to bound ground-state energies.
- Apply first- and second-order time-independent perturbation theory.

Chapters 13 and 14 worked with wavefunctions $\Psi(x,t)$ and differential operators. But the spin of an electron has no wavefunction in this sense: its state is a vector in $\mathbb{C}^2$, as the Stern–Gerlach experiment showed. And the position and momentum wavefunctions $\Psi(x)$ and $\Phi(p)$ are two descriptions of one state. The theory needs a framework in which all of these are instances of a single structure. That framework is linear algebra in Hilbert space, the subject of Chapter 9. This chapter states the postulates of quantum mechanics in that language and derives their most important consequences. It then develops two approximation methods that make the theory usable for problems that cannot be solved exactly, which is almost all of them.

15.1 The Postulates

Definition 15.1 (Postulates of quantum mechanics). (I) **States.** The state of an isolated system is represented by a unit vector $|\psi\rangle$ in a complex Hilbert space $\mathcal{H}$. Vectors that differ by a phase factor $e^{i\alpha}$ represent the same state.

(II) **Observables.** Each measurable quantity A is represented by a Hermitian operator $\hat{A}$ on $\mathcal{H}$.

(III) **Measurement.** A measurement of A yields one of the eigenvalues of $\hat{A}$. If $\hat{A}$ has discrete spectrum with spectral decomposition $\hat{A} = \sum_a a\,\hat{P}_a$ (Corollary 9.28), the probability of the result a is

$$P(a) = \langle\psi|\hat{P}_a|\psi\rangle = \|\hat{P}_a|\psi\rangle\|^2,$$

and immediately after the measurement the state is $\hat{P}_a|\psi\rangle / \|\hat{P}_a|\psi\rangle\|$.

(IV) **Dynamics.** Between measurements, the state evolves by the Schrödinger equation $i\hbar\, \mathrm{d}/\mathrm{d}t\, |\psi(t)\rangle = \hat{H}|\psi(t)\rangle$, where the Hermitian operator $\hat{H}$ is the Hamiltonian.

When the eigenvalue a is nondegenerate, with unit eigenvector $|a\rangle$, then $\hat{P}_a = |a\rangle\langle a|$ and the measurement rule reduces to the familiar Born rule $P(a) = |\langle a|\psi\rangle|^2$. The state after the measurement is $|a\rangle$. The mean of many measurements is

$$\langle A\rangle = \sum_a a\,P(a) = \sum_a a\langle\psi|\hat{P}_a|\psi\rangle = \langle\psi|\hat{A}|\psi\rangle,$$

which is (13.2) in abstract form. Every consequence of Chapters 13–14 follows from these postulates with the choice $\mathcal{H} = L^2(\mathbb{R})$ and $\hat{x}$, $\hat{p} = -i\hbar\, d/dx$.

Example 15.2 (A degenerate measurement). In $\mathbb{C}^3$, let $\hat{A} = \mathrm{diag}(1,1,2)$ and $|\psi\rangle = \tfrac{1}{3}\,\langle 1,2,2\rangle$. The eigenvalue 1 has eigenspace spanned by $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$, with projector $\hat{P}_1 = \mathrm{diag}(1,1,0)$. So $P(1) = \|\hat{P}_1\psi\|^2 = \frac{1+4}{9} = \frac{5}{9}$ and $P(2) = \frac{4}{9}$. If the result is 1, the new state is $\frac{1}{\sqrt{5}}\,\langle 1,2,0\rangle$. The relative amplitudes *within* the eigenspace are preserved. This is the projection postulate of Lüders, and it matters whenever an apparatus cannot distinguish states that share an eigenvalue.

15.1.1 CONTINUOUS SPECTRA

For observables with continuous spectra, such as position and momentum, the sums become integrals and the eigenvectors are normalized to delta functions (Chapters 9 and 11):

$$\hat{x}|x\rangle = x|x\rangle, \qquad \langle x|x'\rangle = \delta(x-x'), \qquad \int |x\rangle\langle x|\,dx = \hat{I};$$
$$\hat{p}|p\rangle = p|p\rangle, \qquad \langle x|p\rangle = \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}}.$$

The wavefunction is the component of the state in the position basis, $\psi(x) = \langle x|\psi\rangle$. The momentum-space wavefunction is $\Phi(p) = \langle p|\psi\rangle = \int \langle p|x\rangle\langle x|\psi\rangle\,dx$, which is the Fourier transform of Chapter 13, now derived by inserting completeness. Probabilities become densities: $|\langle x|\psi\rangle|^2 dx$ is the probability of finding the particle in $[x, x+dx]$.

15.1.2 COMPATIBLE OBSERVABLES

Observables A and B are *compatible* if $[\hat{A},\hat{B}] = 0$. By Theorem 9.30, they then have a common eigenbasis. Measuring A and then B leaves the system in a simultaneous eigenstate, and a further measurement of A returns the same value. For incompatible observables, as in the Stern–Gerlach sequence of Chapter 12, measuring B destroys the value of A.

A *complete set of commuting observables* is a set of mutually commuting observables whose joint eigenvalues label a unique basis state. For the hydrogen atom without spin, $\{\hat{H}, \hat{L}^2, \hat{L}_z\}$ is such a set, and the states are labeled $|n,\ell,m\rangle$ (Chapter 16).

15.2 THE GENERALIZED UNCERTAINTY PRINCIPLE

For an observable A in state $|\psi\rangle$, the *uncertainty* is $\sigma_A^2 = \langle(\hat{A}-\langle A\rangle)^2\rangle = \langle A^2\rangle - \langle A\rangle^2$.

Theorem 15.3 (Robertson uncertainty relation). *For any observables A, B and any state,*

$$\sigma_A\,\sigma_B \geq \frac{1}{2}\big|\langle[\hat{A},\hat{B}]\rangle\big|.$$

Proof. Let $|f\rangle = (\hat{A}-\langle A\rangle)|\psi\rangle$ and $|g\rangle = (\hat{B}-\langle B\rangle)|\psi\rangle$. Since these operators are

Hermitian, $\sigma_A^2 = \langle f|f\rangle$ and $\sigma_B^2 = \langle g|g\rangle$. By Cauchy–Schwarz (Theorem 9.12),

$$\sigma_A^2\sigma_B^2 \geq |\langle f|g\rangle|^2 \geq \big(\operatorname{Im}\langle f|g\rangle\big)^2 = \Big(\frac{\langle f|g\rangle - \langle g|f\rangle}{2i}\Big)^2.$$

Expanding, $\langle f|g\rangle = \langle AB\rangle - \langle A\rangle\langle B\rangle$ and $\langle g|f\rangle = \langle BA\rangle - \langle A\rangle\langle B\rangle$. So $\langle f|g\rangle - \langle g|f\rangle = \langle[\hat{A},\hat{B}]\rangle$, and the inequality becomes $\sigma_A^2\sigma_B^2 \geq \big(\frac{1}{2i}\langle[\hat{A},\hat{B}]\rangle\big)^2$. The quantity in parentheses is real, because the commutator of Hermitian operators is anti-Hermitian. Take square roots. ∎

With $[\hat{x},\hat{p}] = i\hbar$ this gives $\sigma_x\sigma_p \geq \frac{\hbar}{2}$ once again, now as a consequence of a commutator rather than of Fourier analysis. For spin, $[\hat{S}_x,\hat{S}_y] = i\hbar\hat{S}_z$ gives $\sigma_{S_x}\sigma_{S_y} \geq \frac{\hbar}{2}|\langle S_z\rangle|$. This bound depends on the state. In a state with $\langle S_z\rangle = 0$ the bound is trivial, and indeed S_x and S_y can then both have small uncertainty in a suitable superposition.

COMMON ERROR (what the uncertainty principle does not say). The Robertson relation concerns the spreads of results in two separate ensembles of identically prepared systems, one used to measure A and the other B. It does not by itself describe the disturbance one measurement causes to another, though related inequalities do. Nor does it say the particle "really has" a sharp position and momentum that we cannot know: the Stern–Gerlach sequence and, more decisively, the Bell experiments of Chapter 18 rule that out.

15.2.1 ENERGY AND TIME

Time is not an operator in quantum mechanics; it is a parameter. So there is no commutator $[\hat{t},\hat{H}]$ to insert into Theorem 15.3. Nevertheless an energy–time relation holds, with a precise meaning.

Theorem 15.4 (Mandelstam–Tamm). *For any observable Q not explicitly dependent on time, define $\Delta t_Q = \sigma_Q / \left|\, \mathrm{d}\langle Q\rangle/\mathrm{d}t \,\right|$, the time for $\langle Q\rangle$ to change by one standard deviation. Then*

$$\sigma_H\,\Delta t_Q \geq \frac{\hbar}{2}.$$

Proof. By Theorem 15.5 below, $\mathrm{d}\langle Q\rangle/\mathrm{d}t = \frac{i}{\hbar}\langle[\hat{H},\hat{Q}]\rangle$. By Theorem 15.3, $\sigma_H\sigma_Q \geq \frac{1}{2}|\langle[\hat{H},\hat{Q}]\rangle| = \frac{\hbar}{2}\left|\, \mathrm{d}\langle Q\rangle/\mathrm{d}t \,\right|$. Divide by $\left|\, \mathrm{d}\langle Q\rangle/\mathrm{d}t \,\right|$. ∎

A state that changes quickly must have a large spread of energies, and a state with sharp energy (a stationary state) does not change at all. An unstable state that lives for a time τ has an energy spread, its *natural linewidth*, of order $\Gamma \approx \hbar/\tau$. This connects the Lorentzian of Example 8.35 with exponential decay. The Δ meson lives about 5.6×10^{-24} s, and its measured mass is correspondingly spread by about 117 MeV.

15.3 Time Evolution

Theorem 15.5 (Time dependence of expectation values). *For any observable $\hat{Q}$, possibly depending explicitly on time,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Q\rangle = \frac{i}{\hbar}\big\langle [\hat{H},\hat{Q}]\big\rangle + \Big\langle \frac{\partial \hat{Q}}{\partial t}\Big\rangle.$$

Proof. Differentiate $\langle\psi|\hat{Q}|\psi\rangle$ by the product rule and use $\mathrm{d}/\mathrm{d}t\,|\psi\rangle = -\frac{i}{\hbar}\hat{H}|\psi\rangle$ and its adjoint, $\mathrm{d}/\mathrm{d}t\,\langle\psi| = \frac{i}{\hbar}\langle\psi|\hat{H}$ (here $\hat{H}^\dagger = \hat{H}$ is essential):

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Q\rangle = \frac{i}{\hbar}\langle\psi|\hat{H}\hat{Q}|\psi\rangle - \frac{i}{\hbar}\langle\psi|\hat{Q}\hat{H}|\psi\rangle + \langle\psi|\partial_t\hat{Q}|\psi\rangle. \qquad \blacksquare$$

Corollary 15.6 (Conservation laws). *An observable that does not depend explicitly on time and commutes with $\hat{H}$ is* conserved*: its expectation value, indeed its whole probability distribution, is constant in every state. In particular $\hat{H}$ itself is conserved.*

Ehrenfest's theorem is a special case. With $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$, one computes $[\hat{H},\hat{x}] = -\frac{i\hbar}{m}\hat{p}$ and $[\hat{H},\hat{p}] = i\hbar V'(\hat{x})$, using $[\hat{x},\hat{p}] = i\hbar$ and the commutator product rule. Substituting into Theorem 15.5 gives $\mathrm{d}\langle x\rangle/\mathrm{d}t = \langle p\rangle/m$ and $\mathrm{d}\langle p\rangle/\mathrm{d}t = -\langle V'\rangle$.

15.3.1 The evolution operator

For a time-independent Hamiltonian, the Schrödinger equation is solved by

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle, \qquad \hat{U}(t) = e^{-i\hat{H}t/\hbar} = \sum_n e^{-iE_nt/\hbar}|n\rangle\langle n|,$$

where $\hat{H}|n\rangle = E_n|n\rangle$. This is Theorem 13.18 in operator form. By Theorem 9.31, $\hat{U}(t)$ is unitary, which is the abstract reason that probability is conserved.

15.3.2 THE HEISENBERG PICTURE

Expectation values can be computed in two equivalent ways:

$$\langle Q\rangle(t) = \langle\psi(t)|\hat{Q}|\psi(t)\rangle = \langle\psi(0)|\hat{U}^\dagger\hat{Q}\hat{U}|\psi(0)\rangle.$$

In the *Schrödinger picture* the state evolves and operators are fixed. In the *Heisenberg picture* the state is fixed at $|\psi(0)\rangle$ and the operators evolve as $\hat{Q}_H(t) = \hat{U}^\dagger\hat{Q}\hat{U}$. Differentiating gives the *Heisenberg equation of motion*

$$\frac{d\hat{Q}_H}{dt} = \frac{i}{\hbar}[\hat{H}, \hat{Q}_H] + \left(\frac{\partial\hat{Q}}{\partial t}\right)_H, \tag{15.1}$$

which has the form of a classical equation of motion for operators. Heisenberg's matrix mechanics of 1925 was the theory in this picture. Its operator equations closely parallel Hamilton's equations of classical mechanics, with the commutator divided by $i\hbar$ playing the role of the classical Poisson bracket. Dirac recognized this correspondence and made it the basis of *canonical quantization*.

Example 15.7 (Oscillator in the Heisenberg picture). For $\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$, (15.1) gives $\dot{\hat{a}}_H = \frac{i}{\hbar}[\hat{H}, \hat{a}_H] = -i\omega\hat{a}_H$, so $\hat{a}_H(t) = \hat{a}e^{-i\omega t}$. By (14.2),

$$\hat{x}_H(t) = \hat{x}\cos\omega t + \frac{\hat{p}}{m\omega}\sin\omega t,$$

the classical solution with the initial position and momentum promoted to operators. Every expectation value of x follows the classical trajectory, in every state.

15.4 SYMMETRY AND CONSERVATION

In Chapter 4 we saw that $e^{a\,d/dx}f(x) = f(x+a)$: differentiation generates translations. Since $\hat{p} = -i\hbar\,d/dx$,

$$\hat{T}(a) = e^{-ia\hat{p}/\hbar}, \qquad \hat{T}(a)\psi(x) = \psi(x-a),$$

translates a wavefunction by a. $\hat{T}(a)$ is unitary because $\hat{p}$ is Hermitian, and $\hat{p}$ is called the *generator* of translations.

Theorem 15.8 (Symmetry implies conservation). *If the Hamiltonian is invariant under a one-parameter family of unitary symmetries, $\hat{U}(s) = e^{-is\hat{G}/\hbar}$, in the sense that $\hat{U}^\dagger\hat{H}\hat{U} = \hat{H}$ for all s, then the generator $\hat{G}$ is conserved.*

Proof. Differentiate $\hat{U}(s)^\dagger \hat{H} \hat{U}(s) = \hat{H}$ at $s = 0$: $\frac{i}{\hbar}[\hat{G}, \hat{H}] = 0$. Then apply the corollary to Theorem 15.5. ∎

This is the quantum form of Noether's theorem.

- Invariance under spatial translations (no external potential) conserves momentum.
- Invariance under time translations (a time-independent $\hat{H}$) conserves energy, the generator of time evolution.
- Invariance under rotations (a central potential) conserves angular momentum, the generator of rotations (Chapter 17).
- Invariance under the discrete reflection $\hat{\Pi}\psi(x) = \psi(-x)$ (an even potential) conserves parity, and stationary states can be chosen even or odd.

15.5 The Variational Principle

Few Hamiltonians can be diagonalized exactly. The two methods developed in the rest of this chapter are the workhorses for those that cannot.

Theorem 15.9 (Variational principle). *Let E_0 be the ground-state energy of $\hat{H}$. For every normalized state $|\psi\rangle$,*

$$\langle \psi | \hat{H} | \psi \rangle \geq E_0,$$

with equality if and only if $|\psi\rangle$ is a ground state.

Proof. Expand $|\psi\rangle = \sum_n c_n |n\rangle$ in the eigenbasis of $\hat{H}$, with $\sum |c_n|^2 = 1$. Then $\langle \psi | \hat{H} | \psi \rangle = \sum_n |c_n|^2 E_n \geq \sum_n |c_n|^2 E_0 = E_0$, with equality iff $c_n = 0$ whenever $E_n > E_0$. ∎

This is the Rayleigh quotient bound of Theorem 4.25 in infinite dimensions. To use it, choose a family of *trial functions* ψ_λ depending on parameters, compute $E(\lambda) = \langle \psi_\lambda | \hat{H} | \psi_\lambda \rangle$, and minimize. The minimum is an upper bound on E_0, often a remarkably good one.

Example 15.10 (Gaussian trial for the delta well). Estimate the ground state of $V = -\alpha\delta(x)$ with the trial function $\psi = (2b/\pi)^{1/4} e^{-bx^2}$.

Solution. By Example 13.8, $\langle T \rangle = \frac{\hbar^2 b}{2m}$. Also $\langle V \rangle = -\alpha |\psi(0)|^2 =$

$-\alpha\sqrt{2b/\pi}$. Minimizing $E(b) = \frac{\hbar^2 b}{2m} - \alpha\sqrt{\frac{2b}{\pi}}$ gives $\sqrt{b} = \frac{m\alpha}{\hbar^2}\sqrt{\frac{2}{\pi}}$, so $E_{\min} = -\frac{m\alpha^2}{\pi\hbar^2}$. The exact answer is $-\frac{m\alpha^2}{2\hbar^2}$ (Theorem 14.12). The bound is correct in sign and form, and its value is $2/\pi \approx 64\%$ of the true binding energy, which is respectable for a trial function with the wrong shape: it has no cusp at the origin.

Example 15.11 (The quartic oscillator). For $\hat{H} = \frac{\hat{p}^2}{2m} + \lambda\hat{x}^4$ and the same Gaussian trial, $\langle x^4\rangle = 3\langle x^2\rangle^2 = \frac{3}{16b^2}$ (Theorem 5.31), so $E(b) = \frac{\hbar^2 b}{2m} + \frac{3\lambda}{16b^2}$. Minimizing gives $b^3 = \frac{3m\lambda}{4\hbar^2}$ and

$$E_{\min} = \frac{3}{4}\left(\frac{3}{4}\right)^{1/3}\left(\frac{\hbar^4\lambda}{m^2}\right)^{1/3} \approx 0.6814\left(\frac{\hbar^4\lambda}{m^2}\right)^{1/3}.$$

The exact value, found numerically, is $0.6680\left(\hbar^4\lambda/m^2\right)^{1/3}$, so the Gaussian bound is only 2% high.

15.6 Time-Independent Perturbation Theory

Suppose $\hat{H} = \hat{H}^0 + \lambda\hat{H}'$, where $\hat{H}^0$ is a Hamiltonian we can solve exactly, with orthonormal eigenstates $|n^0\rangle$ and nondegenerate energies E_n^0, and $\lambda\hat{H}'$ is a small *perturbation*. Expand the exact eigenvalues and eigenvectors in powers of λ:

$$E_n = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \cdots, \qquad |n\rangle = |n^0\rangle + \lambda|n^1\rangle + \cdots.$$

Substitute into $\hat{H}|n\rangle = E_n|n\rangle$ and collect powers of λ. The λ^0 terms reproduce the unperturbed problem. The λ^1 terms give

$$(\hat{H}^0 - E_n^0)|n^1\rangle = (E_n^1 - \hat{H}')|n^0\rangle. \tag{15.2}$$

Theorem 15.12 (First and second order). *For a nondegenerate level,*

$$E_n^1 = \langle n^0|\hat{H}'|n^0\rangle, \qquad |n^1\rangle = \sum_{m\neq n}\frac{\langle m^0|\hat{H}'|n^0\rangle}{E_n^0 - E_m^0}|m^0\rangle,$$

$$E_n^2 = \sum_{m\neq n}\frac{|\langle m^0|\hat{H}'|n^0\rangle|^2}{E_n^0 - E_m^0}.$$

Proof. Take the inner product of (15.2) with $\langle n^0|$. The left side vanishes, because $\langle n^0|(\hat{H}^0 - E_n^0) = 0$, which gives E_n^1. Taking the inner product with $\langle m^0|$ for $m \neq n$

gives $(E_m^0 - E_n^0)\langle m^0|n^1\rangle = -\langle m^0|\hat{H}'|n^0\rangle$, which determines the components of $|n^1\rangle$. The component along $|n^0\rangle$ can be set to zero by normalization. The λ^2 equation, projected onto $\langle n^0|$, gives $E_n^2 = \langle n^0|\hat{H}'|n^1\rangle$. Substitute $|n^1\rangle$. ■

The first-order shift is simply the expectation value of the perturbation in the unperturbed state. The second-order correction to the ground-state energy is always negative, since every denominator $E_0^0 - E_m^0$ is negative. Levels "repel": each pair of states coupled by $\hat{H}'$ pushes apart. The expansion is useful when the matrix elements are small compared with the level spacings.

Example 15.13 (Charged oscillator in an electric field). Add $\hat{H}' = -qE\hat{x}$ to the oscillator. Find the energy shifts to second order and compare with the exact result.

Solution. $E_n^1 = -qE\langle n|\hat{x}|n\rangle = 0$ by (14.3). At second order only $m = n \pm 1$ contribute:

$$E_n^2 = q^2E^2\left(\frac{\hbar}{2m\omega}\right)\left[\frac{n}{\hbar\omega} + \frac{n+1}{-\hbar\omega}\right] = -\frac{q^2E^2}{2m\omega^2}.$$

Exactly, completing the square, $\frac{1}{2}m\omega^2x^2 - qEx = \frac{1}{2}m\omega^2\left(x - \frac{qE}{m\omega^2}\right)^2 - \frac{q^2E^2}{2m\omega^2}$: a shifted oscillator with every level lowered by $\frac{q^2E^2}{2m\omega^2}$. Second-order perturbation theory is exact here, and all higher orders vanish.

Example 15.14 (Anharmonic oscillator). Add $\hat{H}' = \lambda\hat{x}^4$ to the oscillator. By Example 14.8,

$$E_n \approx \left(n + \frac{1}{2}\right)\hbar\omega + \lambda\left(\frac{\hbar}{2m\omega}\right)^2(6n^2 + 6n + 3).$$

The shift grows like n^2, so the levels are no longer equally spaced, and perturbation theory fails for large n. The full series in λ actually diverges for every $\lambda \neq 0$, a famous result of Bender and Wu. Its first terms are nevertheless an excellent *asymptotic* approximation for small λ.

15.6.1 DEGENERATE LEVELS

If E_n^0 is degenerate, the denominators in Theorem 15.12 vanish. The cure is to choose the unperturbed basis within the degenerate subspace wisely. The first-order energies are the eigenvalues of the matrix $W_{ij} = \langle i^0|\hat{H}'|j^0\rangle$ on that subspace, and the "good" zeroth-order states are its eigenvectors. In practice, find

an operator that commutes with both $\hat{H}^0$ and $\hat{H}'$ and use its eigenstates; then W is already diagonal. This is how the fine structure of hydrogen and the Zeeman effect are computed, and it is why symmetries are so useful in perturbation theory.

Example 15.15 (Two degenerate states). Suppose $|1^0\rangle$, $|2^0\rangle$ share energy E^0 and $\hat{H}'$ has matrix $\begin{pmatrix} 0 & \Delta \\ \Delta & 0 \end{pmatrix}$ on their span. The eigenvalues of W are $\pm\Delta$, so the level splits into $E^0 \pm \Delta$, with good states $\frac{1}{\sqrt{2}}(|1^0\rangle \pm |2^0\rangle)$. This is the mechanism of bonding and antibonding orbitals in the hydrogen molecular ion. Electron tunneling between two equivalent sites splits a degenerate level into a symmetric and an antisymmetric combination (Chapter 18).

15.7 Worked Problems

The problems below apply the postulates to sequences of measurements, use the Heisenberg picture to derive a limit on repeated position measurements, prove Bloch's theorem from a symmetry, and test the variational method and perturbation theory, nondegenerate and degenerate, against exact or numerical results.

Problem 15.1 Measurement disturbs

A spin-$\frac{1}{2}$ particle is prepared in $|\uparrow\rangle$, the $+$ eigenstate of σ_z. It passes through three measurements: σ_z, then $\sigma_\theta = \cos\theta\,\sigma_z + \sin\theta\,\sigma_x$, then σ_z again. Find the probability that the final result is $+1$, whatever the middle result was. Compare with the probability when the middle measurement is omitted.

Solution. $\sigma_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ has eigenvalues ± 1, with eigenvectors $|+\theta\rangle = \langle \cos\frac{\theta}{2}, \sin\frac{\theta}{2}\rangle$ and $|-\theta\rangle = \langle -\sin\frac{\theta}{2}, \cos\frac{\theta}{2}\rangle$, as the half-angle formulas confirm. The first measurement gives $+1$ with certainty and leaves the state unchanged. By Postulate III (Definition 15.1), the middle measurement gives ± 1 with probabilities $\cos^2\frac{\theta}{2}$ and $\sin^2\frac{\theta}{2}$, and collapses the state to $|\pm\theta\rangle$. From $|+\theta\rangle$ the final σ_z gives $+1$ with probability $\cos^2\frac{\theta}{2}$, and from $|-\theta\rangle$ with probability $\sin^2\frac{\theta}{2}$. So

$$P(+) = \cos^4\frac{\theta}{2} + \sin^4\frac{\theta}{2} = 1 - \frac{1}{2}\sin^2\theta.$$

Without the middle measurement, $P(+) = 1$. For $\theta = \frac{\pi}{2}$ the intermediate σ_x measurement reduces it to $\frac{1}{2}$: it erases all memory of the initial z-state. Only for $\theta = 0$ or π, when $\sigma_\theta = \pm\sigma_z$ commutes with σ_z, is there no disturbance.

Notice that the result holds even if nobody reads the middle result. The probabilities were added, not the amplitudes, because the two middle outcomes are in principle distinguishable. Adding amplitudes instead, $\langle\uparrow|+\theta\rangle\langle+\theta|\uparrow\rangle + \langle\uparrow|-\theta\rangle\langle-\theta|\uparrow\rangle = \langle\uparrow|\uparrow\rangle = 1$ by completeness, recovers the undisturbed result. This is the double-slit experiment of Chapter 12 in matrix form.

Problem 15.2 The standard quantum limit

A particle of mass m moves in a uniform force field F, so $\hat{H} = \frac{\hat{p}^2}{2m} - F\hat{x}$. Find $\hat{x}(t)$ in the Heisenberg picture, compute $[\hat{x}(t), \hat{x}(0)]$, and show that $\sigma_{x(0)}\sigma_{x(t)} \geq \frac{\hbar t}{2m}$ in every state.

Solution. By (15.1), $\dot{\hat{x}} = \frac{i}{\hbar}[\hat{H}, \hat{x}] = \frac{\hat{p}}{m}$ and $\dot{\hat{p}} = \frac{i}{\hbar}[\hat{H}, \hat{p}] = F$. These are the classical equations, and they integrate the same way:

$$\hat{p}(t) = \hat{p} + Ft, \qquad \hat{x}(t) = \hat{x} + \frac{\hat{p}}{m}t + \frac{F}{2m}t^2.$$

Here $\hat{x}$ and $\hat{p}$ are the operators at $t = 0$, and F multiplies the identity. Then

$$[\hat{x}(t), \hat{x}(0)] = \left[\hat{x} + \frac{t}{m}\hat{p},\ \hat{x}\right] = \frac{t}{m}[\hat{p}, \hat{x}] = -\frac{i\hbar t}{m}.$$

By the Robertson relation (Theorem 15.3), applied to the two Hermitian operators $\hat{x}(0)$ and $\hat{x}(t)$,

$$\sigma_{x(0)}\,\sigma_{x(t)} \geq \frac{1}{2}\big|\langle[\hat{x}(t), \hat{x}(0)]\rangle\big| = \frac{\hbar t}{2m}.$$

The force drops out completely. A particle localized precisely at time 0 must be poorly localized at time t. The trade-off is minimized when $\sigma_{x(0)} = \sigma_{x(t)} = \sqrt{\hbar t/2m}$. This sets the *standard quantum limit* for monitoring the position of a free mass. It mattered in the design of the LIGO gravitational-wave detectors, whose 40 kg mirrors must be located to about 10^{-19} m. Beating the limit requires measuring something other than position, such as the squeezed light used since 2019.

Problem 15.3 Bloch's theorem from a symmetry

Suppose $V(x + a) = V(x)$ for all x. Using the translation operator $\hat{T}(a) = e^{-ia\hat{p}/\hbar}$, which acts by $(\hat{T}(a)\psi)(x) = \psi(x - a)$, prove that the energy eigenfunctions can be chosen to satisfy

$$\psi(x + a) = e^{iKa}\psi(x) \qquad \text{for some real } K \in (-\pi/a, \pi/a].$$

Solution. First, $\hat{T}(a)$ commutes with $\hat{H}$. The kinetic term commutes with $\hat{T}(a)$ because $\hat{p}$ commutes with any function of $\hat{p}$. For the potential, $(\hat{T}\,V\psi)(x) = V(x-a)\psi(x-a) = V(x)\psi(x-a) = (V\,\hat{T}\psi)(x)$, using periodicity. So $[\hat{T}(a), \hat{H}] = 0$.

By Theorem 9.30, extended to the present setting, $\hat{H}$ and $\hat{T}(a)$ have a common set of eigenfunctions. Since $\hat{T}(a)$ is unitary (Theorem 9.31, as $\hat{p}$ is Hermitian), its eigenvalues have modulus 1, so they can be written e^{-iKa} with K real. Then $\psi(x-a) = e^{-iKa}\psi(x)$ for all x, which, with x replaced by $x+a$, is $\psi(x+a) = e^{iKa}\psi(x)$. Since e^{iKa} is unchanged by $K \to K + 2\pi/a$, K can be taken in an interval of length $2\pi/a$.

Equivalently, $\psi(x) = e^{iKx}u(x)$ with u periodic: a plane wave modulated by a function with the period of the lattice. The *crystal momentum* $\hbar K$ labels the states, and is conserved because the Hamiltonian is invariant under lattice translations (Theorem 15.8), although $\hat{p}$ itself is not conserved. For each K the energies form a discrete set $E_1(K) \le E_2(K) \le \cdots$, and as K varies these trace out the *energy bands* of a solid, separated by gaps. Whether the highest occupied band is full or partly full decides whether a material is an insulator or a metal.

Problem 15.4 Rayleigh–Ritz for the infinite well

For the infinite well on $(0, a)$, use the trial function $\phi_1 = x(a-x)$ to bound the ground-state energy. Then improve the bound with the two-function trial $\psi = c_1\phi_1 + c_2\phi_2$, $\phi_2 = x^2(a-x)^2$, by minimizing over c_1, c_2.

Solution. Work in units $\hbar = m = a = 1$, so energies are in units of $\frac{\hbar^2}{ma^2}$ and the exact ground state is $\frac{\pi^2}{2} \approx 4.934802$. For real trial functions vanishing at the walls, $\langle \hat{H} \rangle = \frac{1}{2}\int \psi'^2 dx$ after integration by parts.

One function. $\frac{1}{2}\int_0^1 (1-2x)^2 dx = \frac{1}{6}$ and $\int_0^1 x^2(1-x)^2 dx = \frac{1}{30}$, so $E \le \frac{1/6}{1/30} = 5$, an error of 1.3%.

Two functions. The ratio $\langle\psi|\hat{H}|\psi\rangle / \langle\psi|\psi\rangle$ is stationary over (c_1, c_2) exactly when $H\mathbf{c} = ES\mathbf{c}$, where $H_{ij} = \frac{1}{2}\int \phi_i'\phi_j' dx$ and $S_{ij} = \int \phi_i\phi_j\,dx$ (the generalized eigenvalue problem; compare Theorem 4.25). Computing,

$$H = \begin{pmatrix} \frac{1}{6} & \frac{1}{30} \\ \frac{1}{30} & \frac{1}{105} \end{pmatrix}, \qquad S = \begin{pmatrix} \frac{1}{30} & \frac{1}{140} \\ \frac{1}{140} & \frac{1}{630} \end{pmatrix}.$$

The condition $\det(H - ES) = 0$ reduces to $E^2 - 56E + 252 = 0$, so

$$E = 28 - 2\sqrt{133} \approx 4.934875.$$

The error is now 7×10^{-5}, or 0.0015%, nearly a thousandfold improvement for

one added function. The other root, $28+2\sqrt{133}\approx 51.07$, is an upper bound for the next state with the same symmetry about the center, the $n=3$ level $\frac{9\pi^2}{2}\approx 44.41$. Rayleigh–Ritz bounds every level, not just the ground state, a result due to Hylleraas, Undheim, and MacDonald. With polynomial or Gaussian basis sets of hundreds of functions, this is how the electronic structure of molecules is computed in practice.

Problem 15.5 A bound for an excited state

For $\hat H=-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+x^4$ (units with $\hbar=m=1$), the Gaussian trial of Example 15.11 bounds the ground state. Use the trial function $\psi=xe^{-bx^2}$ to bound the first excited energy, and compare with the numerical value $E_1=2.39364$.

Solution. The variational principle bounds only the ground state in general. But $\hat H$ is even, so its eigenstates are even or odd, and the ground state is even. Every odd trial function is orthogonal to it, so $\langle\hat H\rangle$ in an odd state is at least the lowest odd energy, E_1. (Expand the trial function in eigenstates: only odd ones appear.) With $\psi=xe^{-bx^2}$ and the Gaussian moments of Theorem 5.31,

$$\frac{\int \frac{1}{2}\psi'^2dx}{\int\psi^2dx}=\frac{3b}{2},\qquad \frac{\int x^4\psi^2dx}{\int\psi^2dx}=\frac{15}{16b^2},\qquad \langle\hat H\rangle=\frac{3b}{2}+\frac{15}{16b^2}=\frac{3(8b^3+5)}{16b^2}.$$

Minimizing, $\frac{3}{2}-\frac{15}{8b^3}=0$ gives $b^3=\frac{5}{4}$, $b=\frac{10^{1/3}}{2}\approx 1.077$, and $E_1\le 2.4237$. The numerical value, from a finite-difference diagonalization of $\hat H$ on $[-6,6]$ with 4000 points, is 2.39364, so the bound is 1.3% high. The Gaussian bound on the ground state, 0.6814, is 2.0% above the numerical 0.66799. Symmetry has turned a ground-state method into a method for the lowest state of each symmetry class, a technique used throughout atomic and molecular physics.

Problem 15.6 A perturbation that can be solved exactly

Add $\hat H'=\epsilon\hat x^2$ to the oscillator. Find the exact energies, expand them to second order in ϵ, and verify that nondegenerate perturbation theory gives the same first- and second-order terms.

Solution. The perturbed potential $\frac{1}{2}m\omega^2x^2+\epsilon x^2=\frac{1}{2}m\omega'^2x^2$ is an oscillator with $\omega'=\omega\sqrt{1+\frac{2\epsilon}{m\omega^2}}$. So

$$E_n=\Big(n+\frac{1}{2}\Big)\hbar\omega\sqrt{1+\frac{2\epsilon}{m\omega^2}}=\Big(n+\frac{1}{2}\Big)\hbar\omega\Big(1+\frac{\epsilon}{m\omega^2}-\frac{\epsilon^2}{2m^2\omega^4}+\cdots\Big).$$

First order. $E_n^{(1)}=\epsilon\langle n|\hat x^2|n\rangle=\epsilon\frac{\hbar}{2m\omega}(2n+1)$ by Example 14.7, which is $\big(n+\frac{1}{2}\big)\hbar\omega\cdot\frac{\epsilon}{m\omega^2}$, in agreement.

Second order. With $x_0^2 = \frac{\hbar}{m\omega}$, $\hat{x}^2 = \frac{x_0^2}{2}(\hat{a}+\hat{a}^\dagger)^2$ connects $|n\rangle$ only to $|n\pm 2\rangle$ off the diagonal: $\langle n+2|\hat{x}^2|n\rangle = \frac{x_0^2}{2}\sqrt{(n+1)(n+2)}$ and $\langle n-2|\hat{x}^2|n\rangle = \frac{x_0^2}{2}\sqrt{n(n-1)}$. The energy denominators are $\mp 2\hbar\omega$. By Theorem 15.12,

$$E_n^{(2)} = \epsilon^2 \frac{x_0^4}{4}\left[\frac{(n+1)(n+2)}{-2\hbar\omega} + \frac{n(n-1)}{2\hbar\omega}\right] = \epsilon^2 \frac{\hbar^2}{4m^2\omega^2}\cdot\frac{-(4n+2)}{2\hbar\omega}$$
$$= -\left(n+\frac{1}{2}\right)\hbar\omega\frac{\epsilon^2}{2m^2\omega^4}.$$

This matches the exact expansion. The series converges for $|\epsilon| < \frac{m\omega^2}{2}$, the radius of convergence of the square root, whose branch point at $\epsilon = -\frac{m\omega^2}{2}$ is where the perturbed potential stops being confining. Perturbation series often diverge, and their radius of convergence is set by such physical thresholds, possibly at complex values of the coupling.

Problem 15.7 Degenerate perturbation theory in a square box

A particle in the square box $0 < x, y < a$ has eigenfunctions $\psi_{n_x n_y} = \frac{2}{a}\sin\frac{n_x\pi x}{a}\sin\frac{n_y\pi y}{a}$ and energies $\propto n_x^2 + n_y^2$, so the first excited level, $(1,2)$ and $(2,1)$, is doubly degenerate. Find its splitting under $\hat{H}' = \epsilon\, xy$, and the correct zeroth-order states.

Solution. The states factor, so the matrix elements of xy factor into one-dimensional integrals. Let $x_{jk} = \frac{2}{a}\int_0^a x\sin\frac{j\pi x}{a}\sin\frac{k\pi x}{a}\,dx$. Then $x_{11} = x_{22} = \frac{a}{2}$ (each density is symmetric about the center), and a short integration gives $x_{12} = -\frac{16a}{9\pi^2}$. In the basis $|12\rangle, |21\rangle$,

$$W = \epsilon\begin{pmatrix} x_{11}x_{22} & x_{12}x_{21} \\ x_{21}x_{12} & x_{22}x_{11}\end{pmatrix} = \epsilon\begin{pmatrix} \frac{a^2}{4} & \gamma \\ \gamma & \frac{a^2}{4}\end{pmatrix}, \qquad \gamma = x_{12}^2 = \frac{256a^2}{81\pi^4}.$$

By degenerate perturbation theory (the two-state example earlier in this chapter), the first-order shifts are the eigenvalues of W:

$$E_\pm^{(1)} = \epsilon a^2\left(\frac{1}{4} \pm \frac{256}{81\pi^4}\right) \approx \epsilon a^2(0.25 \pm 0.0324),$$

with eigenvectors $\frac{1}{\sqrt{2}}(|12\rangle \pm |21\rangle)$. The perturbation splits the level by $2\gamma\epsilon$. The correct zeroth-order states are the combinations symmetric and antisymmetric under $x \leftrightarrow y$, the symmetry that xy preserves. The symmetric one has its density concentrated near the corners $(0,0)$ and (a,a), where xy is smallest and largest; with $\epsilon > 0$ it lies higher. Naive nondegenerate theory would give $\frac{\epsilon a^2}{4}$ for both states and miss the splitting entirely, and its second-order term would divide by zero.

EXERCISES

A. COMPUTATION

15.1 Let $\hat{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ and $|\psi\rangle = \langle 1, 0\rangle$. Find the possible results of measuring A, their probabilities, $\langle A\rangle$, and σ_A.

15.2 A three-level system has $\hat{H} = \operatorname{diag}(E_1, E_2, E_3)$ and starts in $|\psi(0)\rangle = \frac{1}{\sqrt{3}}\langle 1, 1, 1\rangle$. Find $|\psi(t)\rangle$ and the probability of finding it in $|\psi(0)\rangle$ at time t.

15.3 With $\hat{B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, find the eigenvalues and projectors. If $|\psi\rangle = \frac{1}{\sqrt{6}}\langle 1, 2, 1\rangle$, find the probability of each result and the post-measurement state.

15.4 Compute $[\hat{x}, \hat{p}^2]$, $[\hat{x}^2, \hat{p}]$, and $[\hat{x}, f(\hat{p})]$ for analytic f.

15.5 Compute $[\hat{H}, \hat{x}]$ and $[\hat{H}, \hat{p}]$ for $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$ and verify Ehrenfest's theorem.

15.6 Verify the Robertson relation for σ_x and σ_y in the spin state $|\uparrow\rangle$, using $\hat{S}_i = \frac{\hbar}{2}\sigma_i$.

15.7 For which spin-$\frac{1}{2}$ states does $\sigma_{S_x}\sigma_{S_z}$ attain its minimum? Compute $\sigma_{S_x}\sigma_{S_z}$ for $|\psi\rangle = \cos\frac{\theta}{2}|\uparrow\rangle + \sin\frac{\theta}{2}|\downarrow\rangle$.

15.8 Use $\hat{T}(a) = e^{-ia\hat{p}/\hbar}$ to show $\hat{T}(a)^\dagger \hat{x} \hat{T}(a) = \hat{x} + a$.

15.9 Compute $\langle x\rangle(t)$ and $\langle p\rangle(t)$ in the Heisenberg picture for a free particle, and show $\sigma_x^2(t)$ is quadratic in t.

15.10 Estimate the ground-state energy of the harmonic oscillator using the trial function $\psi = A/(x^2 + b^2)$. Compare with $\frac{1}{2}\hbar\omega$.

15.11 Use the trial function $\psi = Ae^{-b|x|}$ to bound the oscillator's ground-state energy.

15.12 Use a Gaussian trial function to bound the ground-state energy of the linear potential $V = F|x|$.

15.13 Apply first-order perturbation theory to the infinite well with $\hat{H}' = V_0$ for $0 < x < a/2$ and 0 otherwise. Find E_n^1.

15.14 Apply first-order perturbation theory to the infinite well with $\hat{H}' = \epsilon\,\delta(x - a/2)$. Which levels are unaffected, and why?

15.15 For the oscillator with $\hat{H}' = \epsilon \hat{x}^3$, show that $E_n^1 = 0$ and compute E_0^2.

15.16 Find the first-order energy shifts of the two lowest levels of the infinite well under $\hat{H}' = \epsilon x$, and the second-order shift of the ground state (leave as a sum).

15.17 A two-level system has $\hat{H} = \begin{pmatrix} E_1 & V \\ V & E_2 \end{pmatrix}$ with $E_1 < E_2$. Find the exact eigenvalues and compare with second-order perturbation theory for $|V| \ll E_2 - E_1$.

B. THEORY AND PROOF

15.18 Prove that the eigenvalues of $\hat{P}_a$ are 0 and 1, and that $\sum_a P(a) = 1$ in Postulate III.

15.19 Prove that the commutator of two Hermitian operators is anti-Hermitian, and that $i[\hat{A}, \hat{B}]$ is Hermitian.

15.20 Prove the commutator identities $[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}]$ and $[\hat{x}, \hat{p}^n] = i\hbar n \hat{p}^{n-1}$.

15.21 Prove that $\sigma_A = 0$ in a state $|\psi\rangle$ if and only if $|\psi\rangle$ is an eigenstate of $\hat{A}$.

15.22 (Schrödinger–Robertson.) Strengthen Theorem 15.3 to

$$\sigma_A^2 \sigma_B^2 \geq \left(\tfrac{1}{2}\langle\{\hat{A}, \hat{B}\}\rangle - \langle A\rangle\langle B\rangle\right)^2 + \left(\tfrac{1}{2i}\langle[\hat{A}, \hat{B}]\rangle\right)^2,$$

where $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$.

15.23 Prove that the Heisenberg-picture operators satisfy the same commutation relations at equal times as the Schrödinger-picture ones: $[\hat{x}_H(t), \hat{p}_H(t)] = i\hbar$.

15.24 Prove that the parity operator $\hat{\Pi}$ is Hermitian and unitary, with eigenvalues ± 1, and that $[\hat{\Pi}, \hat{H}] = 0$ for even V.

15.25 Prove the *Hellmann–Feynman theorem*: if $\hat{H}(\lambda)|n(\lambda)\rangle = E_n(\lambda)|n(\lambda)\rangle$ with normalized states, then $\mathrm{d}E_n/\mathrm{d}\lambda = \langle n|\,\partial\hat{H}/\partial\lambda\,|n\rangle$. Use it with $\lambda = \omega$ to recover $\langle V\rangle = \tfrac{1}{2}E_n$ for the oscillator.

15.26 Show that the variational estimate with a trial function containing the exact ground state as a special case reproduces E_0 exactly.

C. CHALLENGE

15.27 (Second excited state by variation.) Show that if a trial function is orthogonal to the exact ground state, $\langle\psi|\hat{H}|\psi\rangle \geq E_1$. Use an odd trial function xe^{-bx^2} to bound the first excited state of the quartic oscillator.

15.28 (Thomas–Reiche–Kuhn sum rule.) Prove $\sum_m (E_m - E_n)|\langle m|\hat{x}|n\rangle|^2 = \frac{\hbar^2}{2m}$ by evaluating $\langle n|[\hat{x},[\hat{H},\hat{x}]]|n\rangle$ in two ways. Check it for the oscillator.

15.29 Use the Gaussian trial function to estimate the ground state of hydrogen in three dimensions, $\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r}$, and compare with -13.6 eV. (The answer is $-\frac{8}{3\pi} \times 13.6$ eV.)

15.30 The quantum Zeno effect: a system starts in $|\psi_0\rangle$ and is measured N times at intervals t/N. Show that the probability it is still in $|\psi_0\rangle$ after time t is $\left(1 - \sigma_H^2 t^2/\hbar^2 N^2 + \cdots\right)^N \to 1$ as $N \to \infty$.

CHAPTER 16

Quantum Mechanics in Three Dimensions: The Hydrogen Atom

LEARNING OBJECTIVES

- Solve the three-dimensional box by separation of variables, and explain degeneracy.
- Separate the Schrödinger equation for a central potential into angular and radial equations, and interpret the centrifugal term.
- Solve the angular equation: quantize m by single-valuedness and ℓ by regularity, and construct the spherical harmonics.
- Solve the radial equation for hydrogen by power series, and derive $E_n = -13.6\ \text{eV}/n^2$ from the termination condition.
- Write and interpret the hydrogen wavefunctions $\psi_{n\ell m}$, compute expectation values, and count degeneracies.
- Relate the spectrum to observed lines, including selection rules and the reduced-mass correction.
- Describe qualitatively the corrections beyond the Schrödinger model and the structure of many-electron atoms.

The hydrogen atom is the problem quantum mechanics was invented to solve, and its exact solution is the theory's first great success. Every tool developed in this book contributes: spherical coordinates (Chapter 1), the spherical Laplacian (Chapter 7), separation of variables and series solutions (Chapter 10), orthogonal functions (Chapter 9), and the formalism of eigenvalues and expectation values (Chapter 15). The result explains the Balmer lines, the Bohr energies, and the shapes of atomic orbitals that underlie all of chemistry.

16.1 Three Dimensions and Degeneracy

In three dimensions the time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{r})\psi = E\psi, \qquad \int |\psi|^2 dV = 1. \tag{16.1}$$

The momentum operator is $\hat{\mathbf{p}} = -i\hbar\nabla$, with components satisfying $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$.

Example 16.1 (The cubical box). Let $V = 0$ inside the cube $[0,a]^3$ and $V = \infty$ outside. Separating $\psi = X(x)Y(y)Z(z)$ gives three copies of the infinite well, so

$$\psi_{n_x n_y n_z} = \left(\frac{2}{a}\right)^{3/2} \sin\frac{n_x\pi x}{a}\sin\frac{n_y\pi y}{a}\sin\frac{n_z\pi z}{a}, \qquad E = \frac{\pi^2\hbar^2}{2ma^2}\left(n_x^2+n_y^2+n_z^2\right).$$

The ground state $(1,1,1)$ is unique, but the first excited level $E = 6E_0'$, with $E_0' = \frac{\pi^2\hbar^2}{2ma^2}$, is shared by $(2,1,1)$, $(1,2,1)$, and $(1,1,2)$. It is *threefold degenerate*. The degeneracy comes from the symmetry of the cube: rotations that permute the axes map one of these states into another. Squashing the box to unequal sides lifts the degeneracy. Some degeneracies are "accidental"; for example, $(3,3,3)$ and $(5,1,1)$ both have $n_x^2+n_y^2+n_z^2 = 27$.

As a rule, degeneracy signals symmetry. The rotational symmetry of a central potential produces the $(2\ell+1)$-fold degeneracy in m derived below. Hydrogen's extra degeneracy in ℓ reflects a hidden symmetry of the inverse-square force, the one behind conservation of the Runge–Lenz vector (Theorem 2.28).

16.2 Central Potentials and Separation of Variables

For a *central* potential $V = V(r)$, use spherical coordinates and the Laplacian of Corollary 7.29. Write $\nabla^2 = \frac{1}{r^2}\,\partial/\partial r\left(r^2\,\partial/\partial r\right) + \frac{1}{r^2}\Lambda$, where Λ is the angular operator. Seek $\psi = R(r)Y(\theta,\phi)$. Substituting into (16.1), multiplying by $\frac{2mr^2}{\hbar^2 RY}$, and rearranging,

$$\underbrace{\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) - \frac{2mr^2}{\hbar^2}\big(V(r)-E\big)}_{\text{depends only on } r} = \underbrace{-\frac{1}{Y}\Lambda Y}_{\text{depends only on } \theta,\phi} = \ell(\ell+1),$$

where the separation constant is written $\ell(\ell+1)$ for reasons that will become clear. This gives two equations.

The angular equation, $\Lambda Y = -\ell(\ell+1)Y$, does not involve the potential at all. Its solutions, the spherical harmonics, are the same for every central potential.

The radial equation, simplified by the substitution $u(r) = rR(r)$ (as in Chapter 7), is

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u}{\mathrm{d}r^2} + \left[V(r) + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\right]u = Eu, \qquad \int_0^\infty |u|^2 dr = 1. \tag{16.2}$$

This is a one-dimensional Schrödinger equation on $r > 0$, with boundary condition $u(0) = 0$ and an *effective potential* $V_{\text{eff}} = V + \frac{\hbar^2\ell(\ell+1)}{2mr^2}$. The extra term is the *centrifugal barrier*. Classically, a particle with angular momentum L has radial energy $\frac{p_r^2}{2m} + V + \frac{L^2}{2mr^2}$. The match identifies $L^2 = \ell(\ell+1)\hbar^2$, which Chapter 17 confirms.

16.3 The Angular Equation and Spherical Harmonics

Separate once more: $Y = \Theta(\theta)\Phi(\phi)$. The angular equation becomes

$$\frac{\sin\theta}{\Theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\right) + \ell(\ell+1)\sin^2\theta = -\frac{1}{\Phi}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\phi^2} = m^2.$$

The ϕ equation $\Phi'' = -m^2\Phi$ has solutions $e^{im\phi}$. A wavefunction must be single-valued, so $\Phi(\phi + 2\pi) = \Phi(\phi)$, which forces $m \in \mathbb{Z}$. Here m is the *magnetic quantum number*, not the mass. The context always makes clear which is meant.

The θ equation, with $x = \cos\theta$, becomes the *associated Legendre equation*

$$(1-x^2)\Theta'' - 2x\Theta' + \left[\ell(\ell+1) - \frac{m^2}{1-x^2}\right]\Theta = 0. \tag{16.3}$$

For $m = 0$ this is Legendre's equation (10.2). As shown in Chapter 10, it has a solution finite at both poles $x = \pm 1$ only when ℓ is a nonnegative integer. For general m, the regular solutions are the *associated Legendre functions*

$$P_\ell^m(x) = (1-x^2)^{|m|/2}\frac{\mathrm{d}^{|m|}}{\mathrm{d}x^{|m|}}P_\ell(x),$$

which vanish identically unless $|m| \le \ell$, since P_ℓ is a polynomial of degree ℓ.

Theorem 16.2 (Spherical harmonics). *The regular, single-valued solutions of $\Lambda Y = -\ell(\ell+1)Y$ are, for $\ell = 0, 1, 2, \ldots$ and $m = -\ell, \ldots, \ell$,*

$$Y_\ell^m(\theta,\phi) = \epsilon \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-|m|)!}{(\ell+|m|)!}}\, e^{im\phi} P_\ell^m(\cos\theta),$$

with $\epsilon = (-1)^m$ for $m > 0$ and $\epsilon = 1$ for $m \le 0$. They are orthonormal on the sphere:

$$\int_0^{2\pi}\int_0^{\pi} Y_\ell^{m*}\, Y_{\ell'}^{m'} \sin\theta\, d\theta\, d\phi = \delta_{\ell\ell'}\delta_{mm'},$$

and complete: every square-integrable function on the sphere has an expansion $\sum_{\ell,m} c_{\ell m} Y_\ell^m$.

Orthogonality in m comes from $\int e^{i(m'-m)\phi}d\phi$. Orthogonality in ℓ follows from Sturm–Liouville theory, since Λ is Hermitian on the sphere, and from Theorem 9.25. For each ℓ there are $2\ell+1$ values of m, so every level of a central potential with angular momentum ℓ is at least $(2\ell+1)$-fold degenerate.

ℓ	m	$Y_\ell^m(\theta,\phi)$
0	0	$\frac{1}{\sqrt{4\pi}}$
1	0	$\sqrt{\frac{3}{4\pi}}\cos\theta$
1	±1	$\mp\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{\pm i\phi}$
2	0	$\sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1)$
2	±1	$\mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\, e^{\pm i\phi}$
2	±2	$\sqrt{\frac{15}{32\pi}}\sin^2\theta\, e^{\pm 2i\phi}$

Chemists use real combinations. Since $\sqrt{\frac{3}{4\pi}}\cos\theta = \sqrt{\frac{3}{4\pi}}\frac{z}{r}$, and $\frac{1}{\sqrt{2}}(Y_1^{-1} - Y_1^1) = \sqrt{\frac{3}{4\pi}}\frac{x}{r}$, and similarly for y, the three $\ell = 1$ functions are proportional to x/r, y/r, z/r. These are the dumbbell-shaped p_x, p_y, p_z orbitals, pointing along the axes. The five $\ell = 2$ functions are the d orbitals of transition-metal chemistry.

16.4 The Radial Equation for Hydrogen

For the electron in hydrogen, $V(r) = -\frac{e^2}{4\pi\varepsilon_0 r}$. (Strictly, m should be the reduced mass of the electron and proton, which we return to in Section 16.6.) For bound states $E < 0$. Introduce

$$\kappa = \frac{\sqrt{-2mE}}{\hbar}, \qquad \rho = \kappa r, \qquad \rho_0 = \frac{me^2}{2\pi\varepsilon_0\hbar^2\kappa} = \frac{2}{a\kappa},$$

where $a = \frac{4\pi\varepsilon_0\hbar^2}{me^2}$ is the Bohr radius. Then (16.2) becomes

$$\frac{\mathrm{d}^2u}{\mathrm{d}\rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{\ell(\ell+1)}{\rho^2}\right]u. \tag{16.4}$$

Asymptotics. As $\rho \to \infty$, $u'' \approx u$, so $u \sim e^{-\rho}$ for normalizability. As $\rho \to 0$, the centrifugal term dominates, $u'' \approx \frac{\ell(\ell+1)}{\rho^2}u$, with solutions $\rho^{\ell+1}$ and $\rho^{-\ell}$. Only the first is acceptable, since $u(0) = 0$ is required. So peel off both behaviors:

$$u(\rho) = \rho^{\ell+1}e^{-\rho}v(\rho).$$

Substituting into (16.4) gives

$$\rho v'' + 2(\ell + 1 - \rho)v' + \big[\rho_0 - 2(\ell+1)\big]v = 0. \tag{16.5}$$

Series. Try $v = \sum_j c_j\rho^j$. Collecting powers of ρ,

$$c_{j+1} = \frac{2(j+\ell+1) - \rho_0}{(j+1)(j+2\ell+2)}\,c_j.$$

For large j, $c_{j+1} \approx \frac{2}{j}c_j$, the behavior of the series for $e^{2\rho}$. A non-terminating series would make $u \sim \rho^{\ell+1}e^{\rho}$, which cannot be normalized. This is the same mechanism as for the oscillator (Theorem 10.12). So the series must terminate: for some integer $j_{\max} \geq 0$,

$$\rho_0 = 2(j_{\max} + \ell + 1) \equiv 2n,$$

where $n = j_{\max} + \ell + 1$ is the *principal quantum number*, and $n \geq \ell + 1$.

Theorem 16.3 (Energy levels of hydrogen). *The bound-state energies of hydro-*

gen are

$$E_n = -\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \frac{1}{n^2} = -\frac{\hbar^2}{2ma^2n^2} = -\frac{13.6\text{ eV}}{n^2}, \qquad n = 1, 2, 3, \ldots$$

For each n, the allowed values are $\ell = 0, 1, \ldots, n-1$ and $m = -\ell, \ldots, \ell$. The level E_n is n^2-fold degenerate, or $2n^2$-fold including electron spin.

Proof. From $\rho_0 = \frac{2}{a\kappa} = 2n$ we get $\kappa = \frac{1}{na}$, so $E = -\frac{\hbar^2\kappa^2}{2m} = -\frac{\hbar^2}{2ma^2n^2}$. Counting, $\sum_{\ell=0}^{n-1}(2\ell+1) = n^2$. ∎

These are exactly the Bohr energies of Theorem 12.11, now derived without Bohr's ad hoc postulates. The solution differs from the Bohr model in two important ways. First, for a given n there are several angular momenta, and $\ell = 0$ is allowed. The ground state has *zero* orbital angular momentum, so the electron does not orbit in any classical sense. Second, the electron has no definite path; it is described by a probability cloud.

Theorem 16.4 (Hydrogen wavefunctions). *The normalized stationary states are $\psi_{n\ell m}(r, \theta, \phi) = R_{n\ell}(r)\, Y_\ell^m(\theta, \phi)$, with*

$$R_{n\ell}(r) = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-\ell-1)!}{2n\,(n+\ell)!}}\; e^{-r/na}\left(\frac{2r}{na}\right)^\ell L_{n-\ell-1}^{2\ell+1}\left(\frac{2r}{na}\right),$$

where L_q^p are the associated Laguerre polynomials, which are the terminating solutions of (16.5).

The first few radial functions are

$$R_{10} = 2a^{-3/2}e^{-r/a},$$
$$R_{20} = \tfrac{1}{\sqrt{2}}a^{-3/2}\left(1 - \frac{r}{2a}\right)e^{-r/2a}, \qquad R_{21} = \tfrac{1}{2\sqrt{6}}a^{-3/2}\frac{r}{a}\,e^{-r/2a},$$
$$R_{30} = \tfrac{2}{3\sqrt{3}}a^{-3/2}\left(1 - \frac{2r}{3a} + \frac{2r^2}{27a^2}\right)e^{-r/3a}, \quad R_{31} = \tfrac{8}{27\sqrt{6}}a^{-3/2}\frac{r}{a}\left(1 - \frac{r}{6a}\right)e^{-r/3a},$$
$$R_{32} = \tfrac{4}{81\sqrt{30}}a^{-3/2}\frac{r^2}{a^2}e^{-r/3a}.$$

$R_{n\ell}$ has $n - \ell - 1$ radial nodes. In spectroscopic notation, the letters s, p, d, f denote $\ell = 0, 1, 2, 3$, so ψ_{100} is the $1s$ state and ψ_{21m} are the $2p$ states.

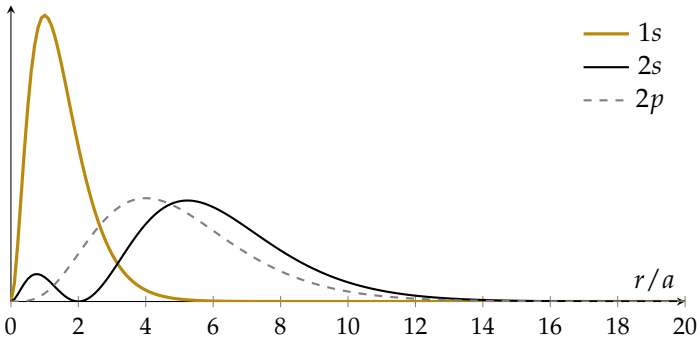


Figure 16.1: Radial probability densities $P(r) = r^2|R_{n\ell}|^2$, in units of $1/a$. The $1s$ density peaks at $r = a$, the Bohr radius. The $2s$ state has one radial node at $r = 2a$; the $2p$ state has none and peaks at $4a$.

16.5 Properties of the Hydrogen States

The probability of finding the electron between r and $r + dr$, at any angle, is $P(r)\,dr = r^2|R_{n\ell}(r)|^2 dr$, since the spherical harmonics are normalized. For the ground state, $P(r) = \frac{4r^2}{a^3}e^{-2r/a}$, which is largest at $r = a$. The Bohr radius survives as the *most probable* distance, although $\langle r\rangle = \frac{3}{2}a$ (Example 5.23).

Theorem 16.5 (Radial expectation values). *In the state $\psi_{n\ell m}$,*

$$\langle r\rangle = \frac{a}{2}\big[3n^2 - \ell(\ell+1)\big], \qquad \left\langle \frac{1}{r}\right\rangle = \frac{1}{n^2 a}, \qquad \left\langle \frac{1}{r^2}\right\rangle = \frac{1}{(\ell+\tfrac{1}{2})n^3 a^2}.$$

The value of $\langle 1/r\rangle$ follows from the virial theorem: for a $1/r$ potential, $\langle V\rangle = 2E_n$, so $-\frac{e^2}{4\pi\varepsilon_0}\left\langle \frac{1}{r}\right\rangle = -\frac{\hbar^2}{ma^2n^2}$. The others follow from the explicit wavefunctions, or more elegantly from the Hellmann–Feynman theorem (Exercise 15.25).

Example 16.6 (Probability of being inside the nucleus). The proton has radius $b \approx 0.84 \times 10^{-15}$ m $\approx 1.6 \times 10^{-5} a$. For $b \ll a$, the ground-state probability is $\int_0^b P(r)\,dr \approx \frac{4}{a^3}\cdot\frac{b^3}{3} = \frac{4}{3}\left(\frac{b}{a}\right)^3 \approx 5 \times 10^{-15}$. Tiny as it is, this overlap produces the measurable *hyperfine* and finite-size shifts of hydrogen levels. The finite-size shift is far larger in muonic hydrogen, whose measurement in 2010 gave a proton radius in sharp conflict with the electron value. That conflict was resolved about a decade later in favor of the smaller radius used here.

16.6 The Spectrum

A transition $n_2 \to n_1$ emits a photon of energy $E_{n_2} - E_{n_1}$, which reproduces Rydberg's formula (12.3). Not every pair of states is connected. For the dominant *electric dipole* transitions, the rate is proportional to $|\langle \psi_f | \hat{\mathbf{r}} | \psi_i \rangle|^2$, and the angular integrals of spherical harmonics vanish unless

$$\Delta \ell = \pm 1, \qquad \Delta m = 0, \pm 1.$$

These are the *selection rules*. The photon carries one unit of angular momentum, which the atom must give up. The $2s$ state cannot decay to $1s$ by dipole emission, and it lives about 0.12 s, compared with 1.6 ns for $2p$.

Reduced mass. The proton is not infinitely heavy. The two-body problem separates into center-of-mass motion and relative motion with *reduced mass* $\mu = \frac{m_e m_p}{m_e + m_p}$, so every energy is multiplied by $\mu/m_e = 0.999456$. For deuterium the factor is 0.999728, which shifts its lines by about 0.03%. In 1932 Harold Urey discovered deuterium by spotting these slightly shifted Balmer lines, beside the far stronger hydrogen lines, in a sample enriched by evaporation.

Hydrogen-like ions. For a single electron bound to charge Ze, replace e^2 by Ze^2: $E_n = -\frac{13.6 Z^2}{n^2}$ eV and $a \to a/Z$.

16.7 Beyond the Schrödinger Atom

The Schrödinger equation with a Coulomb potential gives levels that depend only on n. Precise spectroscopy reveals small splittings, each of which led to new physics:

- *Fine structure*, of relative size $\alpha^2 \approx 5 \times 10^{-5}$, where $\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c} \approx \frac{1}{137}$. It comes from the relativistic correction to kinetic energy and from *spin–orbit coupling* between the electron's spin (Chapter 17) and its orbital motion. Dirac's relativistic equation of 1928 predicts it exactly.
- The *Lamb shift*, of relative size $\sim \alpha^3$, which splits $2s_{1/2}$ from $2p_{1/2}$. It is caused by the electron's interaction with vacuum fluctuations of the electromagnetic field. Its explanation in 1947–49 launched quantum electrodynamics.
- *Hyperfine structure*, from the interaction of the electron and proton spins. The hyperfine transition in the ground state emits the 21 cm line, with which radio astronomers map hydrogen across the galaxy.

Many-electron atoms. With more than one electron, the Schrödinger equation cannot be solved exactly, because of the electron–electron repulsion. Yet the hydrogen states survive as an organizing principle. Each electron moves approximately in an average central potential that is screened by the others. The degeneracy in ℓ is then broken, and s states, which penetrate close to the nucleus, lie lowest. Filling these states in order of energy, with at most two electrons per orbital (Chapter 18), reproduces the structure of the periodic table: $1s$, $2s$, $2p$, $3s$, $3p$, $4s$, $3d$, and so on.

16.8 Worked Problems

The problems below examine the hydrogen atom from several angles not covered in the text: in momentum space, in its states of maximal angular momentum, through its probability current, and through the small effects of the nucleus's size. A spherical well shows that in three dimensions, unlike one, a bound state needs a minimum depth. Here a is the Bohr radius and m_e the electron mass.

Problem 16.1 Hydrogen in momentum space

Find the momentum-space wavefunction of the hydrogen ground state $\psi_{100} = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}$, verify its normalization, and compute $\langle p^2 \rangle$.

Solution. In three dimensions $\phi(\mathbf{p}) = (2\pi\hbar)^{-3/2} \int \psi(\mathbf{r}) e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} d^3r$. For a spherically symmetric ψ, choose the z-axis along $\mathbf{p}$. The angular integral is $\int_0^{2\pi}\int_0^{\pi} e^{-ipr\cos\theta/\hbar} \sin\theta \, d\theta \, d\phi = \frac{4\pi\hbar \sin(pr/\hbar)}{pr}$, so

$$\phi(p) = (2\pi\hbar)^{-3/2} \frac{4\pi\hbar}{p} \int_0^\infty r\psi(r) \sin\frac{pr}{\hbar} \, dr.$$

With $\int_0^\infty r e^{-r/a} \sin(kr)\, dr = \frac{2k/a}{(k^2 + 1/a^2)^2}$ (the imaginary part of $\int_0^\infty r e^{-(1/a - ik)r} dr = \frac{1}{(1/a - ik)^2}$), this gives

$$\phi(p) = \frac{2\sqrt{2}}{\pi} \frac{a^{3/2}\hbar^{5/2}}{(a^2p^2 + \hbar^2)^2}.$$

Normalization: $\int |\phi|^2 d^3p = 4\pi \int_0^\infty p^2 |\phi|^2 dp = 1$, which the substitution $p = \frac{\hbar}{a}\tan u$ reduces to an elementary trigonometric integral. Similarly $\langle p^2 \rangle = 4\pi \int_0^\infty p^4 |\phi|^2 dp = \frac{\hbar^2}{a^2}$, so $\langle T \rangle = \frac{\hbar^2}{2m_e a^2} = 13.6 \text{ eV} = -E_1$, as the virial theorem requires. The typical momentum is $\hbar/a$, so the typical speed is $\frac{\hbar}{m_e a} = \alpha c$, about $c/137$. That is why the nonrelativistic theory works so well for hydrogen, and why relativistic corrections are of relative order α^2.

Problem 16.2 A spherical well needs a minimum depth

A particle of mass μ is in the well $V = -V_0$ for $r < a$, $V = 0$ for $r > a$. Show that an $\ell = 0$ bound state exists only if $V_0 > \frac{\pi^2\hbar^2}{8\mu a^2}$. Apply this to the deuteron, modeled as a neutron and proton in a well of radius 2.0 fm.

Solution. With $u = rR$, the radial equation for $\ell = 0$ is the one-dimensional Schrödinger equation for u on $r > 0$, with the boundary condition $u(0) = 0$ (so that R is finite). Inside, $u = \sin kr$ with $k = \sqrt{2\mu(E + V_0)}/\hbar$. Outside, $u \propto e^{-\kappa r}$. Matching u'/u at $r = a$ gives $k \cot ka = -\kappa$, which is exactly the *odd* condition of the one-dimensional finite well of half-width a (Theorem 14.14). The odd states of the symmetric well vanish at the origin, just as u must.

By the counting in Theorem 14.14, the first odd state appears when $z_0 > \frac{\pi}{2}$, where $z_0 = \frac{a}{\hbar}\sqrt{2\mu V_0}$. So a bound state requires

$$V_0 > \frac{\pi^2\hbar^2}{8\mu a^2}.$$

In one dimension every attractive well binds. In three dimensions a shallow or narrow well binds nothing: the condition $u(0) = 0$ costs kinetic energy that a weak potential cannot repay.

For the deuteron, $\mu = \frac{m_p m_n}{m_p + m_n}$ with $\mu c^2 = 469.5$ MeV, and $\hbar c = 197.3$ MeV fm. Then $\frac{\pi^2(\hbar c)^2}{8\mu c^2 a^2} = \frac{9.870 \times (197.3)^2}{8 \times 469.5 \times 4.0}$ MeV $= 25.6$ MeV. The deuteron's binding energy is only 2.22 MeV, so the nuclear well must be only modestly deeper than this threshold, about 35 MeV. The deuteron is barely bound: it has no excited states, and the diproton and dineutron, whose effective attraction is slightly weaker, are not bound at all.

Problem 16.3 Circular states and the classical limit

For $\ell = n - 1$, the radial function is $R_{n,n-1} \propto r^{n-1}e^{-r/na}$. Show that $\langle r \rangle = n\left(n + \frac{1}{2}\right)a$ and $\sigma_r/\langle r \rangle = 1/\sqrt{2n+1}$. Interpret the result.

Solution. The radial probability density is $P(r) = r^2R^2 \propto r^{2n}e^{-2r/na}$. With $\beta = \frac{2}{na}$ and $\int_0^\infty r^k e^{-\beta r}dr = \frac{k!}{\beta^{k+1}}$ (Theorem 5.30),

$$\langle r \rangle = \frac{(2n+1)!/\beta^{2n+2}}{(2n)!/\beta^{2n+1}} = \frac{2n+1}{\beta} = n\left(n + \frac{1}{2}\right)a, \qquad \langle r^2 \rangle = \frac{(2n+2)(2n+1)}{\beta^2}.$$

Then $\sigma_r^2 = \frac{(2n+2)(2n+1) - (2n+1)^2}{\beta^2} = \frac{2n+1}{\beta^2}$, so $\frac{\sigma_r}{\langle r \rangle} = \frac{\sqrt{2n+1}}{2n+1} = \frac{1}{\sqrt{2n+1}}$. The density peaks at $r = n^2a$, the radius of the nth Bohr orbit.

As n grows the relative spread shrinks: 58% for $n = 1$, 16% for $n = 20$, 1.4% for $n = 2500$. Since ℓ is maximal, $|m| \le n-1$ can equal ℓ, and the state $Y_{n-1}^{\pm(n-1)} \propto \sin^{n-1}\theta\, e^{\pm i(n-1)\phi}$ is concentrated near the equator, more sharply as n grows. So the circular state with $m = \ell = n-1$ approaches a thin ring at the Bohr radius $n^2 a$: the classical circular orbit. Such *circular Rydberg atoms*, with $n \approx 50$, are made in the laboratory and are among the most classical of quantum objects. Note, however, that the electron's position around the ring is completely uncertain, since ψ has a definite L_z.

Problem 16.4 The current of an orbital, and its magnetic moment

For a hydrogen state $\psi_{n\ell m} = R(r)\Theta(\theta)e^{im\phi}$, show that the probability current is $\mathbf{J} = \dfrac{\hbar m}{m_e r \sin\theta}|\psi|^2\,\hat{\boldsymbol{\phi}}$. Hence show that the state has magnetic moment $\mu_z = -m\mu_B$, where $\mu_B = \frac{e\hbar}{2m_e}$.

Solution. The three-dimensional current is $\mathbf{J} = \frac{\hbar}{m_e}\,\mathrm{Im}(\psi^*\nabla\psi)$. In the gradient of Corollary 7.29, the r- and θ-derivatives act on the real factors R and Θ and give real contributions to $\psi^*\nabla\psi$, which drop out. Only the ϕ-term survives: $\frac{1}{r\sin\theta}\partial_\phi\psi = \frac{im}{r\sin\theta}\psi$. So $\mathbf{J} = \frac{\hbar m}{m_e r\sin\theta}|\psi|^2\hat{\boldsymbol{\phi}}$. The probability circulates about the z-axis, counterclockwise for $m > 0$, with angular velocity $\frac{\hbar m}{m_e\rho^2}$ at distance $\rho = r\sin\theta$ from the axis. For $m = 0$ nothing flows, despite $\ell \neq 0$.

The electron has charge $-e$, so the electric current density is $-e\mathbf{J}$. Think of the distribution as a set of circular loops about the z-axis. A loop of radius ρ with cross-section dA in a half-plane $\phi = \text{const}$ carries current $dI = -eJ\,dA$ and has moment $dI\cdot\pi\rho^2$. Summing,

$$\mu_z = -e\int \frac{\hbar m}{m_e\rho}|\psi|^2\pi\rho^2\,dA = -\frac{e\hbar m}{2m_e}\int|\psi|^2\,2\pi\rho\,dA = -\frac{e\hbar m}{2m_e}\int|\psi|^2 d^3r = -m\mu_B,$$

since $2\pi\rho\,dA$ is the volume of the ring. This is $-\frac{e}{2m_e}\langle L_z\rangle$, the classical gyromagnetic ratio for orbital motion. The energy of the moment in a field $B\hat{\mathbf{k}}$ is $-\mu_z B = m\mu_B B$, which splits the $2\ell+1$ states of a level, the normal Zeeman effect of Chapter 17. The electron's spin has twice this gyromagnetic ratio, which is not explained by any circulating current.

Problem 16.5 The size of the proton

Model the proton as a uniformly charged sphere of radius R. Inside, its potential energy for the electron is $V = -\frac{e^2}{4\pi\varepsilon_0}\frac{3R^2-r^2}{2R^3}$ instead of $-\frac{e^2}{4\pi\varepsilon_0 r}$. Find the first-order shift of the $1s$ energy, and evaluate it for $R = 1.09$ fm (corresponding to the measured rms radius 0.841 fm) in hydrogen and in muonic hydrogen.

Solution. The perturbation $\delta V = \frac{e^2}{4\pi\varepsilon_0}\left(\frac{1}{r} - \frac{3R^2 - r^2}{2R^3}\right)$ for $r < R$ (and 0 outside) is positive: the potential is less deep inside the nucleus. Since $R \ll a$, $|\psi_{100}|^2 \approx \frac{1}{\pi a^3}$ throughout the region where $\delta V \neq 0$. So

$$\Delta E \approx \frac{1}{\pi a^3}\int_0^R \delta V\, 4\pi r^2 dr = \frac{1}{\pi a^3}\cdot\frac{e^2}{4\pi\varepsilon_0}\cdot 4\pi\left(\frac{R^2}{2} - \frac{3R^2}{2}\cdot\frac{1}{3} + \frac{R^2}{10}\right)$$
$$= \frac{2}{5}\,\frac{e^2}{4\pi\varepsilon_0 a}\left(\frac{R}{a}\right)^2.$$

With $\frac{e^2}{4\pi\varepsilon_0 a} = 27.21$ eV and $R/a = 1.09\times10^{-15}/5.29\times10^{-11}$, $\Delta E \approx 4.6\times10^{-9}$ eV, about 3 parts in 10^{10} of the binding energy. Small, but larger than the experimental uncertainty of the hydrogen $1s$–$2s$ frequency.

In muonic hydrogen, the muon is 207 times heavier than the electron (reduced mass $185.8m_e$), so $a_\mu = a/185.8 = 285$ fm. The shift scales as μ^3: $\Delta E \approx 4.6\times10^{-9}\times185.8^3$ eV ≈ 0.029 eV, now a substantial correction. In 2010, measurements of muonic hydrogen gave a proton radius about 4% smaller than the accepted value from electronic hydrogen. This *proton radius puzzle* was resolved over the following decade, largely by new hydrogen measurements that agreed with the smaller muonic value, 0.84 fm.

Problem 16.6 Inside the nucleus, by angular momentum

For small r, the radial function behaves as $R_{n\ell} \propto r^\ell$. Find, to leading order in R/a, the probability that the electron is inside a nucleus of radius R for the $2s$ and $2p$ states, and explain why only s states feel the nucleus appreciably.

Solution. Near the origin the centrifugal term $\frac{\ell(\ell+1)}{r^2}$ dominates the radial equation (16.2), and the solution regular at 0 is $u \propto r^{\ell+1}$, so $R \propto r^\ell$. The probability inside radius R is $\int_0^R R_{n\ell}^2 r^2 dr \propto R^{2\ell+3}$.

For $2s$, $R_{20} = \frac{1}{\sqrt{2}}a^{-3/2}\left(1 - \frac{r}{2a}\right)e^{-r/2a} \approx \frac{1}{\sqrt{2}}a^{-3/2}$ near 0, so $P \approx \frac{1}{2}a^{-3}\cdot\frac{R^3}{3} = \frac{1}{6}\left(\frac{R}{a}\right)^3$.
For $2p$, $R_{21} = \frac{1}{\sqrt{24}}a^{-3/2}\frac{r}{a}e^{-r/2a} \approx \frac{1}{\sqrt{24}}a^{-5/2}r$, so $P \approx \frac{1}{24a^5}\cdot\frac{R^5}{5} = \frac{1}{120}\left(\frac{R}{a}\right)^5$.

With $R/a \approx 2\times10^{-5}$, the $2s$ probability is about 10^{-15} and the $2p$ probability about 3×10^{-26}, eleven orders of magnitude smaller. The centrifugal barrier keeps $\ell > 0$ electrons out of the nucleus. That is why finite-size shifts, hyperfine contact interactions, and electron capture by nuclei all involve s electrons, and why the $2s$ and $2p$ levels of hydrogen, degenerate in the Schrödinger theory, are split slightly by effects concentrated at the nucleus (the Lamb shift, described at the end of the chapter).

Problem 16.7 The classically forbidden region

A classical electron with the ground-state energy $E_1 = -\frac{e^2}{8\pi\varepsilon_0 a}$ cannot be farther than r_c from the proton, where $V(r_c) = E_1$. Find r_c and the quantum probability of finding the electron beyond it.

Solution. $-\frac{e^2}{4\pi\varepsilon_0 r_c} = -\frac{e^2}{8\pi\varepsilon_0 a}$ gives $r_c = 2a$. The radial density is $P(r) = \frac{4r^2}{a^3}e^{-2r/a}$, and with $x = 2r/a$,

$$P(r > 2a) = \int_4^\infty \frac{x^2}{2}e^{-x}dx = \frac{1}{2}\left[e^{-x}(x^2+2x+2)\right]_{x=4}$$
$$= \frac{1}{2}e^{-4}(16+8+2) = 13e^{-4} \approx 0.238.$$

Nearly a quarter of the time, the electron is where a classical particle with its energy could never be. In one dimension, the oscillator ground state spends 16% of its probability in the forbidden region. Hydrogen's long exponential tail, compared with the oscillator's Gaussian tail, explains the larger value. Forbidden-region probabilities are not paradoxical: a measurement that finds the electron at $r > 2a$ localizes it, and the localization supplies the kinetic energy that the classical estimate omits.

EXERCISES

A. COMPUTATION

16.1 List the energies and degeneracies of the first six levels of the cubical box. Which degeneracies are accidental?

16.2 A particle is in a box with sides a, a, $2a$. Find the first four energy levels and their degeneracies.

16.3 Verify that Y_1^0 and Y_1^1 satisfy $\Lambda Y = -2Y$.

16.4 Verify that Y_2^0 is normalized and orthogonal to Y_0^0.

16.5 Express $\sin^2\theta\cos 2\phi$ as a combination of spherical harmonics.

16.6 A state has angular wavefunction $\frac{1}{\sqrt{2}}(Y_1^1 + Y_2^0)$. Find the probabilities of each value of L^2 and of L_z.

16.7 Write the d_{z^2} orbital Y_2^0 and the combination $\frac{1}{\sqrt{2}}(Y_2^2 + Y_2^{-2})$ in Cartesian form $(x/r, y/r, z/r)$.

16.8 Verify that R_{10} and R_{20} are normalized and orthogonal (with $r^2\,dr$).

16.9 Verify that $u = rR_{21}$ satisfies the radial equation (16.2) with $\ell = 1$ and $E = E_2$.

16.10 Find the most probable radius for the $2p$ state and for the $3d$ state. Compare with the Bohr radii $4a$ and $9a$.

16.11 Find $\langle r \rangle$ and $\langle r^2 \rangle$ in the $2s$ and $2p$ states, and check Theorem 16.5.

16.12 Find the location of the radial node of the $2s$ state and the two radial nodes of the $3s$ state.

16.13 Compute $\langle V \rangle$ and $\langle T \rangle$ in the hydrogen ground state, and verify the virial theorem.

16.14 Find the probability that the ground-state electron is found at $r < a$.

16.15 Find the wavelengths of the $3 \to 2$ line in hydrogen and in deuterium, and their difference.

16.16 Find the ionization energy and the $2 \to 1$ wavelength of He^+.

16.17 Which of the following transitions are dipole-allowed: $3s \to 2p$, $3d \to 2s$, $3p \to 1s$, $4f \to 3d$, $2s \to 1s$?

16.18 Compute $\langle \psi_{100} | z | \psi_{210} \rangle$. (Use $z = r\cos\theta$.)

16.19 An electron is in the state $\psi = \frac{1}{\sqrt{10}}\big(2\psi_{100} - \psi_{210} + \sqrt{2}\psi_{211} + \sqrt{3}\psi_{21,-1}\big)$. Find $\langle H \rangle$, $\langle L^2 \rangle$, and $\langle L_z \rangle$.

16.20 Estimate the size of the fine-structure splitting of the $n = 2$ level in eV, as $\alpha^2 |E_2|$.

B. THEORY AND PROOF

16.21 Derive (16.2) from (16.1) using Corollary 7.29 and $u = rR$.

16.22 Show that the substitution $u = \rho^{\ell+1} e^{-\rho} v$ transforms (16.4) into (16.5), and derive the recursion relation.

16.23 Show that $Y_\ell^m(\pi - \theta, \phi + \pi) = (-1)^\ell Y_\ell^m(\theta, \phi)$: the spherical harmonics have parity $(-1)^\ell$.

16.24 Prove the selection rule $\Delta m = 0, \pm 1$ for $\langle n'\ell'm' | \hat{\mathbf{r}} | n\ell m \rangle$ by computing the ϕ-integrals of $x \pm iy = r\sin\theta e^{\pm i\phi}$ and $z = r\cos\theta$.

16.25 Use the virial theorem to show $\langle T \rangle = -E_n$ and $\langle V \rangle = 2E_n$ for all hydrogen states.

16.26 Prove $\sum_{m=-\ell}^{\ell} |Y_\ell^m(\theta,\phi)|^2 = \frac{2\ell+1}{4\pi}$ for $\ell = 0, 1$ (Unsöld's theorem). What does it imply about a filled subshell?

16.27 Show that for a two-body system with potential $V(|\mathbf{r}_1 - \mathbf{r}_2|)$, the Hamiltonian separates into center-of-mass and relative parts, and identify the reduced mass.

C. CHALLENGE

16.28 Solve the three-dimensional isotropic oscillator $V = \frac{1}{2}m\omega^2 r^2$ in Cartesian coordinates, show $E = (N + \frac{3}{2})\hbar\omega$ with degeneracy $\frac{1}{2}(N+1)(N+2)$, and identify which values of ℓ occur at each N.

16.29 Find the bound states of the spherical infinite well ($V = 0$ for $r < a$, ∞ outside) with $\ell = 0$, and show that $E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}$.

16.30 Use the Hellmann–Feynman theorem with $\lambda = e^2$ and with $\lambda = \ell$ to derive $\langle 1/r \rangle$ and $\langle 1/r^2 \rangle$ in Theorem 16.5.

16.31 (Kramers' relation.) Show that $\frac{s+1}{n^2}\langle r^s \rangle - (2s+1)a\langle r^{s-1}\rangle + \frac{s}{4}\big[(2\ell+1)^2 - s^2\big]a^2\langle r^{s-2}\rangle = 0$, and use it to derive $\langle r \rangle$ and $\langle r^2 \rangle$.

CHAPTER 17

Angular Momentum and Spin

LEARNING OBJECTIVES

- Derive the commutation relations $[\hat{L}_i, \hat{L}_j] = i\hbar\varepsilon_{ijk}\hat{L}_k$ from $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$.
- Find the spectrum of any angular momentum algebraically using ladder operators, including half-integer values.
- Construct the matrices of angular momentum for small j, and relate angular momentum to rotations.
- Describe spin-$\frac{1}{2}$ with Pauli matrices and spinors, and find spin states along an arbitrary direction (the Bloch sphere).
- Derive Larmor precession of a spin in a magnetic field, and explain the 4π periodicity of spinors.
- Add two angular momenta, construct singlet and triplet states, and state the general rules and the meaning of Clebsch–Gordan coefficients.
- Explain spin–orbit coupling and the normal Zeeman effect qualitatively.

The Stern–Gerlach experiment (Chapter 12) revealed that the electron has an intrinsic angular momentum with only two possible values along any axis. This *spin* has no classical analog: it cannot be the rotation of a small charged ball, as the required surface speed would exceed the speed of light. Yet spin and orbital angular momentum obey the same algebra. This chapter develops that algebra from the commutation relations alone, and shows that it permits exactly the integer and half-integer values seen in nature. The cross product of Chapter 1 and the Levi-Civita symbol of its index notation reappear here as the structure of quantum angular momentum.

17.1 Orbital Angular Momentum

Classically $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. In quantum mechanics define $\hat{L}_i = \varepsilon_{ijk}\hat{x}_j\hat{p}_k$, so $\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$ and cyclically. There is no ordering ambiguity, since $\hat{x}_j$ and $\hat{p}_k$ commute for $j \neq k$.

Theorem 17.1 (Angular momentum commutators).

$$[\hat{L}_i, \hat{L}_j] = i\hbar\,\varepsilon_{ijk}\hat{L}_k, \qquad [\hat{L}^2, \hat{L}_i] = 0,$$

where $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$. *In particular* $[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z$.

Proof. Using $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$ and the product rule for commutators,

$$\begin{aligned}[\hat{L}_x, \hat{L}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y,\ \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] = \hat{y}[\hat{p}_z, \hat{z}]\hat{p}_x + \hat{x}[\hat{z}, \hat{p}_z]\hat{p}_y \\ &= -i\hbar\hat{y}\hat{p}_x + i\hbar\hat{x}\hat{p}_y = i\hbar\hat{L}_z,\end{aligned}$$

since the other cross terms commute. The other relations follow by cyclic permutation. For $\hat{L}^2$, $[\hat{L}_x^2, \hat{L}_z] = \hat{L}_x[\hat{L}_x, \hat{L}_z] + [\hat{L}_x, \hat{L}_z]\hat{L}_x = -i\hbar(\hat{L}_x\hat{L}_y + \hat{L}_y\hat{L}_x)$, and $[\hat{L}_y^2, \hat{L}_z] = +i\hbar(\hat{L}_y\hat{L}_x + \hat{L}_x\hat{L}_y)$. These cancel, and $[\hat{L}_z^2, \hat{L}_z] = 0$. ∎

The components of angular momentum do not commute with each other, so they are incompatible. By Theorem 15.3, $\sigma_{L_x}\sigma_{L_y} \geq \frac{\hbar}{2}|\langle L_z\rangle|$. But $\hat{L}^2$ commutes with each component, so we can specify L^2 together with one component, conventionally L_z, but not two components. A quantum angular momentum vector cannot point in a definite direction.

In spherical coordinates, using the gradient of Corollary 7.29,

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}, \qquad \hat{L}^2 = -\hbar^2\Lambda = -\hbar^2\Big[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial}{\partial\theta}\Big) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\Big].$$

So the spherical harmonics of Theorem 16.2 are simultaneous eigenfunctions:

$$\hat{L}^2 Y_\ell^m = \hbar^2\ell(\ell+1)Y_\ell^m, \qquad \hat{L}_z Y_\ell^m = \hbar m\, Y_\ell^m.$$

This confirms the identification $L^2 = \hbar^2\ell(\ell+1)$ made from the centrifugal term in Chapter 16. The Hamiltonian of a central potential commutes with $\hat{L}^2$ and $\hat{L}_z$, by rotational symmetry (Theorem 15.8). That is why the hydrogen states can be labeled by n, ℓ, and m together.

17.2 The Algebraic Theory of Angular Momentum

We now forget where the operators came from. Call any three Hermitian operators $\hat{J}_x, \hat{J}_y, \hat{J}_z$ satisfying $[\hat{J}_i, \hat{J}_j] = i\hbar\varepsilon_{ijk}\hat{J}_k$ an *angular momentum*, and find all possible eigenvalues from the algebra alone.

Definition 17.2 (Ladder operators). $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$. Then $\hat{J}_\pm^\dagger = \hat{J}_\mp$, and

$$[\hat{J}_z, \hat{J}_\pm] = \pm\hbar\hat{J}_\pm, \qquad [\hat{J}^2, \hat{J}_\pm] = 0, \qquad \hat{J}_\mp\hat{J}_\pm = \hat{J}^2 - \hat{J}_z^2 \mp \hbar\hat{J}_z.$$

Theorem 17.3 (Spectrum of angular momentum). *The simultaneous eigenvalues of $\hat{J}^2$ and $\hat{J}_z$ are $\hbar^2 j(j+1)$ and $\hbar m$, where*

$$j \in \left\{0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \dots\right\}, \qquad m \in \{-j, -j+1, \dots, j-1, j\}.$$

With $|j,m\rangle$ normalized,

$$\hat{J}_\pm|j,m\rangle = \hbar\sqrt{j(j+1) - m(m\pm1)}\;|j, m\pm1\rangle.$$

Proof. Let $\hat{J}^2|\lambda,m\rangle = \hbar^2\lambda|\lambda,m\rangle$ and $\hat{J}_z|\lambda,m\rangle = \hbar m|\lambda,m\rangle$.

Raising and lowering. Since $\hat{J}_z\hat{J}_\pm = \hat{J}_\pm(\hat{J}_z \pm \hbar)$, the vector $\hat{J}_\pm|\lambda,m\rangle$, if nonzero, is an eigenvector of $\hat{J}_z$ with eigenvalue $\hbar(m\pm1)$. It has the same λ, since $\hat{J}^2$ commutes with $\hat{J}_\pm$.

Bounds. $\lambda - m^2 = \frac{1}{\hbar^2}\langle \hat{J}_x^2 + \hat{J}_y^2\rangle \geq 0$, so $|m| \leq \sqrt{\lambda}$. Raising cannot continue forever. There is a top value $m = j$ with $\hat{J}_+|\lambda,j\rangle = 0$. Then $0 = \hat{J}_-\hat{J}_+|\lambda,j\rangle = (\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|\lambda,j\rangle = \hbar^2(\lambda - j^2 - j)|\lambda,j\rangle$, so $\lambda = j(j+1)$. Similarly, there is a bottom value $m = j'$ with $\hat{J}_-|\lambda,j'\rangle = 0$, and $\lambda = j'(j'-1)$. Hence $j(j+1) = j'(j'-1)$, whose solutions are $j' = -j$ or $j' = j+1$. The second is impossible since $j' \leq j$, so $j' = -j$.

Quantization. Lowering from $m = j$ reaches $m = -j$ in an integer number of steps, $2j$. So $2j$ is a nonnegative integer.

Normalization. $\|\hat{J}_+|j,m\rangle\|^2 = \langle j,m|\hat{J}_-\hat{J}_+|j,m\rangle = \hbar^2\big(j(j+1) - m^2 - m\big)$. ∎

The algebra allows half-integer j. Orbital angular momentum uses only the integers, because $e^{im\phi}$ must be single-valued. Nature uses the half-integers too, for spin. The electron, proton, and neutron have $s = \frac{1}{2}$; the photon has spin 1; the Δ

baryons have spin $\frac{3}{2}$. Particles with half-integer spin are *fermions*, and those with integer spin are *bosons*, a distinction with profound consequences (Chapter 18).

Example 17.4 (Matrices for $j = 1$). In the basis $|1,1\rangle, |1,0\rangle, |1,-1\rangle$, Theorem 17.3 gives $\hat{J}_+|1,0\rangle = \sqrt{2}\hbar|1,1\rangle$ and $\hat{J}_+|1,-1\rangle = \sqrt{2}\hbar|1,0\rangle$. So

$$\hat{J}_z = \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}, \qquad \hat{J}_+ = \sqrt{2}\hbar\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}, \qquad \hat{J}_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}.$$

One checks $\hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 = 2\hbar^2 I$, and that $\hat{J}_x$ has eigenvalues $\hbar, 0, -\hbar$, as it must by rotational symmetry.

17.2.1 ANGULAR MOMENTUM AND ROTATIONS

Just as momentum generates translations (Chapter 15), angular momentum generates rotations. The operator

$$\hat{R}_{\hat{\mathbf{n}}}(\alpha) = e^{-i\alpha\,\hat{\mathbf{n}}\cdot\hat{\mathbf{J}}/\hbar}$$

rotates a state by angle α about the axis $\hat{\mathbf{n}}$. For orbital angular momentum about z, $\hat{R}_z(\alpha) = e^{-\alpha\,\partial_\phi}$, which maps $f(\phi)$ to $f(\phi - \alpha)$ by (4.1). On an eigenstate $|j,m\rangle$, a rotation by 2π about z gives

$$\hat{R}_z(2\pi)|j,m\rangle = e^{-2\pi i m}|j,m\rangle = (-1)^{2j}|j,m\rangle.$$

For integer j, a full turn is the identity. For half-integer j, *a rotation by 2π multiplies the state by -1*, and only a rotation by 4π restores it (Example 9.32). The sign is a global phase for a single particle and unobservable. But in an interferometer, where a beam is split and only one half is rotated, the relative phase of -1 is measurable. Neutron interferometry experiments in 1975 confirmed it directly.

17.3 SPIN ONE-HALF

For $s = \frac{1}{2}$ the state space is $\mathbb{C}^2$, with basis $|\uparrow\rangle = |\frac{1}{2},\frac{1}{2}\rangle$ and $|\downarrow\rangle = |\frac{1}{2},-\frac{1}{2}\rangle$. Theorem 17.3 gives $\hat{S}_+|\downarrow\rangle = \hbar|\uparrow\rangle$, so

$$\hat{\mathbf{S}} = \frac{\hbar}{2}\boldsymbol{\sigma}, \qquad \hat{S}_x = \frac{\hbar}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}, \quad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix}0&-i\\i&0\end{pmatrix}, \quad \hat{S}_z = \frac{\hbar}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix},$$

the Pauli matrices of Example 9.24. Then $\hat{S}^2 = \frac{3}{4}\hbar^2 I$, consistent with $s(s+1)\hbar^2$. A general state is a *spinor* $|\chi\rangle = a|\uparrow\rangle + b|\downarrow\rangle$ with $|a|^2 + |b|^2 = 1$. Measuring S_z gives $+\frac{\hbar}{2}$ with probability $|a|^2$.

Theorem 17.5 (Spin along an arbitrary axis). *Let* $\hat{\mathbf{n}} = \langle \sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\rangle$. *The eigenvectors of* $\hat{\mathbf{n}}\cdot\hat{\mathbf{S}}$ *are*

$$|+\hat{\mathbf{n}}\rangle = \cos\tfrac{\theta}{2}\,|\uparrow\rangle + e^{i\phi}\sin\tfrac{\theta}{2}\,|\downarrow\rangle, \qquad |-\hat{\mathbf{n}}\rangle = \sin\tfrac{\theta}{2}\,|\uparrow\rangle - e^{i\phi}\cos\tfrac{\theta}{2}\,|\downarrow\rangle,$$

with eigenvalues $\pm\frac{\hbar}{2}$. *Every spin-*$\frac{1}{2}$ *state is* $|+\hat{\mathbf{n}}\rangle$ *for some direction* $\hat{\mathbf{n}}$, *up to a phase.*

Proof. $\hat{\mathbf{n}}\cdot\boldsymbol{\sigma} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}$. Applying it to $|+\hat{\mathbf{n}}\rangle$, the first component is $\cos\theta\cos\frac{\theta}{2} + \sin\theta\sin\frac{\theta}{2} = \cos\left(\theta - \frac{\theta}{2}\right) = \cos\frac{\theta}{2}$, and the second is $e^{i\phi}\left(\sin\theta\cos\frac{\theta}{2} - \cos\theta\sin\frac{\theta}{2}\right) = e^{i\phi}\sin\frac{\theta}{2}$. So the eigenvalue is $+1$. For the last claim, any normalized $\langle a, b\rangle$ can be multiplied by a phase to make a real and nonnegative; then $a = \cos\frac{\theta}{2}$ and $b = e^{i\phi}\sin\frac{\theta}{2}$ for unique $\theta \in [0, \pi]$ and ϕ. ∎

Every pure state of a spin-$\frac{1}{2}$ particle therefore corresponds to a point on a sphere, the *Bloch sphere*: the spin "points" along $\hat{\mathbf{n}}$. This is the sense in which θ and ϕ, the spherical angles of Chapter 1, parameterize qubit states (Chapter 18). Note the half-angles. Orthogonal states, $|+\hat{\mathbf{n}}\rangle$ and $|-\hat{\mathbf{n}}\rangle$, correspond to *antipodal* points, not perpendicular ones.

Corollary 17.6 (Stern–Gerlach probabilities). *If a spin is prepared in* $|+\hat{\mathbf{n}}\rangle$ *and* S *is then measured along* $\hat{\mathbf{m}}$, *the probability of the result* $+\frac{\hbar}{2}$ *is* $\frac{1}{2}(1 + \hat{\mathbf{n}}\cdot\hat{\mathbf{m}}) = \cos^2\frac{\gamma}{2}$, *where* γ *is the angle between* $\hat{\mathbf{n}}$ *and* $\hat{\mathbf{m}}$.

This gives the probabilities of the sequential Stern–Gerlach experiments of Chapter 12. For perpendicular axes, $\gamma = 90^\circ$ and the probability is $\frac{1}{2}$.

17.4 Spin in a Magnetic Field

A spinning charge has a magnetic moment proportional to its angular momentum, $\hat{\boldsymbol{\mu}} = \gamma\hat{\mathbf{S}}$, where γ is the *gyromagnetic ratio*. For the electron $\gamma = -g\frac{e}{2m_e}$ with $g \approx 2.0023$. The value $g = 2$ is predicted by Dirac's equation, and the small excess was computed by quantum electrodynamics; theory and experiment now agree to about one part in 10^{12}. In a field $\mathbf{B}$ the Hamiltonian is

$$\hat{H} = -\hat{\boldsymbol{\mu}}\cdot\mathbf{B} = -\gamma\mathbf{B}\cdot\hat{\mathbf{S}}.$$

Theorem 17.7 (Larmor precession). *In a uniform field $\mathbf{B} = B\hat{\mathbf{k}}$, the expectation value of spin precesses about $\hat{\mathbf{k}}$ with angular velocity $\boldsymbol{\omega} = -\gamma B\hat{\mathbf{k}}$:*

$$\frac{\mathrm{d}\langle \hat{\mathbf{S}} \rangle}{\mathrm{d}t} = \gamma \, \langle \hat{\mathbf{S}} \rangle \times \mathbf{B},$$

with $\langle \hat{S}_z \rangle$ constant and $\langle \hat{S}_x \rangle$, $\langle \hat{S}_y \rangle$ rotating at the Larmor frequency $\omega_L = \gamma B$.

Proof. By Theorem 15.5 with $\hat{H} = -\gamma B \hat{S}_z$, $\mathrm{d}\langle \hat{S}_x \rangle/\mathrm{d}t = \frac{i}{\hbar}\langle [\hat{H}, \hat{S}_x] \rangle = -\frac{i\gamma B}{\hbar}\langle [\hat{S}_z, \hat{S}_x] \rangle = -\frac{i\gamma B}{\hbar}\, i\hbar \langle \hat{S}_y \rangle = \gamma B \langle \hat{S}_y \rangle$. Similarly $\mathrm{d}\langle \hat{S}_y \rangle/\mathrm{d}t = -\gamma B \langle \hat{S}_x \rangle$, and $\langle \hat{S}_z \rangle$ is constant because $\hat{S}_z$ commutes with $\hat{H}$. These are the components of $\gamma \langle \hat{\mathbf{S}} \rangle \times \mathbf{B}$. ∎

The expectation value obeys exactly the classical equation of motion for a magnetic moment in a field, the torque equation $\mathrm{d}\mathbf{L}/\mathrm{d}t = \boldsymbol{\mu} \times \mathbf{B}$ of Chapter 2. On the Bloch sphere, the state point circles the field axis at constant latitude. Explicitly, starting from $|{+\hat{\mathbf{x}}}\rangle$, the state at time t is $\frac{1}{\sqrt{2}}\big(e^{i\gamma Bt/2}|{\uparrow}\rangle + e^{-i\gamma Bt/2}|{\downarrow}\rangle\big)$. The relative phase advances by γBt, so the spin points in the direction $\phi = -\gamma Bt$.

Example 17.8 (Magnetic resonance). For protons $\gamma_p = 2.675 \times 10^8 \,\mathrm{s^{-1}T^{-1}}$, so in a 1.5 T MRI magnet the Larmor frequency is $f = \frac{\gamma_p B}{2\pi} = 63.9$ MHz, in the radio band. A radio pulse at exactly this frequency tips the proton spins away from the field (Chapter 18). As they precess and relax back, they emit a signal whose frequency encodes position when a field gradient is applied. That is how MRI makes images. For electrons in 1 T, $f \approx 28$ GHz, in the microwave band, used in electron spin resonance.

17.5 Addition of Angular Momenta

A system of two particles with angular momenta $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ has state space the *tensor product* $\mathcal{H}_1 \otimes \mathcal{H}_2$, with basis $|j_1 m_1\rangle \otimes |j_2 m_2\rangle$, written $|m_1 m_2\rangle$ for short. Its dimension is $(2j_1+1)(2j_2+1)$. The total angular momentum $\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2$, with each part acting on its own factor, satisfies the angular momentum algebra. The question is how its eigenstates $|J, M\rangle$ relate to the product states.

17.5.1 TWO SPIN-$\frac{1}{2}$ PARTICLES

The product basis is $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$. Since $\hat{S}_z = \hat{S}_{1z} + \hat{S}_{2z}$, these have $M = 1, 0, 0, -1$. Apply the lowering operator $\hat{S}_- = \hat{S}_{1-} + \hat{S}_{2-}$ to the top state $|\uparrow\uparrow\rangle$, which must have $S = 1$:

$$\hat{S}_- |\uparrow\uparrow\rangle = \hbar\big(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle\big) = \sqrt{2}\hbar\, |1, 0\rangle.$$

Theorem 17.9 (Singlet and triplet). *The states of two spin-$\frac{1}{2}$ particles decompose into a* triplet *with $S = 1$,*

$$|1, 1\rangle = |\uparrow\uparrow\rangle, \qquad |1, 0\rangle = \tfrac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\big), \qquad |1, -1\rangle = |\downarrow\downarrow\rangle,$$

and a singlet *with $S = 0$,*

$$|0, 0\rangle = \tfrac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\big).$$

The triplet states are symmetric under exchange of the two particles, and the singlet is antisymmetric.

Proof. The triplet is generated by lowering as above. The singlet is the remaining state with $M = 0$, orthogonal to $|1, 0\rangle$. To confirm $S = 0$, use $\hat{S}^2 = \hat{S}_1^2 + \hat{S}_2^2 + 2\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2$ and $\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 = \hat{S}_{1z}\hat{S}_{2z} + \frac{1}{2}(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+})$. On the singlet, $\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 = -\frac{3}{4}\hbar^2$, so $\hat{S}^2 = \frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2 - \frac{3}{2}\hbar^2 = 0$. ∎

In the notation of dimensions, $2 \otimes 2 = 3 \oplus 1$. The singlet is rotationally invariant: it has the same form, $\frac{1}{\sqrt{2}}(|+\hat{\mathbf{n}}, -\hat{\mathbf{n}}\rangle - |-\hat{\mathbf{n}}, +\hat{\mathbf{n}}\rangle)$, along *every* axis $\hat{\mathbf{n}}$. Whatever axis both spins are measured along, the results are always opposite. This perfect anticorrelation along every axis is the starting point for the Bell inequality in Chapter 18.

17.5.2 THE GENERAL RULE

Theorem 17.10 (Addition of angular momenta). *$\mathcal{H}_{j_1} \otimes \mathcal{H}_{j_2}$ decomposes into states of total angular momentum*

$$J = |j_1 - j_2|,\ |j_1 - j_2| + 1,\ \ldots,\ j_1 + j_2,$$

each occurring once. The new basis is related to the product basis by the Clebsch–

Gordan coefficients:

$$|J,M\rangle = \sum_{m_1+m_2=M} \langle j_1 m_1; j_2 m_2 | JM\rangle \, |m_1 m_2\rangle .$$

The dimensions check: $\sum_{J=|j_1-j_2|}^{j_1+j_2}(2J+1) = (2j_1+1)(2j_2+1)$. The coefficients are tabulated, and they are computed exactly as for the triplet: lower from the top state of each multiplet, and fill in the remaining states by orthogonality.

17.5.3 SPIN–ORBIT COUPLING

An electron in an atom has orbital angular momentum ℓ and spin $\frac{1}{2}$, so its total angular momentum is $j = \ell \pm \frac{1}{2}$ (or only $j = \frac{1}{2}$ if $\ell = 0$). In the electron's rest frame, the orbiting nucleus produces a magnetic field proportional to $\mathbf{L}$, and the spin's moment couples to it through $\hat{H}_{SO} \propto \frac{1}{r^3}\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$. Since $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}} = \frac{1}{2}(\hat{J}^2 - \hat{L}^2 - \hat{S}^2)$ is diagonal in the $|j\, m_j\rangle$ basis, degenerate perturbation theory (Chapter 15) is easy in that basis. This coupling splits the $2p$ level of hydrogen into $2p_{1/2}$ and $2p_{3/2}$, part of the fine structure. In sodium the same splitting separates the familiar yellow D lines at 589.0 and 589.6 nm.

17.5.4 THE ZEEMAN EFFECT

In a weak magnetic field $B\hat{\mathbf{k}}$, ignoring spin, the orbital moment $-\frac{e}{2m_e}\mathbf{L}$ adds the energy $\frac{eB}{2m_e}\hat{L}_z = \mu_B B m$. Each level of angular momentum ℓ splits into $2\ell+1$ equally spaced sublevels, and a line splits into three components (by the selection rule $\Delta m = 0, \pm 1$). This *normal Zeeman effect* gave the first measurement of e/m_e for electrons inside atoms. The more common *anomalous* Zeeman pattern, with more components, was a decade-long puzzle until spin, with its $g = 2$, explained it in 1925.

17.6 WORKED PROBLEMS

The problems below decompose wavefunctions into angular momentum eigenstates, rotate a spin-1 state, treat spin operators as vectors under rotation, couple angular momenta to explain fine and hyperfine structure, and count the states of three spins.

Problem 17.1 Reading angular momentum off a wavefunction

A particle has wavefunction $\psi = f(r)\,xy$ or $\psi = g(r)\,z^2$, with f, g radial. For each, find the possible results of measuring L^2 and L_z, with their probabilities.

Solution. $\psi = f(r)xy$. In spherical coordinates $xy = r^2\sin^2\theta\cos\phi\sin\phi = \frac{r^2}{2}\sin^2\theta\sin 2\phi$, and $\sin 2\phi = \frac{1}{2i}(e^{2i\phi} - e^{-2i\phi})$. Since $Y_2^{\pm 2} \propto \sin^2\theta\, e^{\pm 2i\phi}$ with equal normalization constants, $\psi \propto Y_2^2 - Y_2^{-2}$. So $L^2 = 6\hbar^2$ with certainty ($\ell = 2$), and $L_z = \pm 2\hbar$ with probability $\frac{1}{2}$ each. A product of two Cartesian coordinates, homogeneous of degree 2 and harmonic ($\nabla^2(xy) = 0$), is a pure $\ell = 2$ function.

$\psi = g(r)z^2$. Now $z^2 = r^2\cos^2\theta$, and $\cos^2\theta = \frac{1}{3} + \frac{2}{3}P_2(\cos\theta)$ with $P_2 = \frac{1}{2}(3\cos^2\theta - 1)$. Since ϕ does not appear, $L_z = 0$ with certainty. But z^2 is not harmonic ($\nabla^2 z^2 = 2$), and it mixes $\ell = 0$ and $\ell = 2$. The probability of $\ell = 0$ is the squared overlap of the normalized angular function with Y_0^0:

$$P(\ell = 0) = \frac{\left|\int Y_0^0 \cos^2\theta\, d\Omega\right|^2}{\int \cos^4\theta\, d\Omega} = \frac{\left(\frac{1}{\sqrt{4\pi}}\cdot\frac{4\pi}{3}\right)^2}{4\pi/5} = \frac{5}{9}, \qquad P(\ell = 2) = \frac{4}{9}.$$

So $L^2 = 0$ with probability $\frac{5}{9}$ and $6\hbar^2$ with probability $\frac{4}{9}$, and $\langle L^2\rangle = \frac{8}{3}\hbar^2$. The radial factor plays no role, since $\hat{L}^2$ and $\hat{L}_z$ act only on angles. In general, a homogeneous polynomial of degree ℓ contains Y_ℓ, $Y_{\ell-2}$, ..., and only its harmonic part is pure ℓ.

Problem 17.2 Rotating a spin-1 state

A spin-1 particle is in $|1,1\rangle$ along z. Its spin is measured along a direction $\hat{\mathbf{n}}$ in the xz-plane at angle θ from z. Find the probabilities of $m_n = 1, 0, -1$.

Solution. The eigenstates of $\hat{\mathbf{n}}\cdot\hat{\mathbf{J}}$ are $|1,m\rangle_n = \hat{R}_y(\theta)|1,m\rangle$, so the amplitudes are ${}_n\langle 1,m|1,1\rangle = \langle 1,m|\hat{R}_y(\theta)^\dagger|1,1\rangle$, the complex conjugates of the entries in the first row of the matrix $d^{(1)}(\theta) = e^{-i\theta\hat{J}_y/\hbar}$. From Example 17.4, $\hat{J}_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$.

One can verify that $(\hat{J}_y/\hbar)^3 = \hat{J}_y/\hbar$, since its eigenvalues are $1, 0, -1$, so the exponential series collapses:

$$e^{-i\theta\hat{J}_y/\hbar} = I - i\sin\theta\,\frac{\hat{J}_y}{\hbar} - (1-\cos\theta)\left(\frac{\hat{J}_y}{\hbar}\right)^2 = \begin{pmatrix} \frac{1+\cos\theta}{2} & -\frac{\sin\theta}{\sqrt{2}} & \frac{1-\cos\theta}{2} \\ \frac{\sin\theta}{\sqrt{2}} & \cos\theta & -\frac{\sin\theta}{\sqrt{2}} \\ \frac{1-\cos\theta}{2} & \frac{\sin\theta}{\sqrt{2}} & \frac{1+\cos\theta}{2} \end{pmatrix}.$$

The probabilities are the squares of the entries of the first row:

$$P(1) = \left(\frac{1+\cos\theta}{2}\right)^2 = \cos^4\frac{\theta}{2}, \qquad P(0) = \frac{\sin^2\theta}{2},$$
$$P(-1) = \left(\frac{1-\cos\theta}{2}\right)^2 = \sin^4\frac{\theta}{2}.$$

They sum to 1. A numerical exponentiation at $\theta = 0.73$ agrees to six digits. Compare spin $\frac{1}{2}$, where $P(+) = \cos^2\frac{\theta}{2}$: the spin-1 probabilities are those of two independent spin-$\frac{1}{2}$ particles, each found up with probability $\cos^2\frac{\theta}{2}$. That is no accident. The state $|1,1\rangle$ is $|\uparrow\uparrow\rangle$ of two spins (Theorem 17.9), and rotations act on each spin separately.

Problem 17.3 Spin is a vector operator

Let $\hat{U}(\varphi) = e^{-i\varphi\hat{S}_z/\hbar} = e^{-i\varphi\sigma_z/2}$. Show that $\hat{U}^\dagger\sigma_x\hat{U} = \cos\varphi\,\sigma_x - \sin\varphi\,\sigma_y$, and find $\hat{U}^\dagger\sigma_y\hat{U}$. Interpret the result in terms of expectation values.

Solution. $\hat{U} = \operatorname{diag}(e^{-i\varphi/2}, e^{i\varphi/2})$. Conjugating an off-diagonal matrix multiplies its entries by phases:

$$\hat{U}^\dagger\sigma_x\hat{U} = \begin{pmatrix} 0 & e^{i\varphi} \\ e^{-i\varphi} & 0 \end{pmatrix} = \cos\varphi\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sin\varphi\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = \cos\varphi\,\sigma_x - \sin\varphi\,\sigma_y.$$

Likewise $\hat{U}^\dagger\sigma_y\hat{U} = \begin{pmatrix} 0 & -ie^{i\varphi} \\ ie^{-i\varphi} & 0 \end{pmatrix} = \sin\varphi\,\sigma_x + \cos\varphi\,\sigma_y$, and σ_z is unchanged.

If $|\chi\rangle$ has spin expectation $\langle\hat{\mathbf{S}}\rangle = \mathbf{s}$, the rotated state $\hat{U}|\chi\rangle$ has $\langle\hat{S}_x\rangle' = \langle\chi|\hat{U}^\dagger\hat{S}_x\hat{U}|\chi\rangle = s_x\cos\varphi - s_y\sin\varphi$ and $\langle\hat{S}_y\rangle' = s_x\sin\varphi + s_y\cos\varphi$: the vector $\mathbf{s}$ rotated by φ about z, counterclockwise. So the expectation value transforms as an ordinary vector, even though the state has only two components and acquires a sign under a 2π rotation (Example 9.32). Operators with this property, $\hat{U}^\dagger\hat{V}_i\hat{U} = R_{ij}\hat{V}_j$, are called *vector operators*. $\hat{\mathbf{r}}$, $\hat{\mathbf{p}}$, $\hat{\mathbf{L}}$, and $\hat{\mathbf{S}}$ all are, and that is what the commutators $[\hat{J}_i, \hat{V}_j] = i\hbar\varepsilon_{ijk}\hat{V}_k$ of the exercises express infinitesimally.

Problem 17.4 Spin–orbit splitting of a p level

For an electron with $\ell = 1$, find the eigenvalues of $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$ and their degeneracies. Verify that the degeneracy-weighted average of the eigenvalues is zero, and explain why.

Solution. With $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$, $\hat{J}^2 = \hat{L}^2 + \hat{S}^2 + 2\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$, so

$$\hat{\mathbf{L}}\cdot\hat{\mathbf{S}} = \frac{1}{2}\left(\hat{J}^2 - \hat{L}^2 - \hat{S}^2\right) = \frac{\hbar^2}{2}\left[j(j+1) - \ell(\ell+1) - \frac{3}{4}\right]$$

on states of definite j. By Theorem 17.10, $\ell = 1$ and $s = \frac{1}{2}$ give $j = \frac{3}{2}$ (four states) and $j = \frac{1}{2}$ (two states). Then:

$$j = \tfrac{3}{2}: \ \frac{\hbar^2}{2}\left(\frac{15}{4} - 2 - \frac{3}{4}\right) = \frac{\hbar^2}{2}, \qquad j = \tfrac{1}{2}: \ \frac{\hbar^2}{2}\left(\frac{3}{4} - 2 - \frac{3}{4}\right) = -\hbar^2.$$

The weighted average is $\frac{4\cdot\frac{1}{2}+2\cdot(-1)}{6}\hbar^2 = 0$. This is the trace of $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$ over the six-dimensional space divided by six, and the trace can be computed in the product basis $|m_\ell, m_s\rangle$, where $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}} = \hat{L}_z\hat{S}_z + \frac{1}{2}(\hat{L}_+\hat{S}_- + \hat{L}_-\hat{S}_+)$. The ladder terms have no diagonal elements, and $\sum_{m_\ell,m_s} m_\ell m_s\hbar^2 = \big(\sum m_\ell\big)\big(\sum m_s\big)\hbar^2 = 0$.

So spin–orbit coupling splits a level without moving its center of gravity: the *Landé interval rule*. For the sodium $3p$ level, the splitting $\frac{3}{2}\hbar^2\times(\text{coupling})$ is 2.1 meV, which separates the yellow D lines at 589.0 and 589.6 nm.

Problem 17.5 The 21-centimeter line

In the hydrogen ground state, the magnetic interaction between the electron and proton spins is $\hat{H}' = \frac{A}{\hbar^2}\hat{\mathbf{S}}_e\cdot\hat{\mathbf{S}}_p$ with $A = 5.874\ \mu\text{eV}$. Find the splitting of the ground state, and the frequency and wavelength of the photon emitted in the transition.

Solution. As for $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$, write $\hat{\mathbf{S}}_e\cdot\hat{\mathbf{S}}_p = \frac{1}{2}\big(\hat{S}^2 - \hat{S}_e^2 - \hat{S}_p^2\big)$, where $\hat{\mathbf{S}}$ is the total spin. Two spins $\frac{1}{2}$ combine to $S=1$ (triplet, three states) or $S=0$ (singlet, one state), by Theorem 17.9. So

$$\hat{\mathbf{S}}_e\cdot\hat{\mathbf{S}}_p = \frac{\hbar^2}{2}\Big[S(S+1) - \frac{3}{4} - \frac{3}{4}\Big] = \begin{cases} +\frac{1}{4}\hbar^2, & S=1,\\ -\frac{3}{4}\hbar^2, & S=0. \end{cases}$$

The triplet lies at $+\frac{A}{4}$ and the singlet at $-\frac{3A}{4}$: the splitting is $A = 5.874\ \mu\text{eV}$. The photon has $\nu = \frac{A}{h} = \frac{5.874\times10^{-6}\ \text{eV}}{4.1357\times10^{-15}\ \text{eV s}} = 1420.4\ \text{MHz}$ and $\lambda = c/\nu = 21.1\ \text{cm}$.

The transition is extraordinarily slow, with a lifetime of about 10^7 years, because it is a magnetic dipole transition with a tiny energy. Yet interstellar hydrogen is so abundant that the line is easily detected. Radio astronomers use it to map the spiral arms of the Galaxy and, from its Doppler shifts, the flat rotation curves of galaxies that are key evidence for dark matter. The line's frequency is known to thirteen digits, and it was chosen as the unit of length and time on the plaques carried by the Pioneer spacecraft.

Problem 17.6 Three spins

Three spin-$\frac{1}{2}$ particles are combined. Find the possible values of the total spin S and their multiplicities, write the state $|S=\frac{3}{2}, M=\frac{1}{2}\rangle$, and find two orthonormal states with $S=\frac{1}{2}$, $M=\frac{1}{2}$.

Solution. Add in two steps. The first two spins give $S_{12} = 1$ or 0. Adding the third spin: $1\otimes\frac{1}{2}$ gives $\frac{3}{2}$ and $\frac{1}{2}$, and $0\otimes\frac{1}{2}$ gives $\frac{1}{2}$. So $S=\frac{3}{2}$ once and $S=\frac{1}{2}$ twice. Dimension check: $4+2+2 = 8 = 2^3$.

The state $|\frac{3}{2},\frac{3}{2}\rangle = |\uparrow\uparrow\uparrow\rangle$. Applying $\hat{S}_- = \hat{S}_{1-} + \hat{S}_{2-} + \hat{S}_{3-}$, each term flips one spin with coefficient $\hbar$, and $\hat{S}_-|\frac{3}{2},\frac{3}{2}\rangle = \hbar\sqrt{3}|\frac{3}{2},\frac{1}{2}\rangle$ by Theorem 17.3. So

$$\left|\tfrac{3}{2},\tfrac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\big(|\downarrow\uparrow\uparrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\uparrow\uparrow\downarrow\rangle\big).$$

The $M = \frac{1}{2}$ subspace is three-dimensional, spanned by the three states with one spin down. The $S = \frac{1}{2}$, $M = \frac{1}{2}$ states form its two-dimensional orthogonal complement. One natural basis:

$$\frac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\big)|\uparrow\rangle, \qquad \frac{1}{\sqrt{6}}\big(2|\uparrow\uparrow\downarrow\rangle - |\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle\big).$$

The first has spins 1 and 2 in a singlet ($S_{12} = 0$), and the second has $S_{12} = 1$. Both are orthogonal to $|\frac{3}{2},\frac{1}{2}\rangle$ (their coefficients sum to zero) and to each other, and both are annihilated by $\hat{S}_+$, as states with $M = S$ must be. The $S = \frac{1}{2}$ subspace is degenerate, and which basis is "right" depends on what else is measured. Exactly this structure, for three quarks, distinguishes the proton and neutron ($S = \frac{1}{2}$) from the Δ baryons ($S = \frac{3}{2}$).

EXERCISES

A. COMPUTATION

17.1 Verify $[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x$ directly from the definitions.

17.2 Show that $\hat{L}_z = -i\hbar\,\partial_\phi$ by writing $\hat{L}_z = -i\hbar(x\partial_y - y\partial_x)$ in spherical coordinates.

17.3 Show that $\hat{L}_+ = \hbar e^{i\phi}\big(\partial_\theta + i\cot\theta\,\partial_\phi\big)$, and verify $\hat{L}_+ Y_1^0 \propto Y_1^1$.

17.4 Construct the matrices $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_z$ for $j = \frac{3}{2}$.

17.5 Verify $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$ for the $j = 1$ matrices of Example 17.4.

17.6 A particle has $\ell = 1$ and is in the state $|1,1\rangle$. Find the probabilities of each result when L_x is measured.

17.7 In the state $|\ell, m\rangle$, compute $\langle L_x\rangle$, $\langle L_x^2\rangle$, and $\sigma_{L_x}\sigma_{L_y}$. Verify the uncertainty principle.

17.8 Find the eigenvalues and eigenvectors of $\hat{S}_y$.

17.9 A spin-$\frac{1}{2}$ particle is in the state $\chi = A\,\langle 3i, 4\rangle$. Normalize it, find $\langle\hat{\mathbf{S}}\rangle$, and find the direction $\hat{\mathbf{n}}$ along which the spin points.

17.10 An electron is prepared with spin up along $\hat{\mathbf{n}}$ at $\theta = 60^\circ$, $\phi = 0$. Find the probabilities of each result when S_x is measured.

17.11 Find the probability that a spin prepared along $+z$ is found along $+\hat{\mathbf{n}}$ for $\hat{\mathbf{n}}$ at 120° to z.

17.12 An electron spin starts in $|\uparrow\rangle$ in a field $\mathbf{B} = B\hat{\mathbf{x}}$. Find $|\chi(t)\rangle$ and the probability of finding spin down at time t.

17.13 Find the Larmor frequency of an electron ($g = 2.0023$) in Earth's field of $50\ \mu$T, and of a proton in a 3 T MRI magnet.

17.14 Compute $\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2$ on each triplet state, and verify $\hat{S}^2 = 2\hbar^2$ there.

17.15 Find the $J = \frac{3}{2}$ and $J = \frac{1}{2}$ states obtained by adding $\ell = 1$ and $s = \frac{1}{2}$, in terms of $|m_\ell, m_s\rangle$.

17.16 What values of total J arise from adding $j_1 = 2$ and $j_2 = \frac{3}{2}$? Verify the dimension count.

17.17 The hydrogen $2p$ level splits into $2p_{3/2}$ and $2p_{1/2}$. How many states are in each, and how many $2p$ states in total?

17.18 Compute the normal Zeeman splitting $\mu_B B$ in eV for $B = 1$ T, and the corresponding wavelength shift of the 656 nm line.

B. THEORY AND PROOF

17.19 Prove $[\hat{L}_i, \hat{x}_j] = i\hbar\varepsilon_{ijk}\hat{x}_k$ and $[\hat{L}_i, \hat{p}_j] = i\hbar\varepsilon_{ijk}\hat{p}_k$: $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ are vector operators.

17.20 Prove $[\hat{L}_i, \hat{r}^2] = 0$ and $[\hat{L}_i, \hat{p}^2] = 0$, and deduce that $\hat{L}_i$ commutes with any Hamiltonian $\frac{\hat{p}^2}{2m} + V(r)$.

17.21 Prove the Pauli identity $(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{b}) = (\mathbf{a}\cdot\mathbf{b})I + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b})$.

17.22 Use the Pauli identity to show $e^{-i\alpha\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}/2} = \cos\frac{\alpha}{2}\,I - i\sin\frac{\alpha}{2}\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}$, and that it rotates $\langle\boldsymbol{\sigma}\rangle$ by α about $\hat{\mathbf{n}}$.

17.23 Show that the singlet has the same form along every axis.

17.24 Prove that $\langle \hat{\mathbf{a}}\cdot\sigma_1\ \hat{\mathbf{b}}\cdot\sigma_2\rangle = -\hat{\mathbf{a}}\cdot\hat{\mathbf{b}}$ in the singlet state.

17.25 Prove that $\sum_J (2J+1) = (2j_1+1)(2j_2+1)$ when J runs from $|j_1 - j_2|$ to $j_1 + j_2$.

C. CHALLENGE

17.26 (Rabi oscillations.) A spin in a static field $B_0\hat{\mathbf{k}}$ is driven by a rotating field $B_1(\cos\omega t\,\hat{\mathbf{i}} - \sin\omega t\,\hat{\mathbf{j}})$. Transform to the frame rotating with the field and show that the probability of a spin flip is $\frac{\Omega^2}{\Omega^2+\Delta^2}\sin^2\left(\frac{1}{2}\sqrt{\Omega^2+\Delta^2}\,t\right)$, where $\Omega = \gamma B_1$ and $\Delta = \omega - \gamma B_0$.

17.27 Find the Clebsch–Gordan coefficients for adding $j_1 = 1$ and $j_2 = 1$ to $J = 0$, and show the state is $\frac{1}{\sqrt{3}}\big(|1,-1\rangle - |0,0\rangle + |-1,1\rangle\big)$.

17.28 Show that there is no 2×2 representation of $[\hat{J}_i,\hat{J}_j] = i\hbar\varepsilon_{ijk}\hat{J}_k$ other than $j = \frac{1}{2}$ (up to change of basis) and the trivial one.

17.29 Derive the Landé g-factor for a state $|j,m_j\rangle$ formed from ℓ and $s = \frac{1}{2}$, $g_J = 1 + \frac{j(j+1)+s(s+1)-\ell(\ell+1)}{2j(j+1)}$, using the projection $\langle\hat{\mathbf{S}}\rangle = \frac{\langle\hat{\mathbf{S}}\cdot\hat{\mathbf{J}}\rangle}{j(j+1)\hbar^2}\langle\hat{\mathbf{J}}\rangle$.

CHAPTER 18

Two-Level Systems, Identical Particles, and Entanglement

LEARNING OBJECTIVES

- Solve the general two-level system, and derive Rabi oscillations and two-flavor neutrino oscillations.
- Describe composite systems with tensor products, distinguish product from entangled states, and use the Schmidt decomposition.
- Use density matrices and the partial trace to describe subsystems, and explain decoherence.
- State the symmetrization postulate, and derive the Pauli exclusion principle and the exchange effect.
- Derive the CHSH inequality for local hidden-variable theories, and show that quantum mechanics violates it.
- Work with qubits and quantum gates, prove the no-cloning theorem, and analyze quantum teleportation.

The last chapter of this book is about the features of quantum mechanics that have no classical counterpart at all. When two systems interact, their joint state generally cannot be described by giving a state for each one separately. The two are *entangled*. When the two are identical particles, the theory requires a symmetry of the joint state that underlies the structure of every atom. In 1964 John Bell showed that the correlations of entangled states are stronger than any theory of pre-existing local properties allows. Experiments since have confirmed quantum mechanics and ruled such theories out, and the 2022 Nobel Prize in Physics was awarded for that work. Entanglement is now a technological resource, the foundation of quantum computing and quantum communication. Throughout, the mathematics is the finite-dimensional linear algebra of Chapter 9.

18.1 Two-Level Systems

Any quantum system in which only two states matter, such as a spin in a field, an atom driven near one transition, or a molecule with two equivalent configurations, has a 2×2 Hermitian Hamiltonian. Every such matrix can be written uniquely as

$$\hat{H} = E_0\, I + \mathbf{b} \cdot \boldsymbol{\sigma} = \begin{pmatrix} E_0 + b_z & b_x - i b_y \\ b_x + i b_y & E_0 - b_z \end{pmatrix},$$

with E_0 real and $\mathbf{b} \in \mathbb{R}^3$, because $\{I, \sigma_x, \sigma_y, \sigma_z\}$ is a basis for the 2×2 Hermitian matrices.

Theorem 18.1 (Two-level spectrum and dynamics). *The eigenvalues of $\hat{H}$ are $E_0 \pm |\mathbf{b}|$, with eigenvectors the spin states $|\pm\hat{\mathbf{b}}\rangle$ along $\hat{\mathbf{b}} = \mathbf{b}/|\mathbf{b}|$. The evolution is a precession of the Bloch vector about $\hat{\mathbf{b}}$ at angular frequency $\omega = 2|\mathbf{b}|/\hbar$.*

Proof. $(\mathbf{b} \cdot \boldsymbol{\sigma})^2 = |\mathbf{b}|^2 I$ by the Pauli identity, so the eigenvalues of $\mathbf{b} \cdot \boldsymbol{\sigma}$ are $\pm|\mathbf{b}|$, and its eigenvectors are those of Theorem 17.5. The Hamiltonian has the form of a spin in a field, with $-\gamma \mathbf{B} \cdot \hat{\mathbf{S}}$ replaced by $\mathbf{b} \cdot \boldsymbol{\sigma}$, so Theorem 17.7 applies with $\gamma B \to 2|\mathbf{b}|/\hbar$. ∎

Example 18.2 (Tunneling between two wells). The ammonia molecule NH_3 is a pyramid, with the nitrogen above or below the plane of the hydrogens: states $|L\rangle$ and $|R\rangle$, with equal energy E_0 by symmetry. The nitrogen can tunnel through the plane, which couples the states with amplitude $-A$: $\hat{H} = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}$. The stationary states are $\frac{1}{\sqrt{2}}(|L\rangle \pm |R\rangle)$ with energies $E_0 \mp A$. This is the degenerate perturbation theory of Chapter 15. A molecule started in $|L\rangle$ oscillates to $|R\rangle$ and back at angular frequency $2A/\hbar$, since

$$P_R(t) = \sin^2(At/\hbar).$$

For ammonia, $2A/h = 23.87$ GHz. The 1954 ammonia maser of Townes, which oscillated at this frequency, was the forerunner of the laser.

18.1.1 Rabi Oscillations

An atom with ground state $|g\rangle$ and excited state $|e\rangle$, separated by $\hbar\omega_0$, is driven by light of frequency ω. In a frame rotating at ω, and dropping rapidly oscillating terms (the *rotating-wave approximation*), the Hamiltonian becomes time-

independent:

$$\hat{H}_{\text{rot}} = \frac{\hbar}{2}\begin{pmatrix} -\Delta & \Omega \\ \Omega & \Delta \end{pmatrix}, \qquad \Delta = \omega - \omega_0,$$

where the *Rabi frequency* Ω is proportional to the field amplitude and the dipole matrix element. By Theorem 18.1, with $|\mathbf{b}| = \frac{\hbar}{2}\sqrt{\Omega^2 + \Delta^2}$, an atom starting in $|g\rangle$ is excited with probability

$$P_e(t) = \frac{\Omega^2}{\Omega^2 + \Delta^2}\sin^2\left(\frac{\sqrt{\Omega^2 + \Delta^2}}{2}\,t\right). \tag{18.1}$$

On resonance ($\Delta = 0$) the atom is driven completely to $|e\rangle$ after a *π-pulse*, of duration $t = \pi/\Omega$. Off resonance the transfer is incomplete and faster. This is the mechanism of NMR and MRI (Example 17.8), atomic clocks, and the single-qubit gates of a quantum computer.

18.1.2 NEUTRINO OSCILLATIONS

Neutrinos are produced in definite *flavor* states, such as $|\nu_e\rangle$ and $|\nu_\mu\rangle$. But the states of definite mass, which evolve with definite energy, are different: $|\nu_e\rangle = \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle$ and $|\nu_\mu\rangle = -\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle$. For ultrarelativistic neutrinos of momentum p, $E_i \approx pc + \frac{m_i^2 c^4}{2pc}$, so the mass states acquire a relative phase as they travel.

Theorem 18.3 (Two-flavor oscillation). *A neutrino produced as ν_e with energy E is detected as ν_μ after distance L with probability*

$$P(\nu_e \to \nu_\mu) = \sin^2 2\theta\,\sin^2\left(\frac{\Delta m^2 c^4 L}{4\hbar c E}\right), \qquad \Delta m^2 = m_2^2 - m_1^2.$$

Proof. After time $t = L/c$, the state is

$$|\psi\rangle = \cos\theta\, e^{-iE_1 t/\hbar}|\nu_1\rangle + \sin\theta\, e^{-iE_2 t/\hbar}|\nu_2\rangle.$$

Then $\langle\nu_\mu|\psi\rangle = \sin\theta\cos\theta\left(e^{-iE_2 t/\hbar} - e^{-iE_1 t/\hbar}\right)$, whose squared modulus is $\sin^2 2\theta\,\sin^2\left(\frac{(E_2 - E_1)t}{2\hbar}\right)$. Substitute $E_2 - E_1 = \frac{\Delta m^2 c^4}{2E}$ and $t = L/c$. ∎

Oscillations can occur only if $\Delta m^2 \neq 0$, so their observation, in solar and atmospheric neutrinos, proved that neutrinos have mass. That discovery earned the 2015 Nobel Prize. It is a purely quantum interference effect, visible over distances of hundreds of kilometers.

18.2 Composite Systems and Entanglement

Definition 18.4 (Tensor product). If systems A and B have Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ with orthonormal bases $\{|i\rangle\}$ and $\{|j\rangle\}$, the joint system has Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ with orthonormal basis $\{|i\rangle \otimes |j\rangle\}$, written $|ij\rangle$. Its dimension is $\dim \mathcal{H}_A \cdot \dim \mathcal{H}_B$. Inner products factor: $\langle ij|kl\rangle = \delta_{ik}\delta_{jl}$. An operator acting on A alone is $\hat{O}_A \otimes I$.

A general state is $|\Psi\rangle = \sum_{ij} c_{ij}|ij\rangle$. It is a *product state* if $|\Psi\rangle = |\alpha\rangle \otimes |\beta\rangle$ for some states of A and B, and *entangled* otherwise. The coefficients of a product state factor, $c_{ij} = a_i b_j$, which means the matrix $C = [c_{ij}]$ has rank one. Most states are entangled.

Example 18.5 (Bell states). The four *Bell states*

$$|\Phi^{\pm}\rangle = \tfrac{1}{\sqrt{2}}\big(|00\rangle \pm |11\rangle\big), \qquad |\Psi^{\pm}\rangle = \tfrac{1}{\sqrt{2}}\big(|01\rangle \pm |10\rangle\big)$$

form an orthonormal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$, and each is maximally entangled. $|\Psi^-\rangle$ is the spin singlet of Theorem 17.9. For $|\Phi^+\rangle$, $C = \frac{1}{\sqrt{2}} I$ has rank two, so it is not a product. Measuring the first qubit gives 0 or 1 with probability $\frac{1}{2}$ each, and afterwards the second qubit is certainly in the same state.

Theorem 18.6 (Schmidt decomposition). *Every state of $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written*

$$|\Psi\rangle = \sum_{k=1}^{r} \sqrt{p_k}\, |u_k\rangle \otimes |v_k\rangle,$$

with $\{|u_k\rangle\}$ and $\{|v_k\rangle\}$ orthonormal, $p_k > 0$, and $\sum p_k = 1$. The Schmidt rank *r equals the rank of C. The state is a product iff $r = 1$.*

Proof. Apply the singular value decomposition $C = U\Sigma V^{T}$ (a consequence of the spectral theorem applied to $C^{\dagger}C$). Here U and V are unitary and Σ has nonnegative diagonal entries s_k. Then $c_{ij} = \sum_k U_{ik} s_k V_{jk}$, so $|\Psi\rangle = \sum_k s_k \big(\sum_i U_{ik}|i\rangle\big) \otimes \big(\sum_j V_{jk}|j\rangle\big)$. Normalization gives $\sum s_k^2 = 1$; set $p_k = s_k^2$. ∎

18.2.1 Density Matrices and Subsystems

If A and B are entangled, A alone is not described by any state vector. It is described by a *density matrix*, which also describes ensembles prepared with classi-

cal uncertainty.

Definition 18.7 (Density matrix). A *density matrix* is a positive semidefinite Hermitian operator $\hat{\rho}$ with $\operatorname{tr}\hat{\rho} = 1$. The expectation value of an observable is $\langle A\rangle = \operatorname{tr}(\hat{\rho}\hat{A})$. A *pure state* has $\hat{\rho} = |\psi\rangle\langle\psi|$; otherwise the state is *mixed*, with $\operatorname{tr}\hat{\rho}^2 < 1$.

Definition 18.8 (Partial trace). The *reduced density matrix* of A is $\hat{\rho}_A = \operatorname{tr}_B \hat{\rho} = \sum_j \langle j|_B \hat{\rho}\, |j\rangle_B$. It gives the correct expectation value of every observable of A alone: $\operatorname{tr}(\hat{\rho}(\hat{O}_A \otimes I)) = \operatorname{tr}(\hat{\rho}_A \hat{O}_A)$.

By Theorem 18.6, $\hat{\rho}_A = \sum_k p_k |u_k\rangle\langle u_k|$. If $|\Psi\rangle$ is entangled, $\hat{\rho}_A$ is mixed. For a Bell state, $\hat{\rho}_A = \frac{1}{2}I$, the maximally mixed state: the first qubit alone is completely random, with no direction on the Bloch sphere. Every bit of information is in the correlations. The *entanglement entropy* $S = -\sum p_k \log_2 p_k$ is 0 for product states and 1 bit for Bell states.

18.2.2 DECOHERENCE AND WHICH-PATH INFORMATION

Return to the double slit (Chapter 12). If a photon scatters off the electron at slit i, the joint state becomes

$$|\Psi\rangle = \tfrac{1}{\sqrt{2}}\big(|\psi_1\rangle|\gamma_1\rangle + |\psi_2\rangle|\gamma_2\rangle\big),$$

where $|\gamma_i\rangle$ is the photon's state. The probability density on the screen is computed from the reduced density matrix of the electron:

$$P(y) = \tfrac{1}{2}\big(|\psi_1(y)|^2 + |\psi_2(y)|^2\big) + \operatorname{Re}\big(\psi_1^*(y)\psi_2(y)\langle\gamma_1|\gamma_2\rangle\big).$$

The interference term is multiplied by the overlap $\langle\gamma_1|\gamma_2\rangle$. If the photon states are orthogonal, so that the photon records which slit was used, interference vanishes completely. If they are identical, interference is complete. In between, it is partial. This process, the leaking of phase information into the environment through entanglement, is *decoherence*. It explains why quantum superpositions of macroscopic objects are never seen: they become entangled with their surroundings almost instantly.

18.3 IDENTICAL PARTICLES

Classically, identical particles can be distinguished by following their trajectories. In quantum mechanics there are no trajectories, and two electrons are truly in-

distinguishable. Let $\hat{P}$ exchange the two particles: $\hat{P}\psi(x_1, x_2) = \psi(x_2, x_1)$, with spin labels exchanged too. If the particles are identical, no measurement can tell ψ from $\hat{P}\psi$, so they represent the same state, and $\hat{P}\psi = \lambda\psi$. Since $\hat{P}^2 = I$, $\lambda = \pm 1$.

Definition 18.9 (Symmetrization postulate). The state of identical *bosons* is symmetric under exchange of any two particles; the state of identical *fermions* is antisymmetric.

The *spin–statistics theorem* of relativistic quantum field theory proves that particles with integer spin are bosons and those with half-integer spin are fermions. Electrons, protons, neutrons, and quarks are fermions; photons, gluons, and helium-4 atoms are bosons.

If two particles occupy single-particle states ψ_a and ψ_b, the allowed two-particle wavefunctions are

$$\psi_{\pm}(x_1, x_2) = \frac{1}{\sqrt{2}}\big[\psi_a(x_1)\psi_b(x_2) \pm \psi_b(x_1)\psi_a(x_2)\big].$$

Theorem 18.10 (Pauli exclusion principle). *Two identical fermions cannot occupy the same single-particle state.*

Proof. If $\psi_a = \psi_b$, then $\psi_- \equiv 0$. ∎

For N fermions in states $\psi_1, \dots, \psi_N$, the antisymmetric state is the *Slater determinant* $\frac{1}{\sqrt{N!}}\det\big[\psi_i(x_j)\big]$, which vanishes when two rows coincide. Including spin, each spatial orbital holds at most two electrons with opposite spins. That is the rule that builds the periodic table (Chapter 16). It also explains why matter is rigid and does not collapse, why white dwarfs are supported against gravity, and why metals conduct.

Theorem 18.11 (Exchange effect). *For two particles in orthonormal states ψ_a, ψ_b in one dimension,*

$$\big\langle (x_1 - x_2)^2 \big\rangle_{\pm} = \langle x^2\rangle_a + \langle x^2\rangle_b - 2\langle x\rangle_a\langle x\rangle_b \mp 2|\langle x\rangle_{ab}|^2,$$

where $\langle x\rangle_{ab} = \int x\,\psi_a^*\psi_b\,dx$. The first three terms are the result for distinguishable particles.

Proof. Expand $\langle x_1^2 + x_2^2 - 2x_1x_2\rangle$ in $\psi_\pm$. The cross terms in $\langle x_1x_2\rangle$ contribute $\pm|\langle x\rangle_{ab}|^2$. ∎

Symmetric (bosonic) spatial wavefunctions bring particles closer together on average, and antisymmetric (fermionic) ones push them apart. No force is involved. This *exchange effect* is a consequence of symmetry alone. Combined with the Coulomb repulsion, it makes electrons with parallel spins (antisymmetric spatial state) avoid each other and lower their repulsion energy. This is Hund's rule in atoms, and the origin of ferromagnetism.

18.4 Bell's Theorem

In 1935 Einstein, Podolsky, and Rosen argued that quantum mechanics is incomplete. Consider two particles in the singlet state, flown far apart. Measuring S_z on particle 1 predicts with certainty the result of measuring S_z on particle 2. Measuring S_x on particle 1 would predict S_x on particle 2 with equal certainty. Since, they argued, no influence can travel faster than light, particle 2 must have had definite values of both S_z and S_x all along, which the quantum state does not describe. The natural completion is a theory of *local hidden variables*. Each pair carries a variable λ that predetermines every result, and each result depends only on the local setting and λ.

In 1964 Bell showed that such theories make predictions that differ from quantum mechanics. We give the version of Clauser, Horne, Shimony, and Holt (CHSH). Alice measures along $\hat{\mathbf{a}}$ or $\hat{\mathbf{a}}'$ and Bob along $\hat{\mathbf{b}}$ or $\hat{\mathbf{b}}'$; each records ± 1. Let $E(\hat{\mathbf{a}}, \hat{\mathbf{b}})$ be the average of the product of their results.

Theorem 18.12 (CHSH inequality). *In any local hidden-variable theory,*

$$|S| \le 2, \qquad S = E(\hat{\mathbf{a}}, \hat{\mathbf{b}}) + E(\hat{\mathbf{a}}, \hat{\mathbf{b}}') + E(\hat{\mathbf{a}}', \hat{\mathbf{b}}) - E(\hat{\mathbf{a}}', \hat{\mathbf{b}}').$$

Proof. For a given λ, the outcomes are $A, A' \in \{\pm 1\}$ for Alice's two settings and $B, B' \in \{\pm 1\}$ for Bob's, all fixed in advance. Then

$$AB + AB' + A'B - A'B' = A(B + B') + A'(B - B').$$

One of $B + B'$ and $B - B'$ is 0 and the other is ± 2, so this quantity is ± 2. Averaging over any distribution of λ gives a number in $[-2, 2]$. ∎

Theorem 18.13 (Quantum violation). *For the singlet state, $E(\hat{\mathbf{a}}, \hat{\mathbf{b}}) = -\hat{\mathbf{a}} \cdot \hat{\mathbf{b}}$. If the four directions lie in one plane at angles $\hat{\mathbf{a}} = 0^\circ$, $\hat{\mathbf{a}}' = 90^\circ$, $\hat{\mathbf{b}} = 45^\circ$, and $\hat{\mathbf{b}}' = -45^\circ$, then $S = -2\sqrt{2}$, so $|S| = 2\sqrt{2} \approx 2.83$.*

Proof. The correlation formula is Exercise 17.24. The angles between the settings are 45° for $(\hat{\mathbf{a}}, \hat{\mathbf{b}})$, 45° for $(\hat{\mathbf{a}}, \hat{\mathbf{b}}')$, 45° for $(\hat{\mathbf{a}}', \hat{\mathbf{b}})$, and 135° for $(\hat{\mathbf{a}}', \hat{\mathbf{b}}')$. So

$$S = -\cos 45^\circ - \cos 45^\circ - \cos 45^\circ + \cos 135^\circ = -\tfrac{3}{\sqrt{2}} - \tfrac{1}{\sqrt{2}} = -2\sqrt{2}. \qquad \blacksquare$$

Quantum mechanics predicts $|S| = 2\sqrt{2}$, exceeding the local bound of 2. (Tsirelson proved that $2\sqrt{2}$ is the largest possible quantum value.) Experiments by Aspect and collaborators in 1982, and loophole-free tests in 2015, measured S in agreement with quantum mechanics and in clear violation of the CHSH bound. The conclusion is inescapable: no theory in which measurement results are determined by local pre-existing properties can reproduce nature. Entanglement cannot be used to send signals faster than light, since each party's local results are random whatever the other does. But the correlations themselves are *nonlocal* in Bell's sense.

18.5 Qubits and Quantum Information

A *qubit* is any two-level quantum system, with computational basis $|0\rangle, |1\rangle$. Its pure states are the points of the Bloch sphere. A register of n qubits has a 2^n-dimensional state space, and that exponential growth is the source of both the difficulty of simulating quantum systems and the power of quantum computers.

Gates are unitary operations. Single-qubit gates include the Pauli matrices $X = \sigma_x$, $Y = \sigma_y$, $Z = \sigma_z$, the *Hadamard* gate $H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, which maps $|0\rangle \mapsto |+\rangle$ and $|1\rangle \mapsto |-\rangle$, and phase rotations. The two-qubit *controlled-NOT* gate flips the second qubit if the first is $|1\rangle$: $\mathrm{CNOT}|a, b\rangle = |a, a \oplus b\rangle$. Applying H to the first qubit and then CNOT turns $|00\rangle$ into the Bell state $|\Phi^+\rangle$, creating entanglement.

> **Theorem 18.14** (No-cloning). *There is no unitary U and fixed state $|e\rangle$ such that $U(|\psi\rangle \otimes |e\rangle) = |\psi\rangle \otimes |\psi\rangle$ for every state $|\psi\rangle$.*

Proof. If U cloned both $|\psi\rangle$ and $|\phi\rangle$, then taking inner products of the two equations and using unitarity gives $\langle\psi|\phi\rangle = \langle\psi|\phi\rangle^2$. So $\langle\psi|\phi\rangle$ is 0 or 1: only mutually orthogonal (or identical) states can be cloned by one device. $\blacksquare$

An unknown quantum state cannot be copied. This is the security basis of quantum key distribution: an eavesdropper cannot copy the transmitted qubits without disturbing them.

Theorem 18.15 (Teleportation). *Alice holds an unknown qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ and shares $|\Phi^+\rangle$ with Bob. If she measures her two qubits in the Bell basis and tells Bob the result (two classical bits), Bob can apply one of I, X, Z, XZ to his qubit to obtain $|\psi\rangle$ exactly.*

Proof. Expanding the three-qubit state $|\psi\rangle \otimes |\Phi^+\rangle$ in the Bell basis of Alice's two qubits gives

$$\tfrac{1}{2}\Big[|\Phi^+\rangle(\alpha|0\rangle + \beta|1\rangle) + |\Phi^-\rangle(\alpha|0\rangle - \beta|1\rangle) \\ + |\Psi^+\rangle(\alpha|1\rangle + \beta|0\rangle) + |\Psi^-\rangle(\alpha|1\rangle - \beta|0\rangle)\Big].$$

Each of Alice's four outcomes occurs with probability $\frac{1}{4}$, and leaves Bob's qubit in $|\psi\rangle$, $Z|\psi\rangle$, $X|\psi\rangle$, or $XZ|\psi\rangle$ up to a sign. The corresponding inverse restores $|\psi\rangle$. ∎

Teleportation transmits a quantum state using entanglement and classical communication. It does not violate no-cloning, because Alice's original is destroyed by her measurement. Nor does it violate relativity, because Bob must wait for her classical message. It has been demonstrated over more than 1000 km, using a satellite.

LOOKING AHEAD. Quantum algorithms exploit superposition, interference, and entanglement. Shor's algorithm factors integers in polynomial time using the quantum Fourier transform, the unitary version of the discrete Fourier transform of Chapter 11. Grover's algorithm searches an unstructured list of N items in about $\sqrt{N}$ steps. Simulating molecules and materials, which is the problem Feynman identified in 1982, remains the most natural application. Building fault-tolerant quantum computers is one of the defining engineering projects of this century. Every one of its components is a two-level system described by this chapter.

18.6 Worked Problems

The problems below cover a driven two-level system, the entanglement structure of three qubits, an exchange effect that underlies magnetism, the Bell inequality for partially entangled states, and three protocols of quantum information: measurement as entanglement, the impossibility of copying, and superdense coding.

Problem 18.1 Power broadening

For the driven atom of (18.1), find the time-averaged excitation probability as a function of the detuning Δ. What is the full width at half maximum of the resulting line?

Solution. The time average of $\sin^2$ is $\frac{1}{2}$, so

$$\overline{P_e} = \frac{1}{2}\,\frac{\Omega^2}{\Omega^2+\Delta^2}.$$

As a function of Δ this is a Lorentzian, with peak value $\frac{1}{2}$ at resonance and half that value where $\Delta^2 = \Omega^2$. So the full width at half maximum is 2Ω. (A numerical average of $P_e(t)$ over 2000 time units, with $\Omega = 1.3$ and $\Delta = 0.7$, agrees with the formula to 10^{-3}.)

The width is proportional to the Rabi frequency, and so to the amplitude of the driving field: driving harder broadens the line. This *power broadening* is why precision spectroscopy uses weak fields, and why the narrowest lines are probed with long, gentle pulses. In practice the atom also decays spontaneously, at rate Γ, and the line then has width $\Gamma\sqrt{1+2\Omega^2/\Gamma^2}$. For weak driving the width is the natural width Γ, fixed by the excited state's lifetime through the energy–time uncertainty relation.

Problem 18.2 The W state and the GHZ state

Let $|W\rangle = \frac{1}{\sqrt{3}}\big(|001\rangle + |010\rangle + |100\rangle\big)$ and $|GHZ\rangle = \frac{1}{\sqrt{2}}\big(|000\rangle + |111\rangle\big)$. For each, find the reduced density matrix of one qubit and of two qubits. If the third qubit is measured in the basis $\{|0\rangle, |1\rangle\}$, what states can the first two be left in?

Solution. *W.* Tracing out qubits 2 and 3: qubit 1 is $|1\rangle$ in one of the three terms and $|0\rangle$ in two, and the cross terms vanish because the other qubits' states are orthogonal. So $\hat\rho_1 = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1|$. Tracing out qubit 3 only, group the terms by its value: $|W\rangle = \frac{1}{\sqrt{3}}|00\rangle|1\rangle + \sqrt{\frac{2}{3}}\,|\Psi^+\rangle|0\rangle$, where $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$. So

$$\hat\rho_{12} = \frac{1}{3}|00\rangle\langle 00| + \frac{2}{3}|\Psi^+\rangle\langle\Psi^+|.$$

Measuring qubit 3 gives 1 with probability $\frac{1}{3}$, leaving $|00\rangle$, and 0 with probability $\frac{2}{3}$, leaving the Bell state $|\Psi^+\rangle$ (Example 18.5).

GHZ. $\hat\rho_1 = \frac{1}{2}I$, maximally mixed. And $\hat\rho_{12} = \frac{1}{2}\big(|00\rangle\langle 00| + |11\rangle\langle 11|\big)$, a classical mixture of two product states, with no entanglement at all. Measuring qubit 3 leaves $|00\rangle$ or $|11\rangle$, product states, with probability $\frac{1}{2}$ each.

The two states are entangled in inequivalent ways. The GHZ state is the extreme case of collective entanglement: lose one qubit and the other two retain only classical correlations. The W state's entanglement is shared pairwise and is robust: with probability $\frac{2}{3}$, the loss or measurement of one qubit leaves the others maximally entangled. No local operations can convert one type into the other, a result of Dür, Vidal, and Cirac (2000).

Problem 18.3 Exchange splitting from a contact interaction

Two electrons in a harmonic trap occupy the oscillator states ψ_0 and ψ_1, and interact through $\hat{V} = g\,\delta(x_1 - x_2)$ with $g > 0$. Using the singlet and triplet combinations of Theorem 18.11, find the first-order energy shift of each spin state.

Solution. The total state must be antisymmetric. The spin singlet is antisymmetric, so it pairs with the symmetric spatial state $\psi_S = \frac{1}{\sqrt{2}}\big(\psi_0(x_1)\psi_1(x_2) + \psi_1(x_1)\psi_0(x_2)\big)$. The three triplet states are symmetric, so they pair with the antisymmetric ψ_A, with a minus sign. The interaction acts only where $x_1 = x_2$:

$$\langle g\,\delta(x_1 - x_2)\rangle = g\int |\psi(x,x)|^2\,dx.$$

For ψ_A, $\psi_A(x,x) = 0$, so the triplet is unshifted. For ψ_S, $\psi_S(x,x) = \sqrt{2}\,\psi_0(x)\psi_1(x)$, so the shift is $2g\int \psi_0^2\psi_1^2\,dx$. With $x_0 = \sqrt{\hbar/m\omega}$, $\psi_0 = \pi^{-1/4}x_0^{-1/2}e^{-x^2/2x_0^2}$ and $\psi_1 = \sqrt{2}\,\frac{x}{x_0}\psi_0$, so

$$\int \psi_0^2\psi_1^2\,dx = \frac{2}{\pi x_0^4}\int x^2 e^{-2x^2/x_0^2}dx = \frac{2}{\pi x_0^4}\cdot\frac{\sqrt{\pi}}{2}\Big(\frac{x_0^2}{2}\Big)^{3/2} = \frac{1}{2\sqrt{2\pi}\,x_0},$$

$$\Delta E_{\text{singlet}} = \frac{g}{\sqrt{2\pi}\,x_0}.$$

The interaction is spin-independent, yet the energy depends on the spin. The Pauli principle ties spin symmetry to spatial symmetry, and spatial symmetry decides whether the particles can meet. With repulsion ($g > 0$) the triplet, with aligned spins, is lower. This is the mechanism of Hund's first rule in atoms and of ferromagnetism in iron: the effective spin–spin coupling, the *exchange interaction*, is electrostatic in origin and is far stronger than the magnetic interaction between the spins.

Problem 18.4 Bell violation for partially entangled states

For $|\psi_\alpha\rangle = \cos\alpha|00\rangle + \sin\alpha|11\rangle$ with $0 < \alpha < \frac{\pi}{2}$, show that a suitable choice of measurement directions gives the CHSH value $S = 2\sqrt{1 + \sin^2 2\alpha}$. Conclude that every entangled pure state of this form violates the CHSH inequality.

Solution. Take all directions in the xz-plane: $\hat{\mathbf{a}} \cdot \boldsymbol{\sigma} = \cos a\, \sigma_z + \sin a\, \sigma_x$, and similarly for b. Compute the correlations of $|\psi_\alpha\rangle$:

$$\langle \sigma_z \otimes \sigma_z \rangle = 1, \qquad \langle \sigma_x \otimes \sigma_x \rangle = 2 \sin\alpha \cos\alpha = \sin 2\alpha,$$
$$\langle \sigma_z \otimes \sigma_x \rangle = \langle \sigma_x \otimes \sigma_z \rangle = 0.$$

(For example, $\sigma_x \otimes \sigma_x$ maps $|00\rangle \leftrightarrow |11\rangle$, so its expectation is $2\cos\alpha \sin\alpha$.) Hence $E(a,b) = \cos a \cos b + \sin 2\alpha \, \sin a \sin b$. Choose $a = 0$, $a' = \frac{\pi}{2}$, and $b = -b' = \beta$:

$$S = E(a,b) + E(a,b') + E(a',b) - E(a',b') = 2\cos\beta + 2\sin 2\alpha \sin\beta.$$

Maximizing over β gives $S = 2\sqrt{1 + \sin^2 2\alpha}$, at $\tan\beta = \sin 2\alpha$. This exceeds 2 whenever $\sin 2\alpha \neq 0$, that is, whenever the state is entangled. For $\alpha = \frac{\pi}{4}$ it is Tsirelson's bound $2\sqrt{2}$ (Theorem 18.13). A numerical maximization over all measurement directions confirms that this is the largest possible value for each α (for example 2.0391 at $\alpha = 0.1$).

So every entangled pure two-qubit state violates a Bell inequality (Gisin's theorem, 1991). For mixed states the situation is subtler: Werner found entangled mixed states whose correlations admit a local hidden-variable model. Entanglement and nonlocality are related, but they are not the same thing.

Problem 18.5 Measurement as entanglement

A qubit in the state $a|0\rangle + b|1\rangle$ interacts with a "pointer" qubit in $|0\rangle$ through the CNOT gate, with the qubit as control. Find the joint state, and the reduced density matrix of the qubit. What happened to the interference between $|0\rangle$ and $|1\rangle$?

Solution. CNOT flips the target when the control is $|1\rangle$, so

$$\big(a|0\rangle + b|1\rangle\big)|0\rangle \;\longmapsto\; a|0\rangle|0\rangle + b|1\rangle|1\rangle.$$

The pointer now records the qubit's value, and the two are entangled unless a or b is 0. Tracing out the pointer,

$$\hat{\rho} = |a|^2 |0\rangle\langle 0| + |b|^2 |1\rangle\langle 1|,$$

because the pointer states are orthogonal. Before the interaction, $\hat{\rho}$ had off-diagonal elements ab^* and a^*b, which carry the relative phase and produce interference (Chapter 12, and the which-path exercise of this chapter). After it, those elements are gone. Measurement statistics of the qubit alone are now those of a classical mixture: 0 with probability $|a|^2$, 1 with probability $|b|^2$.

Nothing in this process is irreversible, and no collapse occurred: applying CNOT again disentangles the pair and restores the interference. But if the pointer is a macroscopic object, or is coupled to an environment of many particles, undoing the correlation becomes impossible in practice. This is *decoherence*. It explains why interference vanishes when which-path information exists anywhere, and why macroscopic superpositions are never seen. It does not, by itself, explain why one outcome occurs, which is the residue of the measurement problem.

Problem 18.6 Why CNOT cannot clone

CNOT copies the basis states: $|0\rangle|0\rangle \mapsto |0\rangle|0\rangle$ and $|1\rangle|0\rangle \mapsto |1\rangle|1\rangle$. Apply it to $|\psi\rangle|0\rangle$ with $|\psi\rangle = a|0\rangle + b|1\rangle$. Find the fidelity $F = \langle\psi|\hat{\rho}_2|\psi\rangle$ of the second qubit with the original, and its average over all $|\psi\rangle$ on the Bloch sphere.

Solution. By the previous problem, the output is $a|00\rangle + b|11\rangle$, not $|\psi\rangle|\psi\rangle = a^2|00\rangle + ab|01\rangle + ab|10\rangle + b^2|11\rangle$. The two agree only when $ab = 0$, as the no-cloning theorem (Theorem 18.14) requires. The "copy" has $\hat{\rho}_2 = |a|^2|0\rangle\langle 0| + |b|^2|1\rangle\langle 1|$, so

$$F = \langle\psi|\hat{\rho}_2|\psi\rangle = |a|^4 + |b|^4.$$

With $a = \cos\frac{\theta}{2}$ and $|b| = \sin\frac{\theta}{2}$ on the Bloch sphere, $F = \cos^4\frac{\theta}{2} + \sin^4\frac{\theta}{2} = 1 - \frac{1}{2}\sin^2\theta$. Averaging over the sphere, with weight $\frac{1}{2}\sin\theta\,d\theta$,

$$\overline{F} = 1 - \frac{1}{2}\cdot\frac{1}{2}\int_0^\pi \sin^3\theta\,d\theta = 1 - \frac{1}{4}\cdot\frac{4}{3} = \frac{2}{3}.$$

The copy is perfect at the poles and worst on the equator, where $F = \frac{1}{2}$, no better than a random guess. The optimal universal quantum cloner, found by Bužek and Hillery in 1996, reaches $\overline{F} = \frac{5}{6}$, still short of 1. The impossibility of perfect copying is what makes quantum key distribution secure: an eavesdropper who copies the transmitted qubits inevitably disturbs them.

Problem 18.7 Superdense coding

Alice and Bob share the Bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, Alice holding the first qubit. Show that by applying one of I, X, Z, XZ to her qubit and sending it to Bob, Alice can transmit two classical bits, though she sends only one qubit.

Solution. Alice's four operations, acting on the first qubit, produce

$$I : \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = |\Phi^+\rangle, \qquad X : \tfrac{1}{\sqrt{2}}(|10\rangle + |01\rangle) = |\Psi^+\rangle,$$
$$Z : \tfrac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = |\Phi^-\rangle, \qquad XZ : \tfrac{1}{\sqrt{2}}(|10\rangle - |01\rangle) = -|\Psi^-\rangle.$$

These are the four Bell states, which are orthonormal (Example 18.5), up to a global phase. After receiving Alice's qubit, Bob holds both, and can measure in the Bell basis: apply CNOT and then H to the first qubit, which maps the four Bell states to the four basis states $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ in some order (this runs the Bell-state circuit of Exercise 18.15 backward), and read off two bits. So the four messages $00, 01, 10, 11$ are perfectly distinguishable.

This does not violate the limit that one qubit carries at most one bit of accessible information (Holevo's bound), because the pre-shared entanglement is a resource, used up in the process. Superdense coding is the dual of teleportation (Theorem 18.15), where one shared Bell pair and two classical bits transmit one qubit. Together they show that entanglement, qubits, and classical bits are interconvertible resources, the starting point of quantum information theory.

EXERCISES

A. COMPUTATION

18.1 Write $\hat{H} = \begin{pmatrix} 3 & 1-i \\ 1+i & 1 \end{pmatrix}$ as $E_0 I + \mathbf{b}\cdot\boldsymbol{\sigma}$, find its eigenvalues, and find the precession frequency.

18.2 For the ammonia model, find $|\psi(t)\rangle$ starting from $|L\rangle$, and the time to transfer completely to $|R\rangle$.

18.3 An atom is driven with Rabi frequency $\Omega = 2\pi \times 1$ MHz and detuning $\Delta = 2\pi \times 1$ MHz. Find the maximum excitation probability and the time at which it first occurs.

18.4 Find the length of a π-pulse for $\Omega = 2\pi \times 50$ kHz, and the state after a $\pi/2$-pulse starting from $|g\rangle$.

18.5 For atmospheric neutrinos with $\Delta m^2 c^4 = 2.5 \times 10^{-3}$ eV^2 and $E = 1$ GeV, find the oscillation length $L_{\text{osc}} = \frac{4\pi\hbar c E}{\Delta m^2 c^4}$ in km.

18.6 Determine which of these are entangled: $\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$, $\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle - |11\rangle)$, $\frac{1}{\sqrt{2}}(|01\rangle + |11\rangle)$.

18.7 Find the Schmidt decomposition and entanglement entropy of $\frac{1}{\sqrt{3}}(|00\rangle + |01\rangle + |11\rangle)$.

18.8 Compute the reduced density matrix of the first qubit for each Bell state.

18.9 For $\hat{\rho} = \frac{1}{2}\begin{pmatrix} 1 & r \\ r^* & 1 \end{pmatrix}$ with $|r| \le 1$, find $\operatorname{tr}\hat{\rho}^2$ and the Bloch vector. When is the state pure?

18.10 In the double slit with partial which-path information, $\langle\gamma_1|\gamma_2\rangle = 0.6$. Find the fringe visibility $V = \frac{P_{\max}-P_{\min}}{P_{\max}+P_{\min}}$, assuming $|\psi_1| = |\psi_2|$.

18.11 Two noninteracting particles of mass m are in an infinite well of width a, in states ψ_1 and ψ_2. Find the energy and $\langle(x_1 - x_2)^2\rangle$ for distinguishable particles, bosons, and fermions.

18.12 Find the ground-state energy and degeneracy of five noninteracting spin-$\frac{1}{2}$ fermions in a one-dimensional infinite well.

18.13 Verify that $E(\hat{\mathbf{a}}, \hat{\mathbf{b}}) = -\cos\gamma$ in the singlet by computing $\langle\Psi^-|\sigma_z \otimes (\cos\gamma\,\sigma_z + \sin\gamma\,\sigma_x)|\Psi^-\rangle$.

18.14 Show that the angles $\hat{\mathbf{a}} = 0$, $\hat{\mathbf{a}}' = 90^\circ$, $\hat{\mathbf{b}} = 45^\circ$, $\hat{\mathbf{b}}' = -45^\circ$ give $|S| = 2\sqrt{2}$, and find S for $\hat{\mathbf{b}}' = 135^\circ$.

18.15 Compute the matrix of CNOT in the basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, and verify that $\mathrm{CNOT}(H \otimes I)|00\rangle = |\Phi^+\rangle$.

18.16 Compute HZH and HXH, and interpret the Hadamard as a change of basis.

B. THEORY AND PROOF

18.17 Prove that $\{I, \sigma_x, \sigma_y, \sigma_z\}$ is a basis for the 2×2 Hermitian matrices over $\mathbb{R}$, with $\frac{1}{2}\operatorname{tr}(\sigma_i\sigma_j) = \delta_{ij}$.

18.18 Prove that a pure state $|\Psi\rangle$ of two qubits is a product state iff $\det C = 0$, where C is its 2×2 coefficient matrix.

18.19 Prove that $\operatorname{tr}\hat{\rho}^2 \le 1$ for every density matrix, with equality iff the state is pure.

18.20 Prove that local unitaries $U_A \otimes U_B$ cannot change the Schmidt coefficients, so they cannot create or destroy entanglement.

18.21 Prove that the exchange operator $\hat{P}$ commutes with any Hamiltonian symmetric in the two particles, so the symmetry of a state is preserved in time.

18.22 Show that no local operation by Bob changes Alice's reduced density matrix (no signaling).

18.23 Complete the algebra in the proof of Theorem 18.15.

C. CHALLENGE

18.24 (Helium.) Using the hydrogen-like orbitals with $Z=2$, the noninteracting ground-state energy of helium is -8×13.6 eV. Treat the electron repulsion as a first-order perturbation, $\langle \frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|}\rangle = \frac{5}{4}Z\times 13.6$ eV, and compare $E\approx -74.8$ eV with the measured -79.0 eV. Explain why the spatial state must be symmetric and the spin state a singlet.

18.25 (GHZ.) For $|GHZ\rangle=\frac{1}{\sqrt{2}}(|000\rangle+|111\rangle)$, compute the expectation values of XXX, XYY, YXY, and YYX. Show that no assignment of predetermined values ±1 to X_i and Y_i can reproduce all four. This is a Bell theorem without inequalities.

18.26 (Deutsch's algorithm.) A function $f:\{0,1\}\to\{0,1\}$ is accessed through the unitary $U_f|x,y\rangle=|x,y\oplus f(x)\rangle$. Show that preparing $|0\rangle|1\rangle$, applying $H\otimes H$, then U_f, then $H\otimes I$, and measuring the first qubit determines whether $f(0)=f(1)$ with one query, which is impossible classically.

18.27 (Quantum Fourier transform.) Show that the $N\times N$ matrix $F_{jk}=\frac{1}{\sqrt{N}}e^{2\pi ijk/N}$ is unitary, and write F for $N=4$. Show that F^2 is the reversal permutation $|j\rangle\mapsto|-j \bmod N\rangle$.

APPENDIX A

PHYSICAL CONSTANTS AND UNITS

Values are the CODATA 2022 recommended values. Since the 2019 redefinition of the SI, the constants c, h, e, k_B, and N_A have exact defined values, and the kilogram, ampere, kelvin, and mole are defined through them. The other constants are measured, and the digits shown are accurate to the last place given. Calculations in this book use as many digits as the data of each problem justify, usually three or four.

DEFINING CONSTANTS (EXACT)

Quantity	Symbol	Value	Unit
Speed of light in vacuum	c	299 792 458	$\mathrm{m\,s^{-1}}$
Planck constant	h	$6.626\,070\,15 \times 10^{-34}$	J s
Reduced Planck constant	$\hbar = h/2\pi$	$1.054\,571\,817\ldots \times 10^{-34}$	J s
Elementary charge	e	$1.602\,176\,634 \times 10^{-19}$	C
Boltzmann constant	k_B	$1.380\,649 \times 10^{-23}$	$\mathrm{J\,K^{-1}}$
Avogadro constant	N_A	$6.022\,140\,76 \times 10^{23}$	$\mathrm{mol^{-1}}$

MEASURED CONSTANTS

Quantity	Symbol	Value	Unit
Electron mass	m_e	$9.109\,383\,71 \times 10^{-31}$	kg
Proton mass	m_p	$1.672\,621\,93 \times 10^{-27}$	kg
Neutron mass	m_n	$1.674\,927\,50 \times 10^{-27}$	kg
Atomic mass constant	m_u	$1.660\,539\,07 \times 10^{-27}$	kg
Vacuum permittivity	ε_0	$8.854\,187\,82 \times 10^{-12}$	$\mathrm{F\,m^{-1}}$
Vacuum permeability	μ_0	$1.256\,637\,06 \times 10^{-6}$	$\mathrm{N\,A^{-2}}$
Fine-structure constant	$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c}$	$1/137.035\,999$	
Bohr radius	$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{m_e e^2}$	$5.291\,772\,11 \times 10^{-11}$	m
Rydberg constant	R_∞	$1.097\,373\,157 \times 10^{7}$	$\mathrm{m^{-1}}$
Rydberg energy	$R_\infty hc$	13.605 693	eV
Bohr magneton	$\mu_B = \frac{e\hbar}{2m_e}$	$9.274\,010\,07 \times 10^{-24}$	$\mathrm{J\,T^{-1}}$
Nuclear magneton	$\mu_N = \frac{e\hbar}{2m_p}$	$5.050\,783\,74 \times 10^{-27}$	$\mathrm{J\,T^{-1}}$
Electron g-factor	g_e	$-2.002\,319\,304$	
Compton wavelength	$\lambda_C = h/m_e c$	$2.426\,310\,24 \times 10^{-12}$	m
Classical electron radius	$r_e = \alpha^2 a_0$	$2.817\,940\,32 \times 10^{-15}$	m
Proton rms charge radius	r_p	0.8408×10^{-15}	m
Gravitational constant	G	6.6743×10^{-11}	$\mathrm{m^3\,kg^{-1}\,s^{-2}}$
Stefan–Boltzmann constant	σ	$5.670\,374\,42 \times 10^{-8}$	$\mathrm{W\,m^{-2}\,K^{-4}}$
Wien displacement constant	$b = \lambda_{\max} T$	$2.897\,772 \times 10^{-3}$	m K

The last two are derived from the exact constants, so they are exact too, though infinite decimals.

CONVENIENT COMBINATIONS

Most atomic and nuclear calculations go fastest in electronvolts and nanometers or femtometers, with these combinations:

hc	1239.842 eV nm	$\hbar c$	197.327 MeV fm
$m_e c^2$	0.510 998 95 MeV	$m_p c^2$	938.272 09 MeV
$m_n c^2$	939.565 42 MeV	$m_\mu c^2$	105.658 38 MeV
$\frac{e^2}{4\pi\varepsilon_0}$	1.439 965 eV nm	$E_h = \frac{e^2}{4\pi\varepsilon_0 a_0}$	27.211 386 eV
μ_B	$5.788\,382 \times 10^{-5}$ eV $\mathrm{T^{-1}}$	m_p/m_e	1836.152 67
$k_B T$ at 300 K	25.852 meV	1 eV$/k_B$	11 604.5 K
1 eV	$1.602\,177 \times 10^{-19}$ J	1 eV$/h$	$2.417\,99 \times 10^{14}$ Hz

For example, the energy of a 500 nm photon is $hc/\lambda = 1239.8/500$ eV $= 2.48$ eV. The Bohr radius is $a_0 = \frac{\hbar c}{\alpha\, m_e c^2} = \frac{197.327\ \text{eV nm}}{(1/137.036)(510\,999\ \text{eV})} = 0.052\,918$ nm. The de Broglie wavelength of an electron of kinetic energy K is $\lambda = \frac{hc}{\sqrt{2 m_e c^2 K}} = \frac{1.2264}{\sqrt{K/\text{eV}}}$

nm.

UNITS AND SCALES

Lengths		Energies	
1 fm (fermi)	10^{-15} m	1 eV	1.602×10^{-19} J
1 Å (ångström)	10^{-10} m $= 0.1$ nm	1 cm^{-1} ($hc\times$)	1.2398×10^{-4} eV
1 barn	10^{-28} m^2	1 hartree	27.2114 eV $= 2$ Ry

Scale	Typical length	Typical energy
Nuclear	1–10 fm	1–10 MeV per nucleon
Atomic (electronic)	0.05–0.3 nm	1–20 eV
Molecular vibration	bond length 0.1 nm	0.05–0.5 eV
Molecular rotation		10^{-4}–10^{-2} eV
Hyperfine		10^{-6}–10^{-5} eV
Thermal, room temperature		0.025 eV

ATOMIC UNITS

Theoretical atomic physics often sets $\hbar = m_e = e = 4\pi\varepsilon_0 = 1$. Then the unit of length is the Bohr radius a_0, the unit of energy is the hartree $E_h = 27.211$ eV, the unit of time is $\hbar/E_h = 2.4189 \times 10^{-17}$ s, and the speed of light is $c = 1/\alpha \approx 137.036$. The hydrogen Hamiltonian becomes $\hat{H} = -\tfrac{1}{2}\nabla^2 - \tfrac{1}{r}$, with ground-state energy $-\tfrac{1}{2}$. To restore units in a result, multiply by the appropriate combination of a_0, E_h, and $\hbar/E_h$, which is fixed by dimensional analysis. The same practice in particle physics ($\hbar = c = 1$, energies in GeV) makes lengths and times inverse energies, with $1\ \text{GeV}^{-1} = 0.1973$ fm.

THE GREEK ALPHABET

α	alpha	η	eta	ν	nu	τ	tau
β	beta	$\theta, \vartheta, \Theta$	theta	ξ, Ξ	xi	υ, Υ	upsilon
γ, Γ	gamma	ι	iota	o	omicron	ϕ, φ, Φ	phi
δ, Δ	delta	κ	kappa	π, Π	pi	χ	chi
ϵ, ε	epsilon	λ, Λ	lambda	ρ, ϱ	rho	ψ, Ψ	psi
ζ	zeta	μ	mu	σ, Σ	sigma	ω, Ω	omega

APPENDIX B

MATHEMATICAL REFERENCE

This appendix collects the formulas used most often in the book, in the book's conventions. Every entry in the tables of special functions was generated by computer algebra, and every transform and integral was checked numerically.

B.1 VECTOR IDENTITIES

Here f, g are scalar fields, $\mathbf{A}, \mathbf{B}$ vector fields, and $\mathbf{r} = \langle x, y, z \rangle$ with $r = |\mathbf{r}|$. Summation over repeated indices is implied.

$$\varepsilon_{ijk}\varepsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}, \qquad \varepsilon_{ijk}\varepsilon_{ijn} = 2\delta_{kn}, \qquad \varepsilon_{ijk}\varepsilon_{ijk} = 6,$$

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \mathbf{B}\cdot(\mathbf{C}\times\mathbf{A}) = \mathbf{C}\cdot(\mathbf{A}\times\mathbf{B}),$$

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B}),$$

$$(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = (\mathbf{A}\cdot\mathbf{C})(\mathbf{B}\cdot\mathbf{D}) - (\mathbf{A}\cdot\mathbf{D})(\mathbf{B}\cdot\mathbf{C}).$$

Products.

$$\nabla(fg) = f\nabla g + g\nabla f, \qquad \boldsymbol{\nabla}\cdot(f\mathbf{A}) = f\boldsymbol{\nabla}\cdot\mathbf{A} + \mathbf{A}\cdot\nabla f,$$

$$\boldsymbol{\nabla}\times(f\mathbf{A}) = f\boldsymbol{\nabla}\times\mathbf{A} + \nabla f\times\mathbf{A}, \qquad \boldsymbol{\nabla}\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) - \mathbf{A}\cdot(\boldsymbol{\nabla}\times\mathbf{B}),$$

$$\boldsymbol{\nabla}\times(\mathbf{A}\times\mathbf{B}) = \mathbf{A}(\boldsymbol{\nabla}\cdot\mathbf{B}) - \mathbf{B}(\boldsymbol{\nabla}\cdot\mathbf{A}) + (\mathbf{B}\cdot\nabla)\mathbf{A} - (\mathbf{A}\cdot\nabla)\mathbf{B},$$

$$\nabla(\mathbf{A}\cdot\mathbf{B}) = \mathbf{A}\times(\boldsymbol{\nabla}\times\mathbf{B}) + \mathbf{B}\times(\boldsymbol{\nabla}\times\mathbf{A}) + (\mathbf{A}\cdot\nabla)\mathbf{B} + (\mathbf{B}\cdot\nabla)\mathbf{A}.$$

Second derivatives.

$$\boldsymbol{\nabla}\times(\nabla f) = \mathbf{0}, \qquad \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) = 0,$$

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A}) = \nabla(\boldsymbol{\nabla}\cdot\mathbf{A}) - \nabla^2\mathbf{A}.$$

Powers of r (for $r \neq 0$).

$$\nabla r^n = n r^{n-2}\mathbf{r}, \qquad \boldsymbol{\nabla}\cdot(r^n\mathbf{r}) = (n+3)r^n,$$

$$\nabla^2 r^n = n(n+1)r^{n-2}, \qquad \nabla^2\frac{1}{r} = -4\pi\delta^3(\mathbf{r}).$$

B.2 Curvilinear Coordinates

Cylindrical (ρ, ϕ, z): $x = \rho\cos\phi$, $y = \rho\sin\phi$; $d\mathbf{r} = d\rho\,\hat{\boldsymbol{\rho}} + \rho\,d\phi\,\hat{\boldsymbol{\phi}} + dz\,\hat{\mathbf{z}}$; $dV = \rho\,d\rho\,d\phi\,dz$.

$$
\begin{aligned}
\nabla f &= \frac{\partial f}{\partial \rho}\hat{\boldsymbol{\rho}} + \frac{1}{\rho}\frac{\partial f}{\partial \phi}\hat{\boldsymbol{\phi}} + \frac{\partial f}{\partial z}\hat{\mathbf{z}}, \\
\boldsymbol{\nabla}\cdot\mathbf{A} &= \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho A_\rho) + \frac{1}{\rho}\frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z}, \\
\boldsymbol{\nabla}\times\mathbf{A} &= \Big(\frac{1}{\rho}\frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z}\Big)\hat{\boldsymbol{\rho}} + \Big(\frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial \rho}\Big)\hat{\boldsymbol{\phi}} \\
&\quad + \frac{1}{\rho}\Big(\frac{\partial}{\partial \rho}(\rho A_\phi) - \frac{\partial A_\rho}{\partial \phi}\Big)\hat{\mathbf{z}}, \\
\nabla^2 f &= \frac{1}{\rho}\frac{\partial}{\partial \rho}\Big(\rho\frac{\partial f}{\partial \rho}\Big) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}.
\end{aligned}
$$

Spherical (r, θ, ϕ): $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$; $d\mathbf{r} = dr\,\hat{\mathbf{r}} + r\,d\theta\,\hat{\boldsymbol{\theta}} + r\sin\theta\,d\phi\,\hat{\boldsymbol{\phi}}$; $dV = r^2\sin\theta\,dr\,d\theta\,d\phi$; $d\Omega = \sin\theta\,d\theta\,d\phi$.

$$
\begin{aligned}
\nabla f &= \frac{\partial f}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial f}{\partial \theta}\hat{\boldsymbol{\theta}} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\hat{\boldsymbol{\phi}}, \\
\boldsymbol{\nabla}\cdot\mathbf{A} &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 A_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(\sin\theta A_\theta) + \frac{1}{r\sin\theta}\frac{\partial A_\phi}{\partial \phi}, \\
\boldsymbol{\nabla}\times\mathbf{A} &= \frac{1}{r\sin\theta}\Big(\frac{\partial}{\partial \theta}(\sin\theta A_\phi) - \frac{\partial A_\theta}{\partial \phi}\Big)\hat{\mathbf{r}} + \frac{1}{r}\Big(\frac{1}{\sin\theta}\frac{\partial A_r}{\partial \phi} - \frac{\partial}{\partial r}(rA_\phi)\Big)\hat{\boldsymbol{\theta}} \\
&\quad + \frac{1}{r}\Big(\frac{\partial}{\partial r}(rA_\theta) - \frac{\partial A_r}{\partial \theta}\Big)\hat{\boldsymbol{\phi}}, \\
\nabla^2 f &= \frac{1}{r^2}\frac{\partial}{\partial r}\Big(r^2\frac{\partial f}{\partial r}\Big) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\Big(\sin\theta\frac{\partial f}{\partial \theta}\Big) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2} \\
&= \frac{1}{r}\frac{\partial^2}{\partial r^2}(rf) - \frac{\hat{L}^2}{\hbar^2 r^2}f.
\end{aligned}
$$

B.3 THE INTEGRAL THEOREMS

For a curve C from $\mathbf{a}$ to $\mathbf{b}$, a surface S with boundary ∂S, and a solid E with boundary ∂E, suitably oriented:

$$\int_C \nabla f \cdot d\mathbf{r} = f(\mathbf{b}) - f(\mathbf{a}), \qquad \oint_{\partial S} \mathbf{A} \cdot d\mathbf{r} = \iint_S (\boldsymbol{\nabla} \times \mathbf{A}) \cdot d\mathbf{S},$$
$$\oiint_{\partial E} \mathbf{A} \cdot d\mathbf{S} = \iiint_E \boldsymbol{\nabla} \cdot \mathbf{A}\, dV, \qquad \oiint_{\partial E} f\, d\mathbf{S} = \iiint_E \nabla f\, dV,$$
$$\oiint_{\partial E} d\mathbf{S} \times \mathbf{A} = \iiint_E \boldsymbol{\nabla} \times \mathbf{A}\, dV, \qquad \oint_{\partial S} f\, d\mathbf{r} = \iint_S d\mathbf{S} \times \nabla f,$$
$$\iiint_E (f\nabla^2 g - g\nabla^2 f)\, dV = \oiint_{\partial E} (f\nabla g - g\nabla f) \cdot d\mathbf{S}.$$

B.4 INTEGRALS

Gaussian integrals ($a > 0$, $n = 0, 1, 2, \ldots$, with $(-1)!! = 1$):

$$\int_{-\infty}^{\infty} x^{2n} e^{-ax^2} dx = \frac{(2n-1)!!}{(2a)^n}\sqrt{\frac{\pi}{a}}, \qquad \int_0^{\infty} x^{2n+1} e^{-ax^2} dx = \frac{n!}{2a^{n+1}},$$
$$\int_{-\infty}^{\infty} e^{-ax^2+bx} dx = \sqrt{\frac{\pi}{a}}\, e^{b^2/4a}.$$

The last holds for complex b, and for complex a with $\operatorname{Re} a > 0$.

Gamma and Beta functions. $\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}dx$; $\Gamma(n+1) = n!$; $\Gamma(\frac{1}{2}) = \sqrt{\pi}$; $\Gamma(s+1) = s\Gamma(s)$; $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$; Stirling: $n! \approx \sqrt{2\pi n}\,(n/e)^n$.

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)},$$
$$\int_0^{\pi/2} \sin^{2p-1}\theta \cos^{2q-1}\theta\, d\theta = \frac{1}{2}B(p,q).$$

Bose–Einstein and Fermi–Dirac integrals ($s > 1$; ζ is the Riemann zeta function):

$$\int_0^\infty \frac{x^{s-1}}{e^x - 1}dx = \Gamma(s)\zeta(s), \qquad \int_0^\infty \frac{x^{s-1}}{e^x + 1}dx = (1 - 2^{1-s})\Gamma(s)\zeta(s),$$

with $\zeta(2) = \frac{\pi^2}{6}$, $\zeta(3) = 1.202\,057$, $\zeta(4) = \frac{\pi^4}{90}$, $\zeta(5) = 1.036\,928$, $\zeta(6) = \frac{\pi^6}{945}$.

Other definite integrals.

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = \int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = \pi, \qquad \int_0^{\infty} x^n e^{-\lambda x} dx = \frac{n!}{\lambda^{n+1}},$$
$$\int_{-\infty}^{\infty} \frac{dx}{x^2 + a^2} = \frac{\pi}{a}, \qquad \int_{-\infty}^{\infty} \frac{dx}{(x^2 + a^2)^2} = \frac{\pi}{2a^3},$$
$$\int_0^{2\pi} \frac{d\theta}{a + b\cos\theta} = \frac{2\pi}{\sqrt{a^2 - b^2}} \ (a > |b|), \qquad \int_0^{\pi} \sin^n \theta \, d\theta = \sqrt{\pi} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2} + 1)}.$$

Delta function. $\int f(x)\delta(x-a)\,dx = f(a)$; $\delta(cx) = \delta(x)/|c|$; $\delta(g(x)) = \sum_i \delta(x - x_i)/|g'(x_i)|$ over simple zeros x_i; $\int f(x)\delta'(x-a)\,dx = -f'(a)$; $\delta(x) = \frac{1}{2\pi}\int e^{ikx} dk$; $\delta^3(\mathbf{r} - \mathbf{r}') = \frac{\delta(r-r')\delta(\theta-\theta')\delta(\phi-\phi')}{r^2 \sin\theta}$.

B.5 Series

$$\frac{1}{1-x} = \sum_{n\geq 0} x^n \ (|x| < 1), \qquad e^x = \sum_{n\geq 0} \frac{x^n}{n!},$$
$$\ln(1+x) = \sum_{n\geq 1} \frac{(-1)^{n+1} x^n}{n} \ (-1 < x \leq 1),$$
$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \cdots, \qquad x\coth x = 1 + \frac{x^2}{3} - \frac{x^4}{45} + \cdots,$$
$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots, \qquad \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots,$$
$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \cdots \ (|x| < 1).$$

Fourier series on $[-\pi, \pi]$: $f \sim \sum_n c_n e^{inx}$, $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f e^{-inx} dx$; Parseval: $\frac{1}{2\pi}\int_{-\pi}^{\pi} |f|^2 dx = \sum |c_n|^2$.

B.6 Fourier Transforms

Convention: $\hat{f}(k) = \frac{1}{\sqrt{2\pi}}\int f(x) e^{-ikx} dx$, $f(x) = \frac{1}{\sqrt{2\pi}}\int \hat{f}(k) e^{ikx} dk$. In quantum mechanics, with $p = \hbar k$, the momentum wavefunction is $\phi(p) = \hbar^{-1/2}\hat{\psi}(p/\hbar)$. Here $a > 0$.

$f(x)$	$\hat{f}(k)$	$f(x)$	$\hat{f}(k)$
e^{-ax^2}	$\frac{1}{\sqrt{2a}}e^{-k^2/4a}$	xe^{-ax^2}	$\frac{-ik}{(2a)^{3/2}}e^{-k^2/4a}$
$e^{-a\lvert x\rvert}$	$\sqrt{\frac{2}{\pi}}\frac{a}{a^2+k^2}$	$\frac{1}{x^2+a^2}$	$\sqrt{\frac{\pi}{2}}\frac{e^{-a\lvert k\rvert}}{a}$
1 for $\lvert x\rvert < a$	$\sqrt{\frac{2}{\pi}}\frac{\sin ka}{k}$	$1-\frac{\lvert x\rvert}{a}$ for $\lvert x\rvert < a$	$\sqrt{\frac{2}{\pi}}\frac{1-\cos ka}{ak^2}$
$\operatorname{sech} ax$	$\sqrt{\frac{\pi}{2}}\frac{1}{a}\operatorname{sech}\frac{\pi k}{2a}$	$\delta(x-x_0)$	$\frac{e^{-ikx_0}}{\sqrt{2\pi}}$
e^{ik_0x}	$\sqrt{2\pi}\,\delta(k-k_0)$	1	$\sqrt{2\pi}\,\delta(k)$

Properties. $f(x-x_0)\mapsto e^{-ikx_0}\hat{f}$; $e^{ik_0x}f\mapsto\hat{f}(k-k_0)$; $f(cx)\mapsto\frac{1}{|c|}\hat{f}(k/c)$; $f'\mapsto ik\hat{f}$; $xf\mapsto i\hat{f}'$; $f*g\mapsto\sqrt{2\pi}\hat{f}\hat{g}$ with $(f*g)(x)=\int f(y)g(x-y)\,dy$; $\int|f|^2dx=\int|\hat{f}|^2dk$; $\sigma_x\sigma_k\geq\frac{1}{2}$.

B.7 Special Functions

Hermite polynomials. $H_n(x)=(-1)^ne^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}e^{-x^2}$; $H_n''-2xH_n'+2nH_n=0$; $\int H_mH_ne^{-x^2}dx=\sqrt{\pi}\,2^n n!\,\delta_{mn}$; $H_{n+1}=2xH_n-2nH_{n-1}$. Oscillator states: $\psi_n(x)=\left(\frac{1}{\sqrt{\pi}\,2^n n!\,x_0}\right)^{1/2}H_n(x/x_0)e^{-x^2/2x_0^2}$, $x_0=\sqrt{\hbar/m\omega}$.

$$H_0=1,\quad H_1=2x,$$
$$H_2=4x^2-2,\quad H_3=8x^3-12x,$$
$$H_4=16x^4-48x^2+12,\quad H_5=32x^5-160x^3+120x,$$
$$H_6=64x^6-480x^4+720x^2-120.$$

Legendre polynomials. $P_\ell(x)=\frac{1}{2^\ell\ell!}\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}\left(x^2-1\right)^\ell$; $(1-x^2)P_\ell''-2xP_\ell'+\ell(\ell+1)P_\ell=0$; $\int_{-1}^{1}P_\ell P_{\ell'}dx=\frac{2}{2\ell+1}\delta_{\ell\ell'}$; $(\ell+1)P_{\ell+1}=(2\ell+1)xP_\ell-\ell P_{\ell-1}$; $\frac{1}{\sqrt{1-2xt+t^2}}=\sum_\ell P_\ell(x)t^\ell$.

$$P_0=1,\quad P_1=x,$$
$$P_2=\tfrac{1}{2}(3x^2-1),\quad P_3=\tfrac{1}{2}(5x^3-3x),$$
$$P_4=\tfrac{1}{8}(35x^4-30x^2+3),\quad P_5=\tfrac{1}{8}(63x^5-70x^3+15x),$$
$$P_6=\tfrac{1}{16}(231x^6-315x^4+105x^2-5).$$

Spherical harmonics (Theorem 16.2), with $Y_\ell^{-m} = (-1)^m Y_\ell^{m*}$:

$$Y_0^0 = \frac{1}{\sqrt{4\pi}}, \qquad Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta,$$
$$Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{i\phi}, \qquad Y_2^0 = \sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1),$$
$$Y_2^1 = -\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\, e^{i\phi}, \qquad Y_2^2 = \sqrt{\frac{15}{32\pi}}\sin^2\theta\, e^{2i\phi},$$
$$Y_3^0 = \sqrt{\frac{7}{16\pi}}(5\cos^3\theta - 3\cos\theta), \qquad Y_3^1 = -\sqrt{\frac{21}{64\pi}}\sin\theta(5\cos^2\theta - 1)e^{i\phi},$$
$$Y_3^2 = \sqrt{\frac{105}{32\pi}}\sin^2\theta\cos\theta\, e^{2i\phi}, \qquad Y_3^3 = -\sqrt{\frac{35}{64\pi}}\sin^3\theta\, e^{3i\phi}.$$

Addition theorem: $P_\ell(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') = \frac{4\pi}{2\ell+1}\sum_m Y_\ell^m(\hat{\mathbf{r}})Y_\ell^{m*}(\hat{\mathbf{r}}')$; Unsöld: $\sum_m |Y_\ell^m|^2 = \frac{2\ell+1}{4\pi}$.

Hydrogen radial functions (Theorem 16.4), in units where $a = 1$ (multiply by $a^{-3/2}$ and read r as r/a):

$$R_{10} = 2e^{-r}, \qquad R_{20} = \tfrac{1}{\sqrt{2}}\big(1 - \tfrac{r}{2}\big)e^{-r/2},$$
$$R_{21} = \tfrac{1}{2\sqrt{6}}\, r\, e^{-r/2}, \qquad R_{30} = \tfrac{2}{3\sqrt{3}}\big(1 - \tfrac{2r}{3} + \tfrac{2r^2}{27}\big)e^{-r/3},$$
$$R_{31} = \tfrac{4\sqrt{6}}{81}\, r\big(1 - \tfrac{r}{6}\big)e^{-r/3}, \qquad R_{32} = \tfrac{2\sqrt{30}}{1215}\, r^2 e^{-r/3},$$
$$R_{40} = \tfrac{1}{4}\big(1 - \tfrac{3r}{4} + \tfrac{r^2}{8} - \tfrac{r^3}{192}\big)e^{-r/4}, \qquad R_{41} = \tfrac{\sqrt{5}}{16\sqrt{3}}\, r\big(1 - \tfrac{r}{4} + \tfrac{r^2}{80}\big)e^{-r/4},$$
$$R_{42} = \tfrac{1}{64\sqrt{5}}\, r^2\big(1 - \tfrac{r}{12}\big)e^{-r/4}, \qquad R_{43} = \tfrac{1}{768\sqrt{35}}\, r^3 e^{-r/4}.$$

Expectation values: $\langle r\rangle = \frac{a}{2}\big[3n^2 - \ell(\ell+1)\big]$, $\langle r^2\rangle = \frac{a^2n^2}{2}\big[5n^2 + 1 - 3\ell(\ell+1)\big]$, $\big\langle\frac{1}{r}\big\rangle = \frac{1}{n^2 a}$, $\big\langle\frac{1}{r^2}\big\rangle = \frac{1}{n^3(\ell+\frac{1}{2})a^2}$, $\big\langle\frac{1}{r^3}\big\rangle = \frac{1}{n^3\ell(\ell+\frac{1}{2})(\ell+1)a^3}$ $(\ell \geq 1)$.

Bessel functions. $x^2y'' + xy' + (x^2 - \nu^2)y = 0$; $J_\nu(x) = \sum_{k\geq 0}\frac{(-1)^k}{k!\,\Gamma(k+\nu+1)}\big(\frac{x}{2}\big)^{2k+\nu}$. First zeros: J_0: 2.4048, 5.5201, 8.6537; J_1: 3.8317, 7.0156. Spherical: $j_0(x) = \frac{\sin x}{x}$, $j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x}$. Airy: $\mathrm{Ai}'' = x\,\mathrm{Ai}$; zeros of Ai: -2.3381, -4.0879, -5.5206; of Ai′: -1.0188, -3.2482, -4.8201.

B.8 Spin and Angular Momentum

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \sigma_i\sigma_j = \delta_{ij}I + i\varepsilon_{ijk}\sigma_k,$$

$$(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{b}) = (\mathbf{a}\cdot\mathbf{b})I + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b}),$$

$$e^{-i\alpha\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}/2} = \cos\tfrac{\alpha}{2}\, I - i\sin\tfrac{\alpha}{2}\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}.$$

Spin up along $\hat{\mathbf{n}}(\theta,\phi)$: $|+\hat{\mathbf{n}}\rangle = \cos\frac{\theta}{2}|\uparrow\rangle + e^{i\phi}\sin\frac{\theta}{2}|\downarrow\rangle$. General angular momentum: $[\hat{J}_i,\hat{J}_j] = i\hbar\varepsilon_{ijk}\hat{J}_k$, $\hat{J}_\pm|j,m\rangle = \hbar\sqrt{j(j+1)-m(m\pm1)}\,|j,m\pm1\rangle$.

Clebsch–Gordan decompositions. Two spins $\frac{1}{2}$:

$$|1,1\rangle = |\uparrow\uparrow\rangle, \qquad |1,0\rangle = \tfrac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\big), \qquad |1,-1\rangle = |\downarrow\downarrow\rangle,$$
$$|0,0\rangle = \tfrac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\big).$$

$\ell = 1$ and $s = \frac{1}{2}$, in the basis $|m_\ell\rangle|m_s\rangle$:

$$|\tfrac{3}{2},\tfrac{3}{2}\rangle = |1\rangle|\uparrow\rangle, \qquad |\tfrac{3}{2},\tfrac{1}{2}\rangle = \sqrt{\tfrac{2}{3}}\,|0\rangle|\uparrow\rangle + \sqrt{\tfrac{1}{3}}\,|1\rangle|\downarrow\rangle,$$
$$|\tfrac{3}{2},-\tfrac{1}{2}\rangle = \sqrt{\tfrac{2}{3}}\,|0\rangle|\downarrow\rangle + \sqrt{\tfrac{1}{3}}\,|-1\rangle|\uparrow\rangle, \qquad |\tfrac{3}{2},-\tfrac{3}{2}\rangle = |-1\rangle|\downarrow\rangle,$$
$$|\tfrac{1}{2},\tfrac{1}{2}\rangle = \sqrt{\tfrac{2}{3}}\,|1\rangle|\downarrow\rangle - \sqrt{\tfrac{1}{3}}\,|0\rangle|\uparrow\rangle, \qquad |\tfrac{1}{2},-\tfrac{1}{2}\rangle = \sqrt{\tfrac{1}{3}}\,|0\rangle|\downarrow\rangle - \sqrt{\tfrac{2}{3}}\,|-1\rangle|\uparrow\rangle.$$

These follow the Condon–Shortley phase convention, in which the coefficient $\langle j_1, j_1; j_2, J-j_1|J,J\rangle$ is positive.

APPENDIX C

NUMERICAL METHODS FOR QUANTUM MECHANICS

Only a handful of potentials can be solved in closed form. For everything else, and to check the exact solutions, quantum mechanics is done numerically. This appendix presents five short Python programs, each built on an idea from the text, and the output each actually produced. They use only `numpy` and `scipy`. Every program uses units with $\hbar = m = 1$ (or $\hbar = 2m = 1$, where stated), so the Schrödinger equation reads $-\frac{1}{2}\psi'' + V\psi = E\psi$.

The five methods attack the three central problems: finding energy levels (finite differences, shooting, basis expansion) and following the time evolution (Crank–Nicolson, split-operator). Each section states what the method approximates, how its error behaves, and how the output confirms it.

C.1 FINITE DIFFERENCES: THE HAMILTONIAN AS A MATRIX

Replace the line by grid points $x_j = x_0 + jh$ and the wavefunction by its values ψ_j. By Taylor's theorem,

$$\psi''(x_j) = \frac{\psi_{j+1} - 2\psi_j + \psi_{j-1}}{h^2} - \frac{h^2}{12}\psi^{(4)}(x_j) + O(h^4).$$

Dropping the error term turns $\hat{H}$ into a symmetric tridiagonal matrix with diagonal $\frac{1}{h^2} + V(x_j)$ and off-diagonal $-\frac{1}{2h^2}$. Its eigenvalues approximate the energies, and its eigenvectors the sampled wavefunctions. The matrix is real symmetric, so by the spectral theorem (Theorem 9.27) the approximate energies are real and the approximate states orthogonal: the discretization preserves the structure of the exact problem. Imposing $\psi = 0$ at the ends of a finite interval $[-L, L]$ is harmless for bound states, whose wavefunctions are exponentially small there.

```
import numpy as np
from scipy.linalg import eigh_tridiagonal
```

```
def levels(V, L, N, k=4):
    """Lowest k eigenvalues of -1/2 psi'' + V psi on [-L, L], psi = 0 at the
    ends."""
    x, h = np.linspace(-L, L, N + 2, retstep=True)
    x = x[1:-1]                                  # interior grid points
    diag = 1.0 / h**2 + V(x)                     # -1/2 * (-2/h^2) + V
    off = -0.5 / h**2 * np.ones(N - 1)           # -1/2 * (1/h^2)
    return eigh_tridiagonal(diag, off, select='i', select_range=(0, k - 1))[0]

for N in (100, 200, 400, 800):
    E = levels(lambda x: 0.5 * x**2, 8.0, N)
    print(N, np.round(E, 6), "error in E0: %.2e" % abs(E[0] - 0.5))
print("quartic:", np.round(levels(lambda x: x**4, 6.0, 4000), 6))
```

OUTPUT

```
100 [0.499215 1.496068 2.489761 3.48028 ] error in E0: 7.85e-04
200 [0.499802 1.499009 2.497423 3.495042] error in E0: 1.98e-04
400 [0.49995  1.499751 2.499353 3.498756] error in E0: 4.98e-05
800 [0.499988 1.499938 2.499838 3.499688] error in E0: 1.25e-05
quartic: [0.667985 2.393638 4.696775 7.335682]
```

For the oscillator, the exact levels are $n + \frac{1}{2}$. Each doubling of N halves h and divides the error by 4, the signature of a method of order h^2. Higher levels have larger errors, because their wavefunctions oscillate faster and $\psi^{(4)}$ is larger. For the quartic oscillator, 4000 points give $E_0 = 0.667985$ and $E_1 = 2.393638$, the values used as benchmarks in the variational problems of Chapter 15. Because the matrix is tridiagonal, the cost is proportional to N, and millions of points are practical.

C.2 Shooting: Quantization as a Boundary Condition

The Schrödinger equation is a second-order ODE. For any E it has solutions; quantization enters only through the requirement that ψ decay at infinity. The *shooting method* makes this literal. Start at $x = 0$ with initial data of definite parity ($\psi = 1, \psi' = 0$ for even states), integrate outward with an accurate ODE solver, and watch $\psi(L)$ at a large L. For most E the solution blows up, to $+\infty$ or $-\infty$. At an eigenvalue it changes from one to the other, so the energies are the sign changes of $\psi(L)$ as a function of E, and a root finder such as bisection locates them.

The program applies this to the finite well with $z_0 = 4$, in units $\hbar^2/2m = 1$ and

half-width $a = 1$, so $V = -16$ inside, and compares with the transcendental equation of Theorem 14.14.

```
import numpy as np
from scipy.integrate import solve_ivp
from scipy.optimize import brentq

z0 = 4.0                                    # well strength; V = -z0^2 for |x| <
    1
V = lambda x: -z0**2 if abs(x) < 1 else 0.0

def tail(E, L=6.0):
    """Integrate psi'' = (V - E) psi from x = 0 (even start) to x = L; return
    psi(L)."""
    rhs = lambda x, y: [y[1], (V(x) - E) * y[0]]
    sol = solve_ivp(rhs, [0, L], [1.0, 0.0], max_step=0.01, rtol=1e-10, atol=1e
    -12)
    return sol.y[0, -1]

Es = np.linspace(-z0**2 + 1e-6, -1e-6, 400)       # scan for sign changes of
    psi(L)
vals = [tail(E) for E in Es]
roots = [brentq(tail, Es[i], Es[i + 1]) for i in range(len(Es) - 1) if vals[i]
    * vals[i + 1] < 0]
print("shooting:     ", np.round(roots, 6))

f = lambda z: z * np.tan(z) - np.sqrt(z0**2 - z**2)   # even condition of the
    finite well
zs = [brentq(f, 1e-6, np.pi / 2 - 1e-9), brentq(f, np.pi + 1e-9, z0)]
print("transcendental:", np.round([z**2 - z0**2 for z in zs], 6))
```

OUTPUT

```
shooting:      [-14.431611  -3.073783]
transcendental: [-14.431611  -3.073783]
```

The agreement is to all six digits printed. The program never used the analytic solution inside or outside the well. It would work for any potential, including ones that make the transcendental-equation approach impossible. The method's weakness is instability: integrating outward, the growing exponential amplifies rounding errors, and L cannot be taken too large. Integrating inward from both ends and matching ψ'/ψ in the middle avoids this, as in the Numerov method used for atomic structure.

C.3 Basis Expansion: Rayleigh–Ritz with Gaussians

The variational principle (Theorem 15.9) turns into linear algebra when the trial function is a linear combination $\psi = \sum c_i \phi_i$ of fixed basis functions. Minimizing $\langle \hat{H} \rangle / \langle \psi | \psi \rangle$ over the coefficients gives the generalized eigenvalue problem $H\mathbf{c} = ES\mathbf{c}$, with $H_{ij} = \langle \phi_i | \hat{H} | \phi_j \rangle$ and overlap $S_{ij} = \langle \phi_i | \phi_j \rangle$, as in the Rayleigh–Ritz problem of Chapter 15. The lowest eigenvalue is an upper bound to the ground-state energy, and it decreases as the basis grows.

Quantum chemistry uses Gaussians $e^{-\alpha r^2}$, because every integral involving them is elementary (Theorem 5.31). For hydrogen in atomic units ($\hat{H} = -\tfrac{1}{2}\nabla^2 - \tfrac{1}{r}$), with $p = \alpha_i + \alpha_j$,

$$S_{ij} = \left(\frac{\pi}{p}\right)^{3/2}, \qquad T_{ij} = \frac{3\alpha_i \alpha_j}{p} S_{ij}, \qquad V_{ij} = -\frac{2\pi}{p}.$$

The exponents are taken in geometric progression, $\alpha_k = \alpha_0 \beta^k$ (an *even-tempered* basis), with α_0 and β optimized.

```
import numpy as np
from scipy.linalg import eigh
from scipy.optimize import minimize

# Hydrogen ground state (atomic units) in a basis of s-type Gaussians exp(-a r
    ^2).
def ground_energy(alphas):
    a = np.array(alphas)[:, None]
    b = np.array(alphas)[None, :]
    p = a + b
    S = (np.pi / p) ** 1.5                      # overlap
    T = 3 * a * b / p * S                       # kinetic energy, -1/2 Laplacian
    V = -2 * np.pi / p                          # nuclear attraction, -1/r
    return eigh(T + V, S, eigvals_only=True)[0]

def even_tempered(n):
    """Exponents a0 * beta^k, k = 0..n-1, with a0 and beta optimized
    variationally."""
    f = lambda q: ground_energy(np.exp(q[0]) * np.exp(q[1]) ** np.arange(n))
    best = minimize(f, [np.log(0.05), np.log(3.0)], method="Nelder-Mead",
                    options={"xatol": 1e-10, "fatol": 1e-14, "maxiter": 20000})
    return best.fun

print(" 1 Gaussian:   E = %.6f   (exact for one Gaussian: -4/(3 pi) = %.6f)" %
    (even_tempered(1), -4 / (3 * np.pi)))
for n in (2, 4, 8, 12):
    E = even_tempered(n)
```

```
    print("%2d Gaussians: E = %.6f   error = %.1e" % (n, E, E + 0.5))
```

OUTPUT

```
 1 Gaussian:   E = -0.424413   (exact for one Gaussian: -4/(3 pi) = -0.424413)
 2 Gaussians: E = -0.485813   error = 1.4e-02
 4 Gaussians: E = -0.498752   error = 1.2e-03
 8 Gaussians: E = -0.499974   error = 2.6e-05
12 Gaussians: E = -0.499999   error = 1.1e-06
```

The exact energy is $-\frac{1}{2}$ hartree. One Gaussian gives $-\frac{4}{3\pi}$, 15% high, because a Gaussian has zero slope at the origin where the true e^{-r} has a cusp. Twelve Gaussians reach -0.499999. Every value is above $-\frac{1}{2}$, as the variational principle guarantees, and the convergence is slow but steady, since no finite sum of Gaussians reproduces the cusp exactly. Basis sets of this kind, with hundreds of functions per molecule, are the workhorse of computational chemistry.

C.4 Time Evolution: The Crank–Nicolson Method

To evolve ψ we need $e^{-i\hat{H}\Delta t}$. The simplest approximation, $1 - i\hat{H}\Delta t$, is not unitary, and it makes the norm grow without bound. The Crank–Nicolson method uses instead the *Cayley form*

$$e^{-i\hat{H}\Delta t} \approx \left(1 + \frac{i\Delta t}{2}\hat{H}\right)^{-1}\left(1 - \frac{i\Delta t}{2}\hat{H}\right),$$

which agrees with the exponential through second order in Δt. For Hermitian $\hat{H}$ it is *exactly* unitary: if λ is an eigenvalue of $\hat{H}$, the corresponding eigenvalue of the Cayley form is $\frac{1-i\lambda\Delta t/2}{1+i\lambda\Delta t/2}$, a complex number of modulus 1. So probability is conserved to rounding error, at any step size, and the method is stable. Each step requires solving one tridiagonal linear system, which costs $O(N)$.

The program sends a Gaussian packet with mean wavenumber $k_0 = 1.2$ (mean energy ≈ 0.72) at a rectangular barrier of height 1 and width 1, and measures the probability on the far side after the collision. Since the packet contains a spread of wavenumbers, the prediction is the plane-wave transmission coefficient of Theorem 14.17, averaged over the packet's momentum distribution $|\phi(k)|^2 \propto e^{-2s^2(k-k_0)^2}$ (valid for $E < V_0$ and, with sinh $\to$ sin, for $E > V_0$).

```
import numpy as np
from scipy.sparse import diags, identity
from scipy.sparse.linalg import splu

# Units hbar = m = 1. Grid, barrier, and initial Gaussian packet.
```

```
L, N, dt = 200.0, 8000, 0.01
x, h = np.linspace(-L / 2, L / 2, N, retstep=True)
V0, w = 1.0, 1.0                                    # barrier height and width
V = np.where((x > 0) & (x < w), V0, 0.0)
k0, s = 1.2, 5.0                                    # mean wavenumber, width
psi = (2 * np.pi * s**2) ** -0.25 * np.exp(-(x + 40) ** 2 / (4 * s**2) + 1j *
    k0 * x)

# H = -1/2 d^2/dx^2 + V; Crank-Nicolson: (1 + i dt H/2) psi_new = (1 - i dt H
    /2) psi_old
H = diags([-0.5 / h**2, 1 / h**2 + V, -0.5 / h**2], [-1, 0, 1], shape=(N, N))
A = (identity(N) + 0.5j * dt * H).tocsc()
B = (identity(N) - 0.5j * dt * H).tocsc()
solve = splu(A).solve
for step in range(int(60 / dt)):                    # evolve to t = 60
    psi = solve(B @ psi)

norm = np.sum(abs(psi) ** 2) * h
T_num = np.sum(abs(psi[x > w]) ** 2) * h
# Compare: average the exact plane-wave T(k) over the packet's momentum
    distribution.
k = np.linspace(0.01, 3, 6000)
E = k**2 / 2
kap = np.sqrt(2 * (V0 - E + 0j))
T_k = np.real(1 / (1 + V0**2 * np.abs(np.sinh(kap * w)) ** 2 / (4 * E * np.abs(
    V0 - E))))
weight = np.exp(-2 * s**2 * (k - k0) ** 2)
print("norm after evolution: %.12f" % norm)
print("transmitted, numerical: %.4f" % T_num)
print("transmitted, from T(k): %.4f" % (np.sum(weight * T_k) / np.sum(weight)))
```

OUTPUT

```
norm after evolution: 0.999999999999
transmitted, numerical: 0.5436
transmitted, from T(k): 0.5438
```

The norm is conserved to 10^{-12} after 6000 steps, and the transmitted probability agrees with the stationary-state prediction to 2×10^{-4}. This confirms numerically that the time-independent scattering theory of Chapter 14, built from non-normalizable plane waves, correctly describes real, localized wave packets. More than half the packet is transmitted although its mean energy is below the barrier, because the faster components tunnel easily and some pass over the top.

C.5 The Split-Operator Method

The kinetic energy is diagonal in momentum space and the potential is diagonal in position space. The fast Fourier transform (FFT) moves between them in $O(N \log N)$ operations. The *split-operator* method approximates

$$e^{-i(\hat{T}+\hat{V})\Delta t} \approx e^{-i\hat{V}\Delta t/2}\, e^{-i\hat{T}\Delta t}\, e^{-i\hat{V}\Delta t/2},$$

with an error of order Δt^3 per step, from the commutator $[\hat{T}, \hat{V}]$. Every factor is exactly unitary. The kinetic factor is applied by transforming to k-space, multiplying by $e^{-ik^2\Delta t/2}$, and transforming back. On a periodic grid, the discrete Fourier transform of Chapter 11 plays the role of the Fourier transform.

```
import numpy as np

# Units hbar = m = 1. Split-operator method on a periodic grid.
N, L = 4096, 400.0
x = (np.arange(N) - N // 2) * (L / N)
k = 2 * np.pi * np.fft.fftfreq(N, d=L / N)
s0, k0 = 1.0, 2.0
psi = (2 * np.pi * s0**2) ** -0.25 * np.exp(-x**2 / (4 * s0**2) + 1j * k0 * x)

def evolve(psi, V, t, dt):
    """exp(-iHt) ~ [e^{-iV dt/2} e^{-iT dt} e^{-iV dt/2}]^n, with T applied in
    k-space."""
    n = max(1, int(round(t / dt)))
    dt = t / n                                  # make n steps land exactly on
    t
    half = np.exp(-0.5j * V * dt)
    kin = np.exp(-0.5j * k**2 * dt)
    for _ in range(n if t > 0 else 0):
        psi = half * np.fft.ifft(kin * np.fft.fft(half * psi))
    return psi

dx = L / N
for t in (0, 2, 5, 10):
    p = evolve(psi, np.zeros(N), t, 0.01)
    rho = abs(p) ** 2
    mean = np.sum(x * rho) * dx
    sig = np.sqrt(np.sum((x - mean) ** 2 * rho) * dx)
    exact = s0 * np.sqrt(1 + (t / (2 * s0**2)) ** 2)
    print("t=%4.1f  <x>=%8.4f (k0 t=%5.1f)  sigma=%.6f  exact=%.6f" % (t, mean,
     k0 * t, sig, exact))

# Harmonic potential: a displaced ground state (coherent state) should
    oscillate rigidly.
V = 0.5 * x**2
phi = np.pi ** -0.25 * np.exp(-(x - 3.0) ** 2 / 2)
```

```
for t in (np.pi / 2, np.pi, 2 * np.pi):
    p = evolve(phi, V, t, 0.001)
    rho = abs(p) ** 2
    mean = np.sum(x * rho) * dx
    sig = np.sqrt(np.sum((x - mean) ** 2 * rho) * dx)
    print("t=%6.4f  <x>=%9.6f  3cos t=%9.6f  sigma=%.6f" % (t, mean, 3 * np.cos
    (t), sig))
```

OUTPUT

```
t= 0.0  <x>=  0.0000 (k0 t=  0.0)  sigma=1.000000  exact=1.000000
t= 2.0  <x>=  4.0000 (k0 t=  4.0)  sigma=1.414214  exact=1.414214
t= 5.0  <x>= 10.0000 (k0 t= 10.0)  sigma=2.692582  exact=2.692582
t=10.0  <x>= 20.0000 (k0 t= 20.0)  sigma=5.099020  exact=5.099020
t=1.5708  <x>=-0.000000  3cos t= 0.000000  sigma=0.707107
t=3.1416  <x>=-3.000000  3cos t=-3.000000  sigma=0.707107
t=6.2832  <x>= 3.000000  3cos t= 3.000000  sigma=0.707107
```

For the free particle $\hat{V} = 0$, so the splitting is exact, and the computed widths agree with $\sigma(t) = \sigma_0\sqrt{1 + (t/2\sigma_0^2)^2}$ of Theorem 13.14 to all printed digits. The center moves at speed k_0, as Ehrenfest's theorem requires. For the oscillator, the displaced ground state oscillates as a rigid body, $\langle x \rangle = 3\cos t$ with constant $\sigma = 1/\sqrt{2}$, which is the coherent-state behavior derived in the exercises of Chapter 14. After one full period, the packet returns to $x = 3$ to six digits.

C.6 Choosing a Method

Method	Best for	Error and cost
Finite differences	Many levels of a 1D or 2D potential	$O(h^2)$; sparse matrix, $O(N)$ per level
Shooting	A few levels, any 1D potential	Set by the ODE solver; unstable for large L
Basis expansion	Ground states, atoms and molecules	Variational upper bound; basis-dependent
Crank–Nicolson	Time evolution with sharp potentials	$O(\Delta t^2)$, exactly unitary, $O(N)$ per step
Split-operator	Time evolution with smooth potentials	$O(\Delta t^2)$, exactly unitary, $O(N \log N)$ per step

In every case, the essential check is the one used throughout this appendix: compare against an exact result, then halve the step and confirm that the answer changes by the predicted amount.

Further Reading

The books below are standard references. Each covers ground this book treats more briefly, or approaches it from a different angle. They are grouped by the parts of this book they extend.

Vector calculus (Part I)

J. E. Marsden and A. Tromba, Vector Calculus. A thorough treatment at about the level of Part I, with more geometry and more applications to physics.

H. M. Schey, Div, Grad, Curl, and All That. A short, informal introduction to the integral theorems through electrostatics. Good for intuition before or after Chapter 7.

M. Spivak, Calculus on Manifolds. The rigorous theory of differentiation and integration in $\mathbb{R}^n$, culminating in the generalized Stokes theorem with differential forms.

Mathematical methods (Part II)

M. L. Boas, Mathematical Methods in the Physical Sciences. The standard undergraduate survey of the mathematics of Part II, with many exercises.

K. F. Riley, M. P. Hobson, and S. J. Bence, Mathematical Methods for Physics and Engineering. More advanced and more complete than Boas, and an excellent reference.

J. W. Brown and R. V. Churchill, Complex Variables and Applications. A careful first course in complex analysis, including contour integration and conformal mapping.

S. Axler, Linear Algebra Done Right. Linear algebra from the operator point of view of Chapter 9, with the spectral theorem at its center.

E. M. Stein and R. Shakarchi, Fourier Analysis: An Introduction. Rigorous

Fourier series and transforms, with applications to differential equations and number theory.

Quantum mechanics (Part III)

D. J. Griffiths and D. F. Schroeter, Introduction to Quantum Mechanics. The most widely used undergraduate text, at a level close to Part III, with more on perturbation theory, scattering, and approximation methods.

J. S. Townsend, A Modern Approach to Quantum Mechanics. Begins with spin and the Stern–Gerlach experiment and builds the formalism from two-state systems, an order that complements this book's.

R. P. Feynman, R. B. Leighton, and M. Sands, The Feynman Lectures on Physics, *Volume III.* Quantum mechanics from amplitudes and two-state systems, with unmatched physical insight.

R. Shankar, Principles of Quantum Mechanics. A graduate-level text with a careful mathematical introduction and a full treatment of path integrals.

J. J. Sakurai and J. Napolitano, Modern Quantum Mechanics. The standard first-year graduate text, organized around symmetry and angular momentum.

M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information. The reference for the material of the last sections of Chapter 18: qubits, gates, algorithms, entanglement, and error correction.

Numerical methods

W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Numerical Recipes. Practical algorithms, with discussion of when each works and why, for every method in Appendix C and many more.

Answers to Selected Exercises

Answers are given for odd-numbered computational exercises. Every numerical answer here was checked by computer algebra. Proof exercises are not answered; check proofs against the definitions and theorems they rely on.

Chapter 1

Vectors and the geometry of space

1.1 14

1.3 Center $(-4, 1, -3)$, radius 5

1.5 $\langle -4, 4, 2\rangle$; norm 6; $\left\langle -\tfrac{2}{3}, \tfrac{2}{3}, \tfrac{1}{3}\right\rangle$

1.7 $40\sqrt{31} \approx 222.7$ mph on a bearing of about 38.9°

1.9 60°

1.11 $k = 10/7$

1.13 $\mathbf{u}_{\parallel} = \left\langle \tfrac{4}{3}, \tfrac{4}{3}, \tfrac{4}{3}\right\rangle$, $\mathbf{u}_{\perp} = \left\langle \tfrac{2}{3}, \tfrac{5}{3}, -\tfrac{7}{3}\right\rangle$

1.15 $6\,\mathrm{J}$

1.17 $\langle -10, 7, 9\rangle$

1.19 $\sqrt{61}/2 \approx 3.91$

1.21 13

1.23 $\boldsymbol{\tau} = \langle 0, 9, 12\rangle$ N m; $|\boldsymbol{\tau}| = 15\,\mathrm{N\,m}$

1.25 $\dfrac{x-1}{3} = \dfrac{y-2}{-2} = \dfrac{z-3}{4}$

1.27 $2x - y + 4z = 12$

1.29 $3x - y + 2z = 8$

1.31 3

1.33 $\mathbf{r}(t) = \langle 1, -1, 0\rangle + t\,\langle 0, 1, 1\rangle$

1.35 Hyperboloid of one sheet about the z-axis. $z = 0$: ellipse $\tfrac{x^2}{4} + y^2 = 1$; $z = 2$: ellipse $\tfrac{x^2}{8} + \tfrac{y^2}{2} = 1$; $x = 0$: hyperbola $y^2 - \tfrac{z^2}{4} = 1$

1.37 Cylindrical $(2,\, 2\pi/3,\, 2)$; spherical $(2\sqrt{2},\, \pi/4,\, 2\pi/3)$

1.39 $r = 4\cos\theta$; $z = \rho^2$

1.49 Distance $\sqrt{3}/3$. Closest points $\left(-\tfrac{1}{3}, -\tfrac{4}{3}, 0\right)$ at $t = -\tfrac{4}{3}$ and $\left(0, -\tfrac{5}{3}, \tfrac{1}{3}\right)$ at $s = -\tfrac{5}{3}$

1.53 (a) $\tfrac{1}{3} + \tfrac{2}{3} = 1$; (b) $\tfrac{1}{3}$; (c) $\left\langle \sqrt{2/3},\, -1/\sqrt{3}\right\rangle$ (or its negative)

Chapter 2

Vector functions and motion in space

2.1 Domain $[-1, 2)$; limit $\langle \ln 4,\, 1,\, 1\rangle$

2.3 $\mathbf{r}(t) = \langle 2 + 2\cos t,\; 2\sin t,\; 8 + 8\cos t\rangle$

2.5 $\mathbf{r}'(0) = \langle 1, 0, 1\rangle$; $\hat{\mathbf{T}}(0) = \left\langle \tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right\rangle$

2.7 $\langle 1,\; \pi/4,\; (e-1)/2\rangle$

2.9 15

2.11 10π; $\mathbf{r}(s) = \left\langle 3\cos\tfrac{s}{5},\; 3\sin\tfrac{s}{5},\; \tfrac{4s}{5}\right\rangle$

2.13 $\kappa(0) = 2$; $\kappa(1) = \sqrt{266}/98 \approx 0.166$

2.15 $\kappa = 1$; osculating circle $(x - \tfrac{\pi}{2})^2 + y^2 = 1$

2.17 $\hat{\mathbf{T}} = \frac{1}{\sqrt{37}}\langle 6\cos 3t, 1, -6\sin 3t\rangle$, $\hat{\mathbf{N}} = \langle -\sin 3t, 0, -\cos 3t\rangle$, $\hat{\mathbf{B}} = \frac{1}{\sqrt{37}}\langle -\cos 3t, 6, \sin 3t\rangle$

2.19 $3x - 3y + z = 1$

2.21 $a_T = e^t - e^{-t}$, $a_N = \sqrt{2}$

2.23 About 35.8 m from the base of the building, after about 3.37 s

2.25 $\mathbf{r}(t) = \langle t^3 + t,\ t^4 - t,\ 3t - t^3\rangle$

2.27 About 92.4 minutes

CHAPTER 3

PARTIAL DERIVATIVES AND THE GRADIENT

3.1 The open elliptical region $\frac{x^2}{9} + y^2 < 1$

3.3 Parabolas $y = x^2 + k$, shifted vertically by k

3.5 $-\frac{2}{5}$

3.7 0 (since $\frac{x^2y^2}{x^2+y^2} \le y^2$)

3.9 1

3.11 $f_x = 3x^2y^2 - 4y - \dfrac{y}{x^2+y^2}$, $f_y = 2x^3y - 4x + \dfrac{x}{x^2+y^2}$

3.13 $f_{xx} = -y^2\sin xy$, $f_{xy} = f_{yx} = \cos xy - xy\sin xy$, $f_{yy} = -x^2\sin xy$

3.17 $z = 3x + 7y - 5$

3.19 $\Delta V \approx 2360\ \text{cm}^3$

3.21 6

3.23 $z_x = -\dfrac{yz + \sin(x+y+z)}{xy + \sin(x+y+z)}$, $z_y = -\dfrac{xz + \sin(x+y+z)}{xy + \sin(x+y+z)}$

3.25 $32/\sqrt{29}$

3.27 $\sqrt{17}/2$, in the direction of $\langle 0, 1, 4\rangle$

3.29 Plane $4x - 2y + 3z = 21$; normal line $\langle 4, -1, 1\rangle + t\langle 4, -2, 3\rangle$

3.31 $4(u^2 + v^2)$; locally invertible everywhere except $(u, v) = (0, 0)$

CHAPTER 4

OPTIMIZATION IN SEVERAL VARIABLES

4.1 xy

4.3 $1 + (x-1) - (y-1) + \frac{1}{2}(x-1)^2 - (x-1)(y-1) + \frac{1}{2}(y-1)^2$

4.5 Minors $1, 1, 3$: positive definite

4.7 Local (and absolute) minimum -5 at $(-3, 3)$

4.9 Saddles at $(0,0)$, $(1,0)$, $(0,1)$; local maximum $\frac{1}{27}$ at $(\frac{1}{3}, \frac{1}{3})$

4.11 Local min 2 at $(1,2)$; local max 38 at $(-1,-2)$; saddles at $(1,-2)$ and $(-1,2)$

4.13 $2/\sqrt{3}$

4.15 Maximum 4 at $(0, \pm 2)$; minimum -1 at $(1, 0)$

4.17 Maximum 2 at $(1, \sqrt{2})$; minimum 0 on the edges $x = 0$ and $y = 0$

4.19 Maximum 6 at $(2,2,2)$; minimum -6 at $(-2,-2,-2)$

4.21 Base 4×4 m, height 2 m

4.23 $2ab$

4.25 Maximum 4 along $\pm\frac{1}{\sqrt{2}}\langle 1, 1\rangle$; minimum 2 along $\pm\frac{1}{\sqrt{2}}\langle 1, -1\rangle$

CHAPTER 5

MULTIPLE INTEGRALS

5.1 $52/3$

5.3 $\frac{5}{2} + \frac{1}{2}e^{-6}$

5.5 $23/84$

5.7 $\frac{2}{9}(2\sqrt{2} - 1)$

5.9 $\frac{1}{2}(e - 1)$

5.11 $\pi/12$

5.13 $\pi/10$

5.15 $\frac{1}{2}Ma^2$

5.17 4

5.19 $(0, 0, \frac{3}{8}a)$

5.21 162π

5.23 $15\pi/16$

5.25 $\frac{2}{3}Ma^2$

5.27 $\frac{1}{4}(e^6 - 7)$

5.29 $\Gamma(\frac{7}{2}) = \frac{15}{8}\sqrt{\pi}$; the integral is 1

5.31 $24a^5$; $e\sqrt{\pi}$

5.33 $5e^{-2} \approx 0.677$

5.43 $\pi/4 \approx 0.785$, $\pi/6 \approx 0.524$, about 0.0025; the fraction tends to 0

5.45 Normalization constant $1/(32\pi a^5)$; $P(z > 0) = \frac{1}{2}$; $\langle r \rangle = 5a$

CHAPTER 6

VECTOR FIELDS AND LINE INTEGRALS

6.1 Flow lines are the hyperbolas $x^2 - y^2 = C$ (and the lines $y = \pm x$)

6.3 $\nabla r = \mathbf{r}/r = \hat{\mathbf{r}}$, the outward unit radial field

6.5 $\frac{2\sqrt{2}\,\pi}{3}(3 + 4\pi^2)$

6.7 $27/28$

6.9 0

6.11 Flux π; circulation 2π

6.13 Conservative; $f = ye^x + x\sin y$

6.15 $f = xe^y + ye^z$; the integral is 2

6.17 $\nabla \cdot \mathbf{F} = yz$; $\nabla \times \mathbf{F} = \langle -x^2,\, 3xy,\, -xz \rangle$

6.19 No: its divergence is $1 \neq 0$, but $\nabla \cdot (\nabla \times \mathbf{G}) = 0$ always

6.21 $\nabla^2(r^n) = n(n+1)r^{n-2}$; harmonic for $n = 0$ and $n = -1$

6.23 -8π

6.25 $3\pi/8$

6.27 $3\pi/2$

6.39 $(0, \frac{4}{3\pi})$

CHAPTER 7

SURFACE INTEGRALS AND THE INTEGRAL THEOREMS

7.1 Plane: $z = (6 - x - 2y)/3$ over $x, y \geq 0$, $x + 2y \leq 6$. Cone: $(\rho\cos\phi, \rho\sin\phi, \rho)$, $1 \leq \rho \leq 3$

7.3 $3\sqrt{14}$

7.5 $2a^2(\pi - 2)$ (both hemispheres)

7.7 8π

7.9 $3\sqrt{14}$

7.11 $\frac{4}{3}\pi a^3$

7.13 -18π

7.15 80π

7.17 24

7.19 $\frac{12}{5}\pi a^5$

7.21 With $Q = \pi\rho_0 R^3/3$: $\mathbf{E} = \dfrac{\rho_0 r(4R - 3r)}{12\varepsilon_0 R}\hat{\mathbf{r}}$ for $r \leq R$, and $\mathbf{E} = \dfrac{Q}{4\pi\varepsilon_0 r^2}\hat{\mathbf{r}}$ for $r \geq R$

7.23 $|\mathbf{b}|^2 + \mathbf{a}\cdot\mathbf{b}$; 4π

7.25 $\nabla \cdot \mathbf{F} = (n+2)r^{n-1}$; divergence-free for $n = -2$

7.27 $h_u = h_v = \sqrt{u^2 + v^2}$, $h_\phi = uv$; $dV = uv(u^2 + v^2)\,du\,dv\,d\phi$

CHAPTER 8

COMPLEX NUMBERS AND COMPLEX ANALYSIS

8.1 $11 - 10i$; $-\frac{1}{5} + \frac{2}{5}i$; 32

8.3 $2e^{-i\pi/3}$; $2\sqrt{2}\,e^{-3i\pi/4}$; $(1 - i\sqrt{3})^9 = -512$

8.5 $e^{ik\pi/3}$, $k = 0, \dots, 5$

8.7 $\sin^5\theta = \frac{1}{16}(10\sin\theta - 5\sin 3\theta + \sin 5\theta)$

8.9 $4\cos(\omega t + \pi/3)$

8.11 $z = \ln 2 + i(2k+1)\pi$; $z = 2\pi k \pm i\ln\left(2+\sqrt{3}\right)$

8.13 Complex-differentiable exactly on the line $y = x$; analytic nowhere

8.15 $v = e^x \sin y + 2xy$; $u + iv = e^z + z^2$

8.17 0; $2\pi i$; $2\pi i$

8.19 π

8.21 Simple pole at 1, residue $\frac{1}{9}$; double pole at -2, residue $-\frac{1}{9}$

8.23 $-4\pi i$

8.25 $\pi/2$

8.27 $2\pi/3$; $4\pi/(3\sqrt{3})$

8.29 πe^{-2}

CHAPTER 9

LINEAR ALGEBRA AND HILBERT SPACE

9.1 Dependent; the span has dimension 2

9.3 Coordinates $(5, -7, 5)$

9.5 $\ker A = \operatorname{span}\{\langle -2, 1, 0\rangle\}$; image spanned by $\langle 1, 2, 3\rangle$, $\langle 1, 3, 4\rangle$; $1 + 2 = 3$

9.7 $\langle \mathbf{u}, \mathbf{v}\rangle = 5 - 5i$, $\langle \mathbf{v}, \mathbf{u}\rangle = 5 + 5i$, $\|\mathbf{u}\| = \sqrt{6}$, $\|\mathbf{v}\| = \sqrt{11}$

9.9 $\frac{1}{\sqrt{2}}\langle 1, 1, 0\rangle$, $\frac{1}{\sqrt{6}}\langle 1, -1, 2\rangle$, $\frac{1}{\sqrt{3}}\langle -1, 1, 1\rangle$

9.11 $\frac{3}{5}x$

9.13 $\lambda = 1$: $\langle 0, -2, 1\rangle$; $\lambda = 2$: $\langle 1, 0, 0\rangle$; $\lambda = 11$: $\langle 0, 1, 2\rangle$

9.15 Only eigenvalue 1, with a one-dimensional eigenspace spanned by $\langle 1, 0\rangle$

9.17 σ_y: ± 1 with $\frac{1}{\sqrt{2}}\langle 1, \pm i\rangle$; $\sigma_x + \sigma_z$: $\pm\sqrt{2}$

9.19 $\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$, a clockwise rotation by angle t

9.21 Eigenvalues ± 1

9.23 $P = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

9.25 $\langle f|g\rangle = -2\pi i$, $\|f\| = \sqrt{2\pi}$, $\|g\| = \sqrt{2\pi^3/3}$

CHAPTER 10

ORDINARY AND PARTIAL DIFFERENTIAL EQUATIONS

10.1 $y = \dfrac{2}{2 - x^2}$, on $-\sqrt{2} < x < \sqrt{2}$

10.3 $y = \dfrac{\sin x - x\cos x + C}{x^2}$

10.5 About 22.4 minutes

10.7 $y = (1 + 3x)e^{-2x}$

10.9 $y = c_1 e^x + c_2 e^{-x} + \frac{1}{3}e^{2x}$

10.11 $y = c_1\cos x + c_2\sin x - \cos x \ln|\sec x + \tan x|$

10.13 $\gamma = 5$; $x = (1 + 5t)e^{-5t}$

10.15 $\mathbf{x} = \frac{1}{2}\langle e^{3t} + e^{-t},\ e^{3t} - e^{-t}\rangle$

10.17 $\omega^2 = \frac{k}{m}(2 - \sqrt{2})$, $\frac{2k}{m}$, $\frac{k}{m}(2 + \sqrt{2})$, with modes $\left\langle 1, \sqrt{2}, 1\right\rangle$, $\langle 1, 0, -1\rangle$, $\left\langle 1, -\sqrt{2}, 1\right\rangle$

10.19 $a_0\left(1 - \frac{x^2}{2} - \frac{x^4}{8} - \frac{7x^6}{240} - \cdots\right) + a_1\left(x + \frac{x^3}{6} + \frac{x^5}{24} + \frac{x^7}{112} + \cdots\right)$

10.21 Verified: $\int (4x^2 - 2)^2 e^{-x^2} dx = 8\sqrt{\pi}$

10.23 $x^3 = \frac{3}{5}P_1(x) + \frac{2}{5}P_3(x)$

10.25 $\lambda_n = \left(\frac{(2n-1)\pi}{2}\right)^2$, $y_n = \sin\frac{(2n-1)\pi x}{2}$, $n \geq 1$

10.27 $\frac{8}{\pi^3}e^{-\pi^2 t}\sin\pi x + \frac{8}{27\pi^3}e^{-9\pi^2 t}\sin 3\pi x + \cdots$

10.29 $u = V_0 a^2\cos\theta/r^2$

CHAPTER 11

FOURIER ANALYSIS

11.1 $|x| = \dfrac{\pi}{2} - \dfrac{4}{\pi}\displaystyle\sum_{n\ \text{odd}} \dfrac{\cos nx}{n^2}$; $\sum_{n\ \text{odd}} 1/n^2 = \pi^2/8$

11.3 $\sin^2 x = \tfrac{1}{2} - \tfrac{1}{2}\cos 2x$

11.5 Sine: $\sum \frac{2(-1)^{n+1}}{n\pi}\sin n\pi x$. Cosine: $\frac{1}{2} - \frac{4}{\pi^2}\sum_{n\ \text{odd}} \frac{\cos n\pi x}{n^2}$

11.7 $\pi^2/8$

11.9 $\pi^2/12$

11.11 $-ik\,e^{-k^2/2}$

11.13 $\sqrt{2/\pi}\,(1-\cos k)/k^2$; the triangle is the box of half-width $\frac{1}{2}$ convolved with itself

11.15 $\sigma_x = \frac{1}{2\sqrt{a}}$, $\sigma_k = \sqrt{a}$, product $\frac{1}{2}$

11.17 e^{2ik}; $\frac{1}{2}e^{-3ik}\big[\delta(k-1)+\delta(k+1)\big]$

11.19 $u = \dfrac{1}{\sqrt{1+4t}}\,e^{-x^2/(1+4t)}$

11.21 $v_g = \dfrac{\hbar k c^2}{\hbar\omega}$, $v_p = \dfrac{\omega}{k}$; $v_g v_p = c^2$, and $v_p > c$ forces $v_g < c$

11.35 $\langle k\rangle = k_0$; $\langle k^2\rangle = k_0^2 + \dfrac{1}{4\sigma^2}$

CHAPTER 12

THE BREAKDOWN OF CLASSICAL PHYSICS

12.1 650 nm: 1.91 eV, 1.02×10^{-27} kg m/s. 0.10 nm: 12.4 keV, 6.63×10^{-24} kg m/s

12.3 About 9.3 μm, in the infrared

12.5 $x \approx 2.821$; for the wavelength peak $x \approx 4.965$

12.7 539 nm; 1.83 V

12.9 $\lambda' = 0.0536$ nm; electron kinetic energy ≈ 1.68 keV

12.11 656.1, 486.0, 433.9 nm; series limit 364.5 nm

12.13 $v_1 = 2.19\times10^6$ m/s $= \alpha c$, with $1/\alpha = 137.04$

12.15 $a \approx 2.85\times10^{-13}$ m; $E_1 \approx -2.53$ keV

12.17 (a) 3.88 pm; (b) 3.70 pm

12.19 Maximum 2, minimum 0

12.21 $\cos^2(\theta/2)$: $\frac{3}{4}$ at 60°; 0 at 180°

CHAPTER 13

THE WAVEFUNCTION AND THE SCHRÖDINGER EQUATION

13.1 $A = \sqrt{\lambda}$; $\langle x\rangle = 0$, $\langle x^2\rangle = \frac{1}{2\lambda^2}$, $\sigma_x = \frac{1}{\sqrt{2}\,\lambda}$

13.3 $\langle x\rangle = \frac{a}{2}$, $\langle x^2\rangle = \frac{2a^2}{7}$, $\sigma_x = \frac{a}{2\sqrt{7}}$, $\langle p\rangle = 0$, $\langle p^2\rangle = \frac{10\hbar^2}{a^2}$; $\sigma_x\sigma_p = \hbar\sqrt{5/14} \approx 0.598\hbar$

13.5 $J = \frac{\hbar k}{m}\big(|A|^2 - |B|^2\big)$: right-moving minus left-moving flux

13.7 $\Phi(p) = \sqrt{\frac{2}{\pi}}\,\dfrac{(\hbar\lambda)^{3/2}}{p^2+\hbar^2\lambda^2}$

13.9 About 1 MeV (taking $\sigma_p \approx \hbar/r$ gives 0.83 MeV)

13.11 About 58 nm after 1 ps; about 58 μm after 1 ns

13.13 $\sigma(t)\sigma_p = \frac{\hbar}{2}\sqrt{1+\big(\hbar t/2m\sigma_0^2\big)^2}$

13.15 $\omega = (E_2 - E_1)/\hbar$

13.17 $E_n = \dfrac{n^2\pi^2\hbar^2}{2mL^2}$

CHAPTER 14

STATIONARY STATES IN ONE DIMENSION

14.1 9.40, 37.6, 84.6 eV; $\lambda_{31} = 16.5$ nm

14.3 $\langle x\rangle = \frac{a}{2} - \frac{16a}{9\pi^2}\cos\omega t$, with $\omega = \frac{3\pi^2\hbar}{2ma^2}$

14.5 $32/(9\pi^2) \approx 0.360$

14.9 $\psi_2 = \big(\frac{m\omega}{\pi\hbar}\big)^{1/4}\frac{1}{\sqrt{2}}(2\xi^2-1)e^{-\xi^2/2}$

14.11 $3\left(\frac{\hbar}{2m\omega}\right)^{3/2}$; $\sqrt{6}\,\frac{\hbar}{2m\omega}$

14.13 $1-\operatorname{erf}(1)\approx 0.157$

14.15 $E=-6.56$ eV; penetration depth $1/\kappa = 0.076$ nm

14.17 $\langle x^2\rangle = \frac{1}{2\kappa^2}$, $\sigma_p = \hbar\kappa$, $\sigma_x\sigma_p = \hbar/\sqrt{2} > \hbar/2$

14.19 $z_0 \approx 2.60$: two bound states

14.21 Exact $T \approx 0.153$; thick-barrier approximation 0.164 ($\kappa L \approx 1.54$ is not large)

CHAPTER 15

THE FORMALISM

15.1 Results 1 and 3, each with probability $\frac{1}{2}$; $\langle A\rangle = 2$, $\sigma_A = 1$

15.3 Eigenvalue 1 (twice) and -1; $P(1) = \frac{11}{12}$ with post-measurement state $\frac{1}{\sqrt{22}}\langle 2,3,3\rangle$; $P(-1) = \frac{1}{12}$ with state $\frac{1}{\sqrt{2}}\langle 0,-1,1\rangle$

15.5 $[\hat{H},\hat{x}] = -\frac{i\hbar}{m}\hat{p}$, $[\hat{H},\hat{p}] = i\hbar V'(\hat{x})$

15.7 $\sigma_{S_x}\sigma_{S_z} = \frac{\hbar^2}{8}|\sin 2\theta|$; zero for eigenstates of S_x or S_z

15.9 $\langle x\rangle = \langle x\rangle_0 + \langle p\rangle_0 t/m$; $\langle p\rangle$ constant

15.11 $E \le \hbar\omega/\sqrt{2} \approx 0.707\,\hbar\omega$

15.13 $E_n^1 = V_0/2$ for every n

15.15 $E_0^2 = -\frac{11}{8}\frac{\epsilon^2\hbar^2}{m^3\omega^4}$

15.17 $E_\pm = \frac{E_1+E_2}{2} \pm \sqrt{\left(\frac{E_2-E_1}{2}\right)^2 + V^2}$; to second order $E_1 - \frac{V^2}{E_2-E_1}$ and $E_2 + \frac{V^2}{E_2-E_1}$

CHAPTER 16

THE HYDROGEN ATOM

16.1 In units of $\frac{\pi^2\hbar^2}{2ma^2}$: 3 (1), 6 (3), 9 (3), 11 (3), 12 (1), 14 (6); none is accidental (the first accidental degeneracy is at 27)

16.5 $\sin^2\theta\cos 2\phi = \sqrt{\frac{8\pi}{15}}\left(Y_2^2 + Y_2^{-2}\right)$

16.7 $Y_2^0 = \sqrt{\frac{5}{16\pi}}\left(\frac{3z^2}{r^2}-1\right)$; $\frac{1}{\sqrt{2}}(Y_2^2 + Y_2^{-2}) = \sqrt{\frac{15}{16\pi}}\frac{x^2-y^2}{r^2}$

16.11 $2s$: $\langle r\rangle = 6a$, $\langle r^2\rangle = 42a^2$; $2p$: $5a$, $30a^2$

16.13 $\langle V\rangle = -27.2$ eV, $\langle T\rangle = 13.6$ eV

16.15 H: 656.47 nm; D: 656.29 nm; difference 0.18 nm

16.17 Allowed: $3s\to 2p$, $3p\to 1s$, $4f\to 3d$. Forbidden: $3d\to 2s$, $2s\to 1s$

16.19 $\langle H\rangle = 0.4E_1 + 0.6E_2 = -7.48$ eV; $\langle L^2\rangle = 1.2\hbar^2$; $\langle L_z\rangle = -0.1\hbar$

CHAPTER 17

ANGULAR MOMENTUM AND SPIN

17.5 Direct multiplication gives $\hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x = i\hbar\hat{J}_z$

17.7 $\langle L_x\rangle = 0$, $\langle L_x^2\rangle = \frac{\hbar^2}{2}\left[\ell(\ell+1) - m^2\right]$; $\sigma_{L_x}\sigma_{L_y} = \frac{\hbar^2}{2}\left[\ell(\ell+1)-m^2\right] \ge \frac{\hbar^2}{2}|m|$

17.9 $\chi = \frac{1}{5}\langle 3i,4\rangle$; $\langle\hat{\mathbf{S}}\rangle = \frac{\hbar}{2}\left\langle 0,-\frac{24}{25},-\frac{7}{25}\right\rangle$, pointing along $\langle 0,-24,-7\rangle/25$

17.11 $\cos^2 60^\circ = \frac{1}{4}$

17.13 Electron: 1.40 MHz; proton: 128 MHz

17.15 $|\frac{3}{2},\frac{3}{2}\rangle = |1,\uparrow\rangle$; $|\frac{3}{2},\frac{1}{2}\rangle = \sqrt{\frac{2}{3}}|0,\uparrow\rangle + \sqrt{\frac{1}{3}}|1,\downarrow\rangle$; $|\frac{3}{2},-\frac{1}{2}\rangle = \sqrt{\frac{2}{3}}|0,\downarrow\rangle + \sqrt{\frac{1}{3}}|-1,\uparrow\rangle$; $|\frac{3}{2},-\frac{3}{2}\rangle = |-1,\downarrow\rangle$; $|\frac{1}{2},\frac{1}{2}\rangle = \sqrt{\frac{2}{3}}|1,\downarrow\rangle - \sqrt{\frac{1}{3}}|0,\uparrow\rangle$; $|\frac{1}{2},-\frac{1}{2}\rangle = \sqrt{\frac{1}{3}}|0,\downarrow\rangle - \sqrt{\frac{2}{3}}|-1,\uparrow\rangle$

17.17 $2p_{3/2}$: 4 states; $2p_{1/2}$: 2 states; 6 in total

CHAPTER 18

TWO-LEVEL SYSTEMS, IDENTICAL PARTICLES, AND ENTANGLEMENT

18.1 $E_0 = 2$, $\mathbf{b} = \langle 1,1,1\rangle$; eigenvalues $2 \pm \sqrt{3}$; $\omega = 2\sqrt{3}/\hbar$

18.3 $P_{\max} = \tfrac{1}{2}$, first reached at $t = \pi/\sqrt{\Omega^2 + \Delta^2} \approx 0.35\ \mu\text{s}$

18.5 About 990 km

18.7 $p_{1,2} = \tfrac{1}{2} \pm \tfrac{\sqrt{5}}{6} \approx 0.873,\ 0.127$; entanglement entropy ≈ 0.55 bit

18.9 $\operatorname{tr}\hat{\rho}^2 = \tfrac{1}{2}(1 + |r|^2)$; Bloch vector $\langle \operatorname{Re} r, -\operatorname{Im} r, 0\rangle$; pure iff $|r| = 1$

18.11 Energy $5\pi^2\hbar^2/2ma^2$ in all cases. $\langle (x_1 - x_2)^2\rangle \approx 0.103a^2$ (distinguishable), $0.038a^2$ (bosons), $0.168a^2$ (fermions)

18.15 $\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

INDEX

A

addition of angular momenta, 321
adjoint, 173
Aharonov–Bohm effect, 109
analytic functions are harmonic, 150
angular momentum, 14, 317
Angular momentum commutators, 316
antilinear, 170
arc length, 31
arc length function, 32
Area from a boundary integral, 112
Argand diagram, 145
associated Legendre equation, 302
associated Legendre functions, 302
axioms of a vector space, 6
azimuthal angle, 18

B

beat frequency, 192
Bell states, 332
Bessel's inequality, 209
binormal, 34
Bloch sphere, 19, 319
Bohr energy levels, 232
Bohr frequency, 255
Bohr magneton, 235
Born rule, 89
Born's rule, 246
Bose–Einstein and Fermi–Dirac integrals, 350
bound and scattering states, 270
bound state, 270
bound state of the delta well, 270
bound states of the finite well, 271
branch cut, 149

C

canonical quantization, 288
Cauchy integral formula, 152
Cauchy's theorem, 151
Cauchy–Riemann equations, 149
Cauchy–Schwarz in inner-product spaces, 170
Cauchy–Schwarz inequality, 8
Cayley form, 359
centrifugal barrier, 302
Chain rule, 53
change of variables, 92
characteristic equation, 190
characteristic polynomial, 172
characterization of conservative fields, 107
CHSH inequality, 335
circulation, 105
Clairaut–Schwarz, 50
classical electron radius, 231
classical limit, 250
Clebsch–Gordan coefficients, 322
coherent states, 282
commutator, 175
complete set of commuting observables, 285
Completeness and Parseval's identity, 209
complex amplitude, 234
complex conjugate, 145
Complex derivative, 149
Complex exponential, 146
complex function, 148
complex number, 145
Complex numbers, 145

complex plane, 145
complex-differentiable, 149
component functions, 29
Compton formula, 230
Compton wavelength, 230
computing residues, 154
conformal, 164
Conservation laws, 287
Conservation of angular momentum, 37
Conservation of energy, 108
conservative, 103
constant-coefficient solutions, 190
Continuity equation, 129
Continuity equation for probability, 247
continuous spectra, 178
continuously differentiable, 52
contour integral, 150
contour map, 47
controlled-NOT, 336
convolution, 213
convolution theorem, 213
coordinate planes, 3
Copenhagen interpretation, 246
correspondence principle, 232
cosine series, 208
critical point, 69
critically damped, 191
Cross product, 11
Curl as circulation density, 126
curvature, 33
curvature formulas, 33
cylindrical coordinates, 17

D

Darboux vector, 44
De Moivre, 146
decoherence, 333
defective, 173
delta function, 129
delta function of a point charge, 130
density matrix, 332
diagonalization, 172
differentiability, 51
differentiable implies continuous, 51
Differential operators in orthogonal coordinates, 132
differentiation rules, 30
differentiation under the integral sign, 94
dimension, 167
Dirac delta, 129
Dirac delta, informal, 129
direction angles, 9
direction vector, 14
directional derivative, 54
directional derivatives from the gradient, 54
Dirichlet's theorem, 208
discrete Fourier transform, 215
dispersion relation, 214
distance formula, 3
distance formulas, 15
Divergence and curl, 109
Divergence as flux density, 128
divergence theorem, 121
Divergence theorem (Gauss), 127
dot product, 7
double integral, 84

E

effective potential, 302
Ehrenfest, 250
eigenspace, 172
eigenvalue, 172
eigenvalue, eigenvector, 172
eigenvector, 172
Einstein summation convention, 13
electric dipole, 307
energy eigenfunction, 253
energy eigenvalue, 253
energy levels of hydrogen, 304
entangled, 329, 332
entanglement entropy, 333
equipartition, 225
equipotential surfaces, 104
essential singularity, 154
Euler's formula, 146
Euler's theorem, 64

even-tempered, 358
exchange effect, 335
exchange lemma, 167
existence and uniqueness (Picard–Lindelöf), 189
Expansion in an orthonormal basis, 171
Expansion in stationary states, 254
expectation value, 248
exponentials of Hermitian operators are unitary, 176
extreme value theorem, 71
extremes of a quadratic form, 75

F

fast Fourier transform, 215
Fermat, 69
field line, 104
fine structure, 307
fine-structure constant, 242
first octant, 3
flow line, 104
flow lines, 104
flux across a plane curve, 106
flux integral, 124
Fourier inversion theorem, 210
Fourier series, 207, 351
Fourier transform, 210
Fourier uncertainty principle, 214
Frenet frame, 34
Frenet–Serret formulas, 34
Fubini, 84
Function of several variables, 47
fundamental integral, 151
Fundamental theorem for line integrals, 106
Fundamental theorem of algebra, 153
Fundamental theorem of space curves, 35

G

Gamma function, 93
Gauss's law for a point charge, 128
Gaussian integral, 87
Gaussian integrals, 350
Gaussian moments, 93
Gaussian transforms to a Gaussian, 210
Geiger–Nuttall law, 275
generalized Stokes theorem, 134
geometric form of the dot product, 7
Geometry of the gradient, 55
Gibbs phenomenon, 209
global phase, 246
gradient, 46
gradient field, 103
Gradient field, potential, 103
Gram–Schmidt, 171
Green, 111
Green's first identity, 119
Green's second identity, 119
grid curves, 122
group velocity, 215
gyromagnetic ratio, 319

H

half-life, 189
harmonic conjugate, 150
harmonic oscillator, 191
heat kernel, 65, 213
Heisenberg equation of motion, 288
Heisenberg picture, 288
Heisenberg uncertainty principle, 214, 250
Hellmann–Feynman theorem, 298
Hermite polynomials, 193
Hermitian operators, 174
hermitian, unitary, normal, 173
Hessian, 67
Hessian matrix, 67
Hilbert space, 170
homogeneous of degree n, 64
Hydrogen wavefunctions, 305
hyperfine structure, 307

I

imaginary part, 145
Implicit function theorem, 56
index notation, 13
indicial equation, 194
induced orientation, 125
Infinite square well, 264

infinite-dimensional, 168
initial value problem, 189
inner product, 7, 170
integral curve, 104
integrating factor, 189
intercept form, 15
interference pattern, 234
Invariance of length, 32
Inverse function theorem, 56
isolated singularities and residues, 154

J
Jacobian matrix, 53
joint density, 88
joint probability density, 88
Jordan, 156
jump condition, 270

K
Kepler's first law, 37
Kepler's third law, 38
kronecker delta and Levi-Civita symbol, 13

L
ladder algebra, 266
ladder operators, 266, 317
Lagrange multiplier, 72
Lagrange multipliers, 72
Lamb shift, 307
Laplace–Runge–Lenz vector, 38
Larmor frequency, 320
Larmor precession, 320
Laurent series, 153
leading principal minors, 69
least squares line, 74
Legendre polynomials, 171, 194
level curve, 47
level surfaces, 47
limit and continuity, 29
limits are computed componentwise, 29
Line integral of a vector field, 105
Line integral with respect to arc length, 104
linear combination, 6
Linear map, 168
linear operator, 168
Linear second-order equation, 190
linearization, 52
linearly independent, 167
Liouville, 153
local and absolute extrema, 69
local hidden variables, 335
local maximum, 69
logarithm, 148

M
magnetic quantum number, 302
Mandelstam–Tamm, 286
matrix element, 177
matrix mechanics, 236
mean value property, 164
mean value theorem for double integrals, 86
mixed state, 187
Möbius strip, 124
mode density, 226
modulus, 145
moments of inertia, 88
Momentum is Hermitian, 178
momentum-space wavefunction, 249
monkey saddle, 71

N
natural linewidth, 287
negative definite, 68
no-cloning, 336
Normal and binormal, 34
normal distribution, 88
normal equations, 74
normal line, 55
normal modes, 192
normal vector, 15
normal Zeeman effect, 322
normalization, 6
normalization is preserved, 247
number operator, 267

O
observable, 249
octants, 3
ordinary differential equation, 189
ordinary point, 193
orientation, 124
orthogonal decomposition, 9
orthogonal projection, 172
orthonormal, 171
orthonormal set, 171
oscillator eigenfunctions, 268
osculating circle, 34
osculating plane, 34
outer product, 177
overdamped, 191

P
parallelogram law, 5
parametric equations, 14, 29
parametric surface, 122
Parseval's identity, 171
Partial derivative, 49
partial differential equation, 50
Partial trace, 333
partition function, 73
path independence, 107
path independent, 107
Pauli exclusion principle, 334
penetration depth, 272
perturbation, 290
phase velocity, 215
phasor, 147
piecewise smooth, 104
Plancherel, 212
Planck's mean energy, 227
Planck's radiation law, 227
Planck–Einstein relations, 229
planimeter, 112
polar angle, 18
polar coordinates, 86
polar form, 146
polar rectangle, 86
polarization identity, 26, 186
pole of order m, 154
position vector, 5
positive (negative) semidefinite, 68
positive definite, 68
positive orientation, 111
positively oriented, 124
Postulates of quantum mechanics, 284
potential energy, 103
potential function, 103
principal argument, 146
principal logarithm, 148
principal part, 154
principal quantum number, 304
principal unit normal, 34
principle of least action, 75
probability current, 129, 247
Product rules, 110
product state, 332
projection, 9
properties of conjugate and modulus, 145
properties of determinants, 169
properties of stationary states, 253
properties of the cross product, 11
properties of the Fourier transform, 211
pure state, 333

Q
quadratic form, 68
quadratic form and definiteness, 68
quadric surface, 16
quantum tunneling, 274
Quantum violation, 335

R
Rabi frequency, 331
Radial expectation values, 306
radius of curvature, 33
Rank–nullity, 169
Rayleigh quotient, 75
Rayleigh–Jeans law, 227
real part, 145
real-valued function of n variables, 47
rectangular barrier, 273
recurrence relation, 193
reduced density matrix, 333
reduced mass, 307

regular singular point, 194
removable singularity, 154
residue, 154
residue theorem, 154
revivals, 265
right-handed system, 3
Robertson uncertainty relation, 285
roots of unity, 147
rotating-wave approximation, 330

S

saddle point, 70
sample point, 84
scalar component, 9
scalar product, 7
Scalar triple product, 12
scalars, 5
scale factors, 131
scanning tunneling microscope, 274
Scattering by the delta well, 270
scattering state, 270
Schmidt decomposition, 332
Schmidt rank, 332
Schrödinger equation, 245
Schrödinger picture, 288
second derivative test, 70
second-derivative identities, 110
second-order Taylor formula, 67
selection rule, 267
self-adjoint, 173
sensitivity, 73
shadow price, 73
shooting method, 356
sifting property, 130
simple pole, 154
simply connected, 109
simply connected regions, 113
simultaneous diagonalization, 175
sine series, 208
singlet and triplet, 321
skew, 14
Slater determinant, 334
smoothness and decay, 209
solid angle, 90, 123
space curve, 29
spectral decomposition, 175
spectral energy density, 226
spectral theorem, 174
Spectrum of angular momentum, 317
Spectrum of the oscillator, 266
Spherical coordinates, 132
spherical coordinates, 18
Spherical coordinates, physics convention, 18
Spherical harmonics, 303, 353
spherical harmonics, 133
spherical shells, 94
Spin along an arbitrary axis, 319
spin–orbit coupling, 307
spin–statistics theorem, 334
spinor, 318
split-operator, 361
spreading of a Gaussian, 251
standard basis vectors, 6
stationary state, 253
Stefan–Boltzmann and Wien laws, 227
step potential, 273
Stern–Gerlach probabilities, 319
Stokes, 125
Stokes' theorem, 121
stopping potential, 229
structure of solutions, 190
Sturm–Liouville, 195
Sturm–Liouville problem, 195
subspace, 167
subspace, span, independence, basis, 167
superposition principle, 167, 245
surface area, 123
surface integral of a scalar function, 123
surface of revolution, 122
Sylvester's criterion and the eigenvalue test, 69
symmetric equations, 14
symmetrization postulate, 334
symmetry implies conservation, 288

T

tangent plane, 52
tangent plane to a level surface, 55

tangent vector, 30
tangential and normal components of acceleration, 36
Taylor series, 153
teleportation, 337
tensor product, 320, 332
termination and growth, 193
test for conservative fields, 109
threshold frequency, 229
time dependence of expectation values, 287
time-independent Schrödinger equation, 253
torsion, 34
total differential, 52
trace and determinant, 172
trial functions, 289
triangle inequality, 8
triangle law, 5
two-flavor oscillation, 331
two-level spectrum and dynamics, 330
two-path test, 48
two-variable criterion, 68

U

ultraviolet catastrophe, 227
underdamped, 191
unit tangent vector, 30
unit vector, 5
Unitary operators, 174

V

Variational principle, 289
variational principle, 76
Vector, 4
vector field, 103
Vector function, 28
Vector operations, 5
vector projection, 9
vector space, 167
Vector space, 167
Vector triple product, "BAC–CAB", 13
vector-space properties, 5
vector-valued function, 28
velocity of the mean, 248
virial theorem, 261
Volume elements, 90

W

wave mechanics, 236
wave packet, 214
Wigner function, 262
winding number, 113
work function, 229
Wronskian, 190

Z

zero-point energy, 251