Quantum Physics: A Text for Graduate Students

Roger G. Newton

Springer

Graduate Texts in Contemporary Physics

Series Editors:

Springer
New York
Berlin
Heidelberg
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

This page intentionally left blank

Roger G. Newton

Quantum Physics

A Text for Graduate Students

With 27 Figures

Springer

Roger G. Newton
Department of Physics
Swain Hall West 117
Indiana University
Bloomington, IN 47405
USA
newton@indiana.edu

Library of Congress Cataloging-in-Publication Data
Newton, Roger G.
Quantum physics : a text for graduate students / Roger G. Newton.
p. cm. — (Graduate texts in contemporary physics)
Includes bibliographical references and index.
ISBN 0-387-95473-2 (alk. paper)
1. Quantum theory. I. Title. II. Series.
QC174 .12 .N49 2002
530.12—dc21 2002020945

ISBN 0-387-95473-2 Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 10874003

Typesetting: Pages created by the author using a Springer TEX macro package.

www.springer-ny.com

Springer-Verlag New York Berlin Heidelberg
A member of BertelsmannSpringer Science+Business Media GmbH

To Daniel,

in the hope that he, too, will grow up to love science

This page intentionally left blank

Preface

The combination of quantum mechanics and quantum field theory constitutes the most revolutionary and influential physical theory of the twentieth century. Its impact is felt not only in almost all other sciences, but the fruits of its application are ubiquitous in everyday life. This textbook is designed to teach graduate students the underlying quantum-physical ideas, their mathematical formulations, and the basic problem-solving techniques of the discipline. It assumes they have taken at least one introductory course in quantum mechanics as undergraduates and are familiar with the history of the subject and the basic experimental evidence that led to its adoption, as well as with many of its fundamental notions. In contrast to most other authors, I am therefore not introducing the quantum theory via an historical survey of its early successes. Instead, following the models of books on classsical mechanics or electromagnetism, I develop the theory from its basic assumptions, beginning with statics, followed by the dynamics and details of its specific areas of use as well as the needed mathematical techniques.

Although this book, inevitably, deals largely with the behavior of point particles under various conditions, I do not regard particles as the fundamental entities of the universe: the most basic object is the quantum field, with the observed particles arising from the field as its quanta. For this reason I introduce quantum fields right from the beginning and demonstrate, in the first chapter, how particles originate. However, this volume is not intended to be a full-fledged text of quantum field theory; confining itself to the fundamental ideas of field theory and their consequences, glossing over its mathematical difficulties and pitfalls, it does not deal with any of the

subtleties of that large subject. For the development of the basic quantum dynamics, the book employs the Lagrangian technique with the principle of stationary action. The roots of this approach, which includes generating the canonical commutation rules, go back to a course taught long ago by Julian Schwinger, filtered through and modified by many years of teaching the subject. Similarly for the manner of introducing the γ-matrices in Chapter 10.

The general physics preparation of the students for whom this book is intended should comprise classical mechanics (including its Lagrangian and Hamiltonian formulations), Maxwell's theory of electromagnetism, and the special theory of relativity, as well as some understanding of statistical mechanics. As far as mathematics is concerned, they are expected to have a basic knowledge of linear algebra, linear differential equations—both ordinary and partial—and the theory of functions of a complex variable. The book's extensive mathematical appendices contain the needed elements of vector spaces (including Hilbert space), the Dirac delta function, linear integral equations, the required classical special functions, and group representations.

Rather than mixing this mathematical material with the explanation of the physics—though some details that the occasion demands are interspersed, set off by horizontal lines and printed in a smaller font—I collected it at the end, and the instructor will have to present or assign it when needed. These appendices are not meant to take the place of mathematics courses appropriate for graduate students in physics or chemistry; they can serve at best as crutches or reminders. However, the inclusion of group representations allows me to apply this invaluable tool more extensively than do most other comparable graduate texts. A large number of exercises are sprinkled as boldfaced footnotes throughout the text, others collected at the end of each chapter.

A word about notation: Many books and research articles employ the very convenient *Dirac notation*, in which a general state vector Ψ is denoted by $|\,\rangle$, an eigenvector Ψ_A of $\mathbf{A}$ with the eigenvalue A is written $|A\rangle$, and the inner product becomes $(\Psi_A, \Psi) = \langle A|\,\rangle$. Although I shall not always use this notation, when it comes in handy, I shall employ it.

Since the subject of this book now pervades almost all of physics, it is hard to pin down my diffuse indebtedness. I certainly owe much to the four teachers from whom I learned quantum mechanics and quantum field theory as a student: Norman Ramsey, Walter Kohn, John H. Van Vleck, and, above all, Julian Schwinger. Others to whom I owe a debt of gratitude for their infusion of wisdom over the years are Res Jost, Joseph Jauch, and my colleagues at Indiana University, especially Emil Konopinski, John Challifour, and Don Lichtenberg.

Bloomington, Indiana, USA Roger G. Newton
December 2001

Contents

Preface **vii**

I Physics xv

1 Quantum Basics: Statics **1**
1.1 States and Dynamical Variables 1
1.1.1 Measurements 3
1.1.2 Fields as dynamical variables 5
1.2 Probabilities 6
1.2.1 Correlations 7
1.2.2 Interference 9
1.2.3 Expectation values and variance 11
1.3 Mixed States 12
1.4 Representations 16
1.4.1 The configuration representation 19
1.4.2 The momentum representation 22
1.4.3 The number representation 23
1.5 Transformation Theory 27
1.6 Problems and Exercises 32

2 Quantum Basics: Dynamics **33**
2.1 Time Dependence and Commutation Rules 33
2.1.1 The principle of stationary action 34
2.1.2 The Lagrangian equations of motion 36

2.1.3 The canonical commutation relations 37
2.1.4 The Hamiltonian equations of motion 39
2.1.5 The Schrödinger picture 40
2.2 Systems of Particles 43
2.2.1 Linear and angular momentum 43
2.2.2 The equations of motion 44
2.2.3 The time-independent Schrödinger equation 46
2.2.4 Example: Gaussian wave packets 48
2.3 Fields 51
2.3.1 The matter field 51
2.3.2 Infinitely many particles as a field 56
2.3.3 The electromagnetic field 58
2.4 Canonical and Gauge Transformations 63
2.5 Problems and Exercises 66

3 The Schrödinger Equation in One Dimension 67
3.1 A Free Particle 67
3.2 A Particle With Potential Energy 69
3.2.1 The square potential barrier 69
3.2.2 The potential well 74
3.2.3 General transmission and reflection 77
3.2.4 The double well 78
3.2.5 The Kronig-Penney potential 80
3.2.6 The simple harmonic oscillator 82
3.2.7 Coherent states 85
3.3 Problems and Exercises 89

4 One- and Two-Particle Systems in Three Dimensions 91
4.1 Free Particles 91
4.1.1 The propagator 96
4.1.2 Two particles 98
4.2 Potentials with Spherical Symmerty 99
4.2.1 The three-dimensional square well 100
4.2.2 The scattering amplitude 103
4.2.3 Transformation to the laboratory system 110
4.2.4 The Coulomb potential 112
4.2.5 The hydrogen atom 112
4.2.6 Coulomb scattering 117
4.3 The Inverse Problem 121
4.3.1 Solution procedure 122
4.3.2 The phase of the scattering amplitude 125
4.4 Potentials without Spherical Symmetry 127
4.4.1 Green's functions 127
4.4.2 The scattering amplitude 131
4.4.3 Solving the Lippmann-Schwinger equation 134

4.5 Problems and Exercises 137

5 Symmetries **143**
5.1 The Angular-Momentum Operator 143
5.1.1 Active and passive rotations 144
5.1.2 An oscillator model 148
5.1.3 States of spin-1 and spin-1/2 150
5.2 The Rotation Group 151
5.2.1 Angular momentum and the carrier space 155
5.2.2 Polarization and the spin density matrix 160
5.2.3 The magnetic moment 161
5.2.4 Addition of angular momenta 163
5.2.5 Spherical tensors and selection rules 167
5.3 Time Reversal 172
5.3.1 The time-reversal operator 172
5.3.2 Time-reversal invariance 174
5.4 Problems and Exercises 177

6 Stationary Approximation Methods **181**
6.1 The WKB Approximation 181
6.1.1 The connecting formulas 184
6.1.2 Examples 187
6.1.3 Scattering in the WKB approximation 190
6.2 Bound-State Perturbation Theory 195
6.2.1 The nondegenerate case 196
6.2.2 Closely spaced levels 198
6.2.3 The case of degeneracy 199
6.2.4 Second-order shifts from first-order states 203
6.2.5 Effects of symmetry 204
6.2.6 Relativistic correction to hydrogenic levels 205
6.2.7 The Stark effect 206
6.3 Variational Methods 208
6.3.1 Bound states 208
6.3.2 The helium atom 210
6.3.3 Scattering 212
6.4 Problems and Exercises 214

7 Static Magnetic Fields **217**
7.1 The Larmor Precession 217
7.2 The Aharanov-Bohm Effect 221
7.3 Charged Particles with Spin 222
7.3.1 Spin-orbit coupling and the fine structure 223
7.3.2 The Zeeman effect 226
7.3.3 The hyperfine structure 229
7.4 Problems and Exercises 233

8 Time-Dependent Hamiltonians **235**
8.1 Perturbation Theory . . . 235
8.1.1 Application to exponential decay . . . 240
8.1.2 Application to scattering . . . 242
8.1.3 Second order perturbation theory . . . 243
8.1.4 Interaction with electromagnetic fields . . . 244
8.2 Slowly Changing Hamiltonians . . . 253
8.3 Problems and Exercises . . . 256

9 Multiparticle Systems **257**
9.1 Two Particles Interacting with a Center . . . 257
9.2 Identical particles . . . 264
9.2.1 Fock space . . . 269
9.2.2 Bose-Einstein and Fermi-Dirac distributions . . . 276
9.3 Problems and Exercises . . . 278

10 Relativistic Electron Theory **279**
10.1 Preliminaries . . . 279
10.1.1 Relativistic spin . . . 281
10.1.2 The γ-matrices . . . 282
10.2 The Dirac Equation . . . 292
10.2.1 Electrons in an electromagnetic field . . . 299
10.3 The Dirac Equation for a Central Potential . . . 302
10.3.1 The hydrogen atom . . . 305
10.4 Problems and Exercises . . . 308

II Mathematical Appendices **309**

A The Dirac Delta Function **311**

B Hilbert Space **315**
B.1 Linear Vector Spaces . . . 315
B.2 Infinite-Dimensional Linear Vector Spaces . . . 318
B.3 Operators . . . 319
B.3.1 Hermitian operators and their spectra . . . 322
B.3.2 More on operators . . . 328
B.4 Problems and Exercises . . . 332

C Linear Integral Equations **333**
C.1 Fredholm Methods . . . 333
C.2 The Born Series . . . 337
C.3 Application to the Lippmann-Schwinger Equation . . . 338

D Special Functions **341**
D.1 Spherical Harmonics . . . 341

D.1.1 Legendre polynomials . . . 342
D.1.2 Associated Legendre functions . . . 345
D.1.3 Spherical harmonics . . . 345
D.1.4 Zonal harmonics . . . 346
D.2 Spherical Bessel Functions . . . 347
D.3 Hermite Polynomials . . . 348
D.4 The Hypergeometric Function . . . 350
D.5 Laguerre Polynomials . . . 351
D.6 Problems and Exercises . . . 353

E Group Representations **355**
E.1 Groups . . . 355
E.2 Representations . . . 357
E.2.1 Character . . . 361
E.2.2 Real representations . . . 364
E.2.3 Kronecker products . . . 366
E.2.4 Carrier spaces . . . 367
E.2.5 Clebsch-Gordan coefficients . . . 371
E.3 Lie Groups . . . 373
E.3.1 Coordinate transformations . . . 373
E.3.2 Infinitesimal generators . . . 374
E.3.3 The Lie algebra . . . 379
E.4 Representations of Lie Groups . . . 380
E.4.1 Multiple valuedness . . . 385
E.4.2 The group O(2) . . . 386
E.4.3 The group SO(3) . . . 388
E.5 Problems and Exercises . . . 397

References **399**

Index **402**

This page intentionally left blank

Part I

Physics

This page intentionally left blank

1
Quantum Basics: Statics

1.1 States and Dynamical Variables

The aim of quantum physics, like that of classical physics, is to give an explanatory and predictive description of the motions of physical systems: we want to *understand* their actions and be able to *predict* their future behavior. Therefore, the first two tasks (though not necessarily historically the first as the theory developed) are to define the parameters in terms of which such systems are to be described and to lay down the equations of motion of these parameters so that the future course of any given system can be determined. The two are not independent, because a wrong choice in the first will prevent the second from being realized. Thus the most basic decision is how to specify the *state* of any physical system at a given time.

Whereas in classical physics the state of a system, defined as precisely as nature allows, is described by a point in the system's phase space, in quantum physics, such a state (called a *pure* state) is described by a point (i.e., a vector) in an appropriate Hilbert space $\mathfrak{H}$. Just as the precise form of the phase space depends on the details of the system—the number of its degrees of freedom, the range of its dynamical variables, etc.—so does $\mathfrak{H}$. Conversely, every vector in $\mathfrak{H}$ is assumed to describe a possible state of the system,[1] with Ψ and $a\Psi$ denoting the same state, *when in isolation,* for any

[1]There are a few exceptions to this, called *superselection rules,* but they need not concern us for the time being.

complex constant a, which is mathematically expressed by the statement that an isolated state is specified by a *ray* in Hilbert space.[2]

The Hilbert-space description implies that the states are *additive:* if Ψ_1 and Ψ_2 are two states of a given system, $\Psi = \Psi_1 + \Psi_2$ is also a possible state of the system. You should especially note that this *superposition principle*, whose physical significance will emerge shortly, implies that if the two states Ψ_1 and Ψ_2 are combined into another state $\Psi = \Psi_1 + \Psi_2$ by superposition, they are, in Ψ, in some sense *correlated,* and although Ψ_1 and $a\Psi_1$ specify the same state in isolation, the state $\Psi' = a\Psi_1 + \Psi_2$ generally differs from Ψ. (Both the magnitude and the phase of the inner product (Ψ_1, Ψ_2) of the two *entangled* states Ψ_1 and Ψ_2 have observational consequences, as we shall see.) In other words, in order to describe a state "as precisely as nature allows" it has to be specified not just in isolation, but in such a way that its potential relation to all other possible "completely specified" states of the same system is also given, and for this it is not sufficient to describe it as a *ray* in $\mathfrak{H}$; it must be given as a *vector.* In a figurative sense, the vector Ψ "has hair," while the ray spanned by Ψ "is shaved." It is the hair on Ψ_1 and Ψ_2 that describes their correlation and *entanglement* when their superposition forms Ψ.

For example, consider the famous two-slit experiment with electrons. If only slit #1 is open, the electron is in the state Ψ_1; if only slit #2 is open, it is in the state Ψ_2. Both of these states by themselves are equally well described by $a\Psi_1$ and $b\Psi_2$, respectively, with arbitrary a and b. However, if both slits are open, the electron is in the state $\Psi = c_1\Psi_1 + c_2\Psi_2$, with a fixed ratio c_1/c_2: the two states Ψ_1 and Ψ_2 in Ψ are now entangled.

A vector in Hilbert space therefore describes the state of a system not only as perfectly specified in physical isolation but also its readiness for any potential correlation with other systems and states. There is no analogue of such an entanglement for classical particle states, but for wave systems such as light, its analogue is the phase information contained in the full description of light by means of the electromagnetic field, as necessary, for example, for the description of Young's two-slit experiment. (In this sense all quantum systems behave analogously to both "particles" and "waves"—the notorious *particle-wave duality*—though this analogy must not be pushed too far.)

When two physical systems I and II, whose Hilbert spaces are $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{H}^{\mathrm{II}}$, respectively, are combined into one, the Hilbert space of the combined system is the tensor product (see Appendix B) $\mathfrak{H} = \mathfrak{H}^{\mathrm{I}} \bigotimes \mathfrak{H}^{\mathrm{II}}$, and its states are linear combinations of those special tensor product states of the form $\Psi = \Psi^{\mathrm{I}} \otimes \Psi^{\mathrm{II}}$ in which the two subsystems are independent of one another. For example, the Hilbert space of a two-particle system is the

[2] A ray $\mathcal{R}$ is a one-dimensional subspace in $\mathfrak{H}$. Consequently there exists a vector $\Psi \in \mathfrak{H}$ that spans $\mathcal{R}$, so that $\mathcal{R}$ consists of all vectors of the form $a\Psi$.

tensor product of the two Hilbert spaces of its two one-particle subsystems, formed out of linear combinations $\Psi = \sum_{nm} a_{nm} \Psi^{\mathrm{I}}_n \otimes \Psi^{\mathrm{II}}_m$ of special states $\Psi^{\mathrm{I}}_n \otimes \Psi^{\mathrm{II}}_m$ of two independent particles, the first one in state Ψ^{I}_n and the second in Ψ^{II}_m. Whatever correlation there is between the states of the two particles in the combined system is expressed in the coefficients a_{nm}. After the two subsystems cease to interact, they may retain information about their earlier conjunction even when they are far apart, information that expresses itself in a correlation, so that none of these subsystems *by itself* is then specified "as precisely as nature allows"; it is not in a pure state and cannot be described by a state vector. (We shall discuss the appropriate means for its description a little later.)

The radioactive decay of an atomic nucleus in which an electron and an antineutrino are emitted is a typical example. If the original nucleus was in a pure state, so is the state of the combined nucleus-electron-antineutrino system after the decay. However, the state of the electron itself, which is the observationally most easily accessible subsystem, is correlated with those of the recoiling nucleus and the antineutrino and, regarded without attention to the antineutrino and the nucleus, is not pure and cannot be described by a state vector.

The dynamical variables that are observable, such as the locations and momenta of point particles, which are conventionally denoted by the vectors $\vec{q}$ and $\vec{p}$, are quantum-mechanically represented by *Hermitian operators*,[3] such as $\vec{\mathbf{q}}$ and $\vec{\mathbf{p}}$ (boldface letters will be used to denote operators), acting on the vectors in $\mathfrak{H}$. The most fundamental difference between classical and quantum physics arises from the fact that, whereas a classical state at a given time t_0 uniquely determines the values of all dynamical variables at t_0, a quantum state determines some of them only as non-sharp probability distributions. From this follow many of the important differences between classical and quantum physics, and in particular, the probabilistic rather than strictly causal predictions of the latter. The acausal nature of quantum physics arises not from any non-deterministic feature in its dynamics, but from the very definition of the quantum state of a system.

1.1.1 Measurements

When it comes to performing measurements, quantum physics posits three basic postulates, schematically stated as follows:

[3] Mathematicians distinguish between *symmetric* and *self-adjoint* operators, the distinction depending on their domain of definition. The word *Hermitian*, universally customary among physicists, is meant to be essentially equivalent to *self-adjoint*, although we are not always as careful about the domain of definition as we should be. See Appendix B.3 for more details on this.

1. The result of a measurement of the observable A can only be an eigenvalue A of the corresponding Hermitian operator $\mathbf{A}$.

2. After the measurement with the outcome A, a system originally in the state Ψ is left in the eigenstate $\mathsf{P}_A\Psi$ of $\mathbf{A}$ with the eigenvalue A, where P_A is the projection on the eigenspace of $\mathbf{A}$ at A, i.e., the space spanned by the eigenvectors Ψ_A (this is called the *projection postulate*).

3. If the system is in an eigenstate of $\mathbf{A}$ with the eigenvalue A, the outcome of a measurement of A is A with certainty.

The physical justification of the second assumption, together with the third, is that an immediate repetition of the measurement had better yield the same result, else the measurement would have no significance; if the system were not left in an eigenstate, there would be no such assurance. It follows from the second postulate that a measurement is, in effect, identical to the *preparation of a state:* by measuring the variable A and finding the value A, we have prepared the system in an eigenstate of the operator $\mathbf{A}$ with the eigenvalue A. (This is really no different from the classical situation. To prepare a system in a precisely specified state implies that all the dynamical variables needed for its unique determination have been measured, and *vice versa.*) As a result, the states before and after the measurement are generally different, and this is not necessarily because of a disturbance caused by the measurement. Although the nature of this change is special to quantum physics, the intrusive effect of measurements is a common feature of probabilistic theories. (In classical statistical mechanics, probabilistic predictions after a measurement are based on the assumption of uniform distribution within the phase-space coarse grain of the measurement, even though without the measurement that distribution might have been quite nonuniform.)

These postulates, however, must be regarded as no more than schemata, because they cannot always be implemented literally. The reason is that most of the operators representing observables, for example, the $\vec{\mathbf{q}}$s and $\vec{\mathbf{p}}$s, have spectra that are, at least in part, continuous, and "eigenstates" or "eigenvectors" corresponding to points in a continuous spectrum, which we shall refer to as *quasi-eigenvectors* and *quasi-eigenvalues*, respectively, are not normalizable and hence are not members of the system's Hilbert space $\mathfrak{H}$. If points in the continuous spectrum of the Hermitian operator $\mathbf{A}$ are denoted by A, the quasi-eigenvectors $|A\rangle$ are such that their inner products $\langle\,|A\rangle$ with all vectors $|\,\rangle \in \mathfrak{H}$ are well defined as square-integrable functions of A. (In other words, they are not necessarily finite for each A; for more about their definition, see Appendix B.3.1.) If we are to avoid idealized measurements of such variables, which would result in quasi-states described outside $\mathfrak{H}$—the inclusion of the description of such idealized measurements and quasi-states, however, is often very convenient, and we shall

not go out of our way to avoid them—the postulates have to be formulated more carefully as follows: (1) The result A of a measurement of an observable A must lie in the spectrum of the operator $\mathbf{A}$ (the spectrum of a Hermitian operator always lies along the real line!), and (2), (3) if A is the result of a measurement of A with an expected error $\leq \Delta A$ and A lies in the continuous spectrum of $\mathbf{A}$, the system is left in a state in which A has the *expected value* A with the *variance* (or *mean-square deviation*, also sometimes called *dispersion*) ΔA. (If you don't know what "expected value" and "variance" are, they will be defined shortly.) However, such a state, as we shall see explicitly in Section 2.2.2, is not determined to within a constant factor; it takes at least the measurement of a second observable to make it so.

Since the necessary and sufficient condition for two Hermitian operators to have a complete set of eigenvectors in common is that they commute (see Appendix B), a state can in general not be prepared so as to lead, with certainty, to specified outcomes of the measurement of two different observables A and B unless the corresponding operators $\mathbf{A}$ and $\mathbf{B}$ commute. In order to identify states of a given physical system uniquely by means of measurements we have to have a *complete set of commuting observables* at our disposal, so that there exists a complete set of states in $\mathfrak{H}$, each of which is labeled by the eigenvalues—and therefore the measurement outcomes—of these simultaneously measurable dynamical variables. Other observables of the system, which fail to commute with this set, are then not sharply defined.

1.1.2 Fields as dynamical variables

While the dynamical variables of particle systems, the $\vec{q}$s and $\vec{p}$s, are observables, this need not be the case for all systems. The most fundamental description of submicroscopic physics is given in terms of a *field*, which we shall denote by $\Psi(\vec{r})$, an *operator-function* of space.[4] This operator is generally not Hermitian and hence does not describe an observable. If we think of Ψ as analogous to $\vec{\mathbf{q}}$ for a particle, then the Hermitian conjugate $\Psi^\dagger$ turns out to be the analogue of $\vec{\mathbf{p}}$, as we shall see in Section 2.3.1. Since the spatial coordinate vector $\vec{r}$ in $\Psi(\vec{r})$, which you should think of as analogous to the index i on $\vec{\mathbf{q}}_i$, $\ ,i = 1, \ldots, n$, for an n-particle system, ranges over a continuum, the number of degrees of freedom of a system described by a field is necessarily infinite.

[4]Certain mathematical difficulties make it necessary to define Ψ, strictly speaking, not at a point but on a small smeared-out region, thus making it a distribution, but we shall ignore this problem here.

1.2 Probabilities

The probabilistic nature of quantum states implies that two *different* states need not be mutually exclusive: if a system is in state Ψ, it may also have many of the characteristics of the linearly independent state Φ. Indeed, the probabilistic assumption stated earlier is made explicit and precise by the postulate that if the system is in the state Ψ, the probability $P(\Phi|\Psi)$ that it also simultaneously has the properties of the state Φ is given by

$$P(\Phi|\Psi) = |(\Phi, \Psi)|^2 / \parallel \Phi \parallel^2 \parallel \Psi \parallel^2 . \tag{1.1}$$

In particular, the probability for a measurement of the observable A to come up with the result A is $|(\Psi_A, \Psi)|^2 / \parallel \Psi_A \parallel^2 \parallel \Psi \parallel^2$, where Ψ_A is the eigenvector of $\mathbf{A}$ corresponding to the eigenvalue A. (It is therefore often convenient to normalize all state vectors so that $\parallel \Psi_A \parallel = 1$ and $\parallel \Psi \parallel = 1$.) Note that the probability (1.1) depends only on the rays determined by Ψ and Φ; it does not depend on their "hair." In case $\mathbf{A}$ has several mutually orthogonal eigenvectors Ψ_A^n with the same eigenvalue A, which is called *degeneracy,* this probability equals $\sum_n |(\Psi_A^n, \Psi)|^2 / \parallel \Psi_A^n \parallel^2 \parallel \Psi \parallel^2$.

If $P(\Phi, \Psi) = 1$, so that the state Ψ is certain to have all the characteristics of the state Φ, then you expect physically that the two states must be identical (at least in isolation), and this expectation is mathematically reflected in Schwarz's inequality (see Appendix B), according to which $P(\Phi, \Psi) = 1$ implies $\Phi = c\Psi$ for some $c \neq 0$. Only if two state vectors are orthogonal, i.e., if $(\Psi, \Phi) = 0$, are the corresponding states mutually exclusive; if the system is in one of these states, the probability of finding that it has the characteristics of the other is zero. The mathematical formalism incorporates this in the fact that two eigenvectors of a Hermitian operator $\mathbf{A}$ that correspond to different eigenvalues are necessarily orthogonal to one another: if the result of a measurement is A_1, the probability of obtaining the result $A_2 \neq A_1$ immediately afterwards vanishes, as it physically should. It is also noteworthy that $P(\Phi, \Psi) = P(\Psi, \Phi)$: the probability that the state Ψ has the properties of the state Φ is the same as the probability that Φ has those of Ψ.

That all measurement results of an observable have to lie in the spectrum of the corresponding operator is mathematically reflected in the *completeness* property of the spectra of all Hermitian operators, which means that any vector Ψ in $\mathfrak{H}$ can be expanded on the basis of the eigenvectors Ψ_A of any Hermitian operator $\mathbf{A}$:

$$\Psi = \sum_A c_A \Psi_A, \tag{1.2}$$

where the sum runs over the spectrum of $\mathbf{A}$, which we assume for the moment to consist of discrete points only. If Ψ and all the eigenvectors Ψ_A are normalized to unity, it follows from the mutual orthogonality of the

latter [just take the inner product of (1.2) with Ψ_A] that

$$c_A = (\Psi_A, \Psi), \tag{1.3}$$

which implies $\sum_A |c_A|^2 = 1$ and $|c_A|^2$ is the probability $P(A|\Psi)$ of obtaining the result A upon measurement of A in the state Ψ, so that appropriately $\sum_A P(A|\Psi) = 1$. Thus every state of a system can be decomposed into eigenstates of any given observable A of the system, and the coefficients c_A are the "probability amplitudes" (the squares of whose magnitudes are the probabilities) of obtaining the results A upon measurement of A: $P(A|\Psi) = |c_A|^2 = |(\Psi_A, \Psi)|^2$. If the spectrum of $\mathbf{A}$ is continuous, the sum in (1.2) has to be replaced by an integral, analogous to replacing a Fourier series by a Fourier integral,

$$\Psi = \int dA\, c(A)\Psi_A, \tag{1.4}$$

where Ψ_A is a *quasi-eigenvector* of $\mathbf{A}$ that satisfies the equation $\mathbf{A}\Psi_A = A\Psi_A$ with the quasi-eigenvalue A. Here, however, $\Psi_A \notin \mathfrak{H}$; since it does not have a finite norm, it is *not normalizable*. [The quasi-eigenvectors are analogous to $\exp(ikx)$, with (1.4) analogous to a Fourier integral.] These quasi-eigenvectors are still mutually orthogonal, and they can be so chosen that

$$(\Psi_A, \Psi_B) = \delta(A - B), \tag{1.5}$$

where $\delta(A-B)$ is Dirac's delta function (see Appendix A for the definition). We shall call them "δ-normalized" and refer to a set of vectors that satisfy (1.4) and (1.5) as an *orthonormal quasi-basis.* Equations (1.4) and (1.5) imply that

$$c(A) = \int dA\, (\Psi_A, \Psi),$$

and the physical interpretation of $|c(A)|^2$ is that of a *probability density*: $|c(A)|^2 dA$ is the probability of finding the result of a measurement of the observable A to lie in the interval dA around A. If an operator has a spectrum that is in part discrete and in part continuous, or if we wish to write a formula applicable to either kind of spectrum, we shall indicate the summation over the spectrum by the symbol $\int\!\!\!\!\!\sum$.

1.2.1 Correlations

Suppose a system with the Hilbert space $\mathfrak{H}$ consists of two subsystems with the Hilbert spaces $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{H}^{\mathrm{II}}$, so that $\mathfrak{H} = \mathfrak{H}^{\mathrm{I}} \bigotimes \mathfrak{H}^{\mathrm{II}}$, and a pure state of the combined system is of the form $\Psi = \sum_{nm} a_{nm} \Psi^{\mathrm{I}}_n \otimes \Psi^{\mathrm{II}}_m$. (We will assume that Ψ is normalized and the sets of states $\{\Psi^{\mathrm{I}}_n\}$ and $\{\Psi^{\mathrm{II}}_m\}$ are orthonormal bases in $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{H}^{\mathrm{II}}$, respectively.) In such a state the probability of finding the two subsystems in states Ψ^{I}_n and Ψ^{II}_m, respectively,

is $P(n,m|\Psi) = |a_{nm}|^2$, and the probability of obtaining the results A and B upon measurement of the variable A of system I and of B of system II, respectively, is given by

$$P(A,B|\Psi) = |(\Psi^{\rm I}_A \otimes \Psi^{\rm II}_B, \Psi)|^2 = \left|\sum_{nm} a_{nm}(\Psi^{\rm I}_A, \Psi^{\rm I}_n)(\Psi^{\rm II}_B, \Psi^{\rm II}_m)\right|^2 .$$

Therefore the probability of finding system I in the state $\Psi^{\rm I}_n$, irrespective of what state the other system is in, is given by $\sum_m |a_{nm}|^2$. If the two systems are totally independent and uncorrelated, we should have $P(n,m|\Psi) = P(n^{\rm I}|\Psi)P(m^{\rm II}|\Psi)$, which will be the case if and only if a_{nm} is a product $a_{nm} = b_n c_m$, in which instance Ψ is of the form

$$\Psi = \Psi^{\rm I} \otimes \Psi^{\rm II}. \tag{1.6}$$

In that case the probability of finding system I in the state $\Phi^{\rm I}$ is independent of the state of system II.

The general state of the form $\Psi = \sum_{nm} a_{nm}\Psi^{\rm I}_n \otimes \Psi^{\rm II}_m$ of the combined system may also be written in the alternative form

$$\Psi = \sum_m a_m \Phi^{\rm I}_m \otimes \Psi^{\rm II}_m, \tag{1.7}$$

where $a_m\Phi^{\rm I}_m \stackrel{\rm def}{=} \sum_n a_{nm}\Psi^{\rm I}_n$; the states $\Phi^{\rm I}_m$ can be assumed to be normalized but they are not necessarily mutually orthogonal. Eq.(1.7) clearly shows the two systems to be so *entangled*, even though they may no longer be in interaction with one another, that when a test finds system II in state $\Psi^{\rm II}_m$, then system I must be in state $\Phi^{\rm I}_m$, no matter how far separated the two systems may be. The states $\Psi^{\rm II}_m$ may, for example, be eigenstates of an operator $\mathbf{A}$ with the eigenvalue A_m, i.e., the result of a measurement of A on system II with the outcome A_m, and the states $\Phi^{\rm I}_m$ may be eigenstates of $\mathbf{B}$ with the eigenvalue B_m. When the entire system is in the state Ψ, neither system I nor system II can be said to be in any eigenstate of $\mathbf{A}$ or $\mathbf{B}$. However, a measurement of A on system II allows us to infer what the result of a measurement of $\mathbf{B}$ on system I would have been, had it been performed, and thus casting it into an eigenstate of $\mathbf{B}$ without ever coming near the latter, and therefore never disturbing it.

An example of such a situation was used in the famous EPR paper[EPR]. In the somewhat simpler version given by David Bohm,[5] the *Gedanken* experiment goes as follows. A particle of spin 0 at rest decays into two spin-1/2 fragments, which fly off in opposite directions. If at a later time, when they are far apart, the vertical spin projection of fragment A is measured, the probability for it to be found *up* is $\frac{1}{2}$, and similarly for fragment B. But

[5][Bohm, D.] p. 614.

since angular-momentum conservation dictates that the total angular momentum of the two must be 0, the spin states of the two fragments must be correlated so as to be opposite to one another. Therefore, if the vertical spin projection of A is found to be *up*, we know for certain, without coming near B for a measurement, that its spin must be *down.* On the other hand, if the horizontal spin component of A had been measured and found to be *left*, we would know with equal certainty, without approaching it, that the horizontal spin component of B is *right.* (Since, by EPR's definition, these two statements about the vertical and horizontal spin projections of B imply that both spin projections are "elements of reality" but quantum mechanics does not allow them to be simultaneously specified because they fail to commute, the authors conclude that this theory cannot be a "complete description of reality," a conclusion that rests on their specific initial, philosophically based definition of reality.)

All philosophical preconceptions aside, the existence of such correlations between distant entities that allow reliable inferences to replace actual physical measurements demonstrates that quantum mechanics *does not* necessarily imply some irreducible disturbance in the act of measurement, as is sometimes claimed. Notice also that if two systems are correlated as in (1.7), a measurement on one of them instantaneously changes the state of the other, although there may be no interaction (at the present time) between the constituents of the two systems and no matter how great the distance between them. This feature, too, of quantum mechanics is a simple consequence of its probabilistic nature and not of its Hilbert-space formulation.

1.2.2 Interference

The obvious physical interpretation of the superposition of two vectors, as in $\Psi = (\Psi_1 + \Psi_2)/\sqrt{2}$, is that if the system is in the state Ψ, it can be found to have the properties both of the state Ψ_1 and of the state Ψ_2. Assuming that Ψ_1 and Ψ_2 are both normalized and mutually orthogonal, the probability of finding that the system has the characteristics of either state is $\frac{1}{2}$, which would be equally true, however, for the linearly independent state $\Psi' = (e^{i\varphi}\Psi_1 + \Psi_2)/\sqrt{2}$ with any real $\varphi \neq 0$.

Suppose now that a system is in the state described by the normalized vector Φ and we are asking for the probability that, upon measurement, it will be found in agreement with the state $\Psi = (\Psi_1 + \Psi_2)/\sqrt{2}$. This probability is

$$\begin{aligned} P(\Psi|\Phi) &= |(\Psi,\Phi)|^2 = \frac{1}{2}|(\Psi_1,\Phi) + (\Psi_2,\Phi)|^2 \\ &= \frac{1}{2}|(\Psi_1,\Phi)|^2 + \frac{1}{2}|(\Psi_2,\Phi)|^2 + \Re[(\Psi_1,\Phi)^*(\Psi_2,\Phi)], \quad (1.8) \end{aligned}$$

where $\Re$ denotes the *real part* of a complex number and the star indicates the complex conjugate. So the probability of finding the state of the system to agree with Ψ is *not* necessarily equal to the sum of the probabilities $P(\Psi_1|\Phi)$ of finding it in agreement with Ψ_1 and $P(\Psi_2|\Phi)$ of finding it in agreement with Ψ_2, even if the two states Ψ_1 and Ψ_2 are mutually exclusive; instead there is an *interference term* $\Re[(\Psi_1, \Phi)^*(\Psi_2, \Phi)]$, an effect that is characteristic of quantum physics and which indicates the *correlation* between the two states Ψ_1 and Ψ_2 when they form a superposition. Since the interference term is the real part of the product of the two complex numbers $(\Psi_1, \Phi)^*$ and (Ψ_2, Φ), the correlation would be different for $\Psi' = (e^{i\varphi}\Psi_1 + \Psi_2)/\sqrt{2}$. For example, if $P(\Psi_1, \Phi) = P(\Psi_2, \Phi) = p$, then $P(\Psi, \Phi)$ may vary between 0 and $2p$, depending upon the relative phases of (Ψ_1, Φ) and (Ψ_2, Φ).

Whereas this result appears to differ from the classical addition of probabilities for independent systems, it is important to remember that for two states of a given system to form another state by superposition means that they are *entangled*; they are *not independent*. As already noted earlier, this kind of quantum entanglement has its analogue in wave systems (the spectral decomposition, giving the magnitude of the weight of each spectral component, of an electromagnetic wave is not sufficient to specify the wave uniquely), but it has no classical analogue for systems consisting of particles. It should therefore be emphasized that the interference term in (1.8), rather than showing the inapplicability of classical probability theory to quantum mechanics, represents the effect of a kind of correlation between states of quantum systems that most classical systems lack (and those classical systems that have such correlations, such as wave systems, are not usually treated probabilistically). There is no need to invent a special theory of probability applicable only to quantum physics.

The kind of correlation between states of physical systems—even between systems that may be far apart from one another—expressed by the superposition principle and the resulting interference effects implies a certain *nonlocality* inherent in quantum physics. Such nonlocal action-at-a-distance, if postulated as existing between classical particles, is usually regarded as physically unacceptable; for waves, however, we do not abhor it and indeed are quite accustomed to its presence. There is, therefore, no reason to find puzzling such nonlocal effects as are implied by the quantum superposition principle, and the concomitant, experimentally well confirmed interferences. They simply indicate that quanta, though we call them "particles," are not little billiard balls.[6]

[6]In the alternative interpretation of quantum physics put forward by David Bohm, which tries to avoid the quantum acausality by postulating an underlying quasi-classical substratum described by hidden variables, the unavoidable nonlocality appears as a result of an action-at-a-distance between unobservable particles that behave classically, which is physically much more objectionable. Far from being intuitively more appealing,

1.2.3 Expectation values and variance

The *average* value (also called the *mean*, the *expectation value* or the *expected value*) of the observable A in the state Ψ is, by definition, given by

$$\langle \mathbf{A} \rangle \stackrel{\text{def}}{=} \sum\!\!\!\!\!\!\int_A A\, P(A|\Psi),$$

the right-hand side of which is easily evaluated by means of (1.2) and (1.3), with the result

$$\langle \mathbf{A} \rangle = \frac{(\Psi, \mathbf{A}\Psi)}{\| \Psi \|^2} = \frac{\langle\, |\mathbf{A}|\, \rangle}{\langle\, |\, \rangle}. \tag{1.9}$$

The mean value, of course, depends on the state Ψ, but this dependence is customarily not indicated in the notation $\langle \mathbf{A} \rangle$. The *variance*, *dispersion,* or *root-mean-square deviation* of A in the state Ψ is defined by

$$\Delta A \stackrel{\text{def}}{=} \sqrt{\langle (\mathbf{A} - \langle \mathbf{A} \rangle)^2 \rangle} = \sqrt{\langle \mathbf{A}^2 \rangle - \langle \mathbf{A} \rangle^2}. \tag{1.10}$$

If Ψ is an eigenstate of $\mathbf{A}$, its variance is zero: it is *sharp.* To see this, assume that $\mathbf{A}\Psi = A\Psi$ and $\| \Psi \| = 1$; then

$$\langle \mathbf{A}^2 \rangle = (\Psi, \mathbf{A}^2\Psi) = A^2 = \langle \mathbf{A} \rangle^2,$$

and hence $\Delta A = 0$. Conversely, it follows from the Hermiticity of $\mathbf{A}$ and Schwarz's inequality[7] that if $\Delta A = 0$, Ψ must be an eigenstate of $\mathbf{A}$: assuming that $\mathbf{A}\Psi$ and Ψ are linearly independent, so that Ψ is not an eigenstate of $\mathbf{A}$, and that $\| \Psi \| = 1$, we have

$$\langle \mathbf{A} \rangle^2 = |(\Psi, \mathbf{A}\Psi)|^2 < \| \Psi \|^2 \| \mathbf{A}\Psi \|^2 = (\mathbf{A}\Psi, \mathbf{A}\Psi) = (\Psi, \mathbf{A}^2\Psi) = \langle \mathbf{A}^2 \rangle,$$

and therefore $\Delta A > 0$. Thus the only states that are dispersion-free for a given observable are its eigenstates; in the eigenstates of a given observable $\mathbf{A}$, the probability distributions of other observables that do not commute with $\mathbf{A}$ cannot be sharp (see the remark at the end of Section 1.1.1).

A general result with an important special case is that the product of the dispersions of two Hermitian operators is bounded below by one-half of the average of their *commutator*, $[\mathbf{A}, \mathbf{B}] \stackrel{\text{def}}{=} \mathbf{AB} - \mathbf{BA}$. In order to prove this, let us subtract from these operators their expectation values in the state to be considered (which we take normalized to unity) and define $\mathbf{A}' \stackrel{\text{def}}{=} \mathbf{A} - \langle \mathbf{A} \rangle$ [this should really be written as $\mathbf{A}' \stackrel{\text{def}}{=} \mathbf{A} - \langle \mathbf{A} \rangle \mathbb{1}$, where $\mathbb{1}$ is the unit operator (which has the property that for all $\Psi \in \mathfrak{H}$, $\mathbb{1}\Psi = \Psi$), but we shall forego this bit of pedantry] and $\mathbf{B}' \stackrel{\text{def}}{=} \mathbf{B} - \langle \mathbf{B} \rangle$, so that $(\Delta A)^2 =$

---

in fact the Bohm theory is, on these grounds alone, apart from objections based on its hidden variables, more counterintuitive than the orthodox quantum theory.

[7] See Appendix B.

$\langle(\mathbf{A}')^2\rangle = \| \mathbf{A}'\Psi \|^2$ and $(\Delta B)^2 = \| \mathbf{B}'\Psi \|^2$, as well as $[\mathbf{A}, \mathbf{B}] = [\mathbf{A}', \mathbf{B}']$. We then have by Schwarz's inequality

$$\begin{aligned} |(\Psi, [\mathbf{A}, \mathbf{B}]\Psi)| &= |(\Psi, \mathbf{A}'\mathbf{B}'\Psi) - (\Psi, \mathbf{B}'\mathbf{A}'\Psi)| \\ &\leq |(\Psi, \mathbf{A}'\mathbf{B}'\Psi)| + |(\Psi, \mathbf{B}'\mathbf{A}'\Psi)| \\ &\leq 2 \| \mathbf{A}'\Psi \| \| \mathbf{B}'\Psi \| = 2\Delta A \Delta B; \end{aligned}$$

so

$$\Delta A \, \Delta B \geq \frac{1}{2} |\langle [\mathbf{A}, \mathbf{B}] \rangle|. \tag{1.11}$$

In this inequality, the right-hand side as well as the variances on the left are generally state-dependent, except when the commutator between $\mathbf{A}$ and $\mathbf{B}$ is a constant (that is, a multiple of the unit operator), as it is in the special case of a particle's position $\vec{\mathbf{q}}$ and momentum $\vec{\mathbf{p}}$. In that instance we have for each component, as we shall justify in Section 2.1.3,

$$[\mathbf{q}_k, \mathbf{p}_l] = i\hbar\delta_{kl} \qquad k, l = 1, 2, 3, \tag{1.12}$$

(where δ_{kl} is the Kronecker symbol, which equals 1 when $k = l$ and 0 otherwise) and therefore, by (1.11),

$$\Delta q_k \Delta p_k \geq \frac{1}{2}\hbar, \qquad k = 1, 2, 3, \tag{1.13}$$

the right-hand side of which is universal, that is, state-independent. This, of course, is nothing but Heisenberg's uncertainty (or indeterminacy) relation.

If the two observables $\mathbf{A}$ and $\mathbf{B}$ commute, (1.11) shows that there is no lower bound on their uncertainty product. There are, then, states in which both are specified as sharply as the precision of the measurement apparatus permits, in agreement with the fact that such commuting Hermitian operators have a complete set of eigenvectors in common. Commuting observables for a given system are physically compatible with one another and can be specified simultaneously. This may be necessary, for example, if there is degeneracy in their spectra, so that giving the eigenvalue of only one of them does not determine a state to within a constant factor.

1.3 Mixed States

The states of physical systems are, of course, not always specified as precisely as nature allows. In order to represent such less-than-precisely given states of particle systems, classical statistical mechanics uses the technique of *coarse graining*: the phase space is subdivided into small grains, whose size depends on the accuracy of an experimenter's measuring instruments or on the sharpness of an observer's focus. When a system is specified to be in one of these grains, its dynamical variables are fixed only approximately,

and it is the task of statistical mechanics to formulate the laws governing its behavior. Analogously in quantum physics, a system may be specified only to be in a "coarse grain of the Hilbert space $\mathfrak{H}$," that is, to be in one of the mutually exclusive (normalized) states $\Psi_n = |n\rangle$ with the probability p_n. In that case, instead of using (1.1) to calculate the probability for the state of the system to agree with some given (normalized) state $\Phi = |\,\rangle$, one would calculate this probability by

$$P(\Phi) = \sum_n p_n P(\Phi|\Psi_n) = \sum_n p_n |(\Phi, \Psi_n)|^2 = \langle\,| \sum_n p_n |n\rangle\langle n|\,|\,\rangle.$$

Now $\mathsf{P}_n \stackrel{\text{def}}{=} |n\rangle\langle n|$ may be regarded as an operator in the sense that $\mathsf{P}_n\Phi \stackrel{\text{def}}{=} (\Psi_n, \Phi)\Psi_n$. Because it projects any arbitrary vector on Ψ_n, P_n is called a *projection operator,* whose two defining characteristics are that it is Hermitian and *idempotent*, which means that $\mathsf{P}_n^2 = \mathsf{P}_n$.

It is therefore convenient to define the *density operator* of a system by

$$\boldsymbol{\rho} \stackrel{\text{def}}{=} \sum_n p_n |n\rangle\langle n| = \sum_n p_n \mathsf{P}_n, \tag{1.14}$$

where the numbers p_n are non-negative, such that $\sum_n p_n = 1$, and $\mathsf{P}_n\mathsf{P}_m = \mathsf{P}_n\delta_{nm}$. If the state of a given system is described by $\boldsymbol{\rho}$, the probability of finding it to have the properties of the state Φ is expressed by

$$P(\Phi, \boldsymbol{\rho}) = (\Phi, \boldsymbol{\rho}\Phi) = \langle\,|\boldsymbol{\rho}|\,|\,\rangle = \langle\boldsymbol{\rho}\rangle. \tag{1.15}$$

In the special instance in which $p_n = 0$ for all n but one, the density operator degenerates into a single projection onto the ray representing an isolated, "shaved" state of the system.

For example, if the system is definitely in one of the two *uncorrelated*, truly independent states $|1\rangle = \Psi_1$ or $|2\rangle = \Psi_2$ with probability $\frac{1}{2}$ each, its density operator is $\boldsymbol{\rho} = \frac{1}{2}(\mathsf{P}_1 + \mathsf{P}_2)$ and the probability of finding it to agree with the state Φ is $P(\Phi|\boldsymbol{\rho}) = \frac{1}{2}[|(\Phi, \Psi_1)|^2 + |(\Phi, \Psi_2)|^2]$, without the interference term present if it were in a *superposition* of the two states Ψ_1 and Ψ_2. Equation (1.15) also implies that if a system is in the state described by $\boldsymbol{\rho}$, the probability of obtaining the result A upon measurement of the variable A is given by

$$P(A\,|\,\boldsymbol{\rho}) = \frac{(\Psi_A, \boldsymbol{\rho}\Psi_A)}{(\Psi_A, \Psi_A)} = \frac{\langle A|\boldsymbol{\rho}|A\rangle}{\langle A|A\rangle} = \mathrm{tr}(\mathsf{P}_A\,\boldsymbol{\rho}), \tag{1.16}$$

where P_A is the projection on the eigenspace of $\mathbf{A}$ at the eigenvalue A and tr denotes the trace (see Appendix B).

Suppose we perform a measurement of the variable A on a system in a pure state Φ (assumed normalized), so that the probability of obtaining the result A is given by $P(A|\Phi) = |(\Psi_A, \Phi)|^2$ if Ψ_A is a normalized eigenstate

of **A** with the eigenvalue A. (Assume for simplicity that the spectrum of **A** is discrete and without degeneracies.) After the measurement with the outcome A, the state of the system is pure and given by $\mathsf{P}_A\Phi = \Psi_A(\Psi_A, \Phi)$. On the other hand, the state of the system after the measurement *irrespective of the outcome*[8] is not pure but described by the density operator

$$\boldsymbol{\rho}_{\mathbf{A}} = \sum_A |(\Psi_A, \Phi)|^2 \mathsf{P}_A = \sum_A |A\rangle\,\langle\,|A\rangle\,\langle A|\,\rangle\,\langle A|, \tag{1.17}$$

whose eigenvalues are $|(\Psi_A, \Phi)|^2$ and whose eigenfunctions are $\Psi_A = |A\rangle$.[9] After the measurement, the system is therefore no longer in a pure state but in a mixed state.[10] More generally, if the state of the system before the measurement is described by the density operator $\boldsymbol{\rho}$, then its state after the measurement of A, with the "outcome ignored," is described by

$$\boldsymbol{\rho}_{\mathbf{A}} = \sum_A \mathsf{P}_A \boldsymbol{\rho} \mathsf{P}_A. \tag{1.18}$$

If a system, consisting of two subsystems with Hilbert spaces $\mathfrak{H}^{\mathrm{I}}$ and $\mathfrak{H}^{\mathrm{II}}$, is in the pure state $\Psi = \sum_{nm} a_{nm} \Psi^{\mathrm{I}}_n \otimes \Psi^{\mathrm{II}}_m$, then the state of system I, with system II ignored, is described by the density operator[11]

$$\boldsymbol{\rho}^{(\mathrm{I})} = \sum_n b_{nm} \mathsf{P}^{(\mathrm{I})}_{nm}, \qquad b_{nm} = \sum_k a_{nk} a^*_{mk}, \tag{1.19}$$

where

$$\mathsf{P}^{(\mathrm{I})}_{nm} = |n\rangle^{(\mathrm{I})\ (\mathrm{I})}\langle m|,$$

meaning that $\mathsf{P}^{(\mathrm{I})}_{nm}\Phi^{\mathrm{I}} = \Psi^{\mathrm{I}}_n(\Psi^{\mathrm{I}}_m, \Phi^{\mathrm{I}})$. This may be expressed in the more general form applicable also to instances in which the system as a whole is in a mixed state,

$$\boldsymbol{\rho}^{(\mathrm{I})} = \mathrm{tr}_{\mathrm{II}} \boldsymbol{\rho}, \tag{1.20}$$

where $\mathrm{tr}_{\mathrm{II}}$ means that the trace is taken only over the states in $\mathfrak{H}^{\mathrm{II}}$. Even when the complete system is in a pure state, a given subsystem is generally in a mixed state, represented by a density operator.

---

[8] This is often expressed by saying "the outcome is ignored." Such a formulation, however, gives the misleading impression that what the experimenter knows or has in mind influences the state of the physical system. No such inference should be drawn.

[9] **Suppose a standard two-slit experiment is performed with electrons made visible on a screen and photographed. Describe the state in which "the outcome is ignored"; how would you record it?**

[10] In terms of ensembles, the measurement of A with the result A yields a new ensemble of states in each of which A has the same value A (this ensemble is described by an eigenvector of **A** with the eigenvalue A). The measurement of A "with the result ignored" yields an ensembles in which each of the possible eigenvalues A is represented in proportion to the number of times A was obtained in the measurement. This set is described by the density operator (1.17).

[11] **Prove this as an exercise**.

As an illustration, consider the example of the radioactive decay of an infinitely massive (therefore non-recoiling) nucleus in a pure state, with the emission of an electron and an antineutrino, so that the final state of the two-particle system consisting of the electron and the antineutrino is schematically given by

$$\Psi = \sum\!\!\!\!\!\!\int_{\alpha_1\alpha_2} f(\alpha_1, \alpha_2)\Psi_{\rm el}(\alpha_1) \otimes \Psi_{\rm neu}(\alpha_2),$$

where $\Psi_{\rm el}(\alpha_1)$ is the state of the electron, $\Psi_{\rm neu}(\alpha_2)$ is the state of the antineutrino, and α_1 and α_2 denote the eigenvalues or quasi-eigenvalues of the needed dynamical variables: the energy, the momentum direction, the angular momentum, etc. If the antineutrino is ignored, the probability of obtaining the result β upon a measurement on the electron alone is given by

$$\begin{aligned}\sum\!\!\!\!\!\!\int_{\gamma} |(\Psi, \Psi_{\rm el}(\beta) \otimes \Psi_{\rm neu}(\gamma))|^2 &= \sum\!\!\!\!\!\!\int_{\gamma} \Big| \sum\!\!\!\!\!\!\int_{\alpha_1\alpha_2} f(\alpha_1, \alpha_2)(\Psi_{\rm el}(\alpha_1), \Psi_{\rm el}(\beta)) \\ &\qquad \times (\Psi_{\rm neu}(\alpha_2), \Psi_{\rm neu}(\gamma))\Big|^2 \\ &= (\Psi_{\rm el}(\beta), \boldsymbol{\rho}_{\rm el}\Psi_{\rm el}(\beta)),\end{aligned}$$

where

$$\boldsymbol{\rho}_{\rm el} = \sum\!\!\!\!\!\!\int_{\alpha_1'\alpha_1''} \left[\sum\!\!\!\!\!\!\int_{\gamma} f(\alpha_1', \gamma) f^*(\alpha_1'', \gamma)\right] |\alpha_1'\rangle_{\rm el}\langle\alpha_1''|_{\rm el}.$$

This shows that, as remarked earlier, the state of the electron will be a mixture, rather than a superposition, of different momenta and angular momentum projections.

For another example, take the vectors $|E_n\rangle = \Psi_{E_n}$ to be normalized eigenstates of the Hamiltonian $\mathbf{H}$ (the energy operator), and let $\mathsf{P}_n \stackrel{\rm def}{=} |E_n\rangle\langle E_n|$ be the projection on $|E_n\rangle$. If the probability of finding the system in the state $|E_n\rangle$ is $p(E_n)$, then the density operator of the mixed state in which there are no correlations between these eigenstates of $\mathbf{H}$ is given by

$$\boldsymbol{\rho} = \sum_n p(E_n)\mathsf{P}_n = \sum_n p(\mathbf{H})\mathsf{P}_n = p(\mathbf{H})$$

because $\mathbf{H}\mathsf{P}_n = E_n\mathsf{P}_n$ and $\sum_n \mathsf{P}_n = \mathbb{1}$ [see (1.23) below]. For a system in thermodynamic equilibrium at the temperature T, with constituents distributed according to the Boltzmann distribution $p(E) = e^{-E/kT}$ (where k is the Boltzmann constant), the density operator can therefore be written in the simple form

$$\boldsymbol{\rho} = e^{-\mathbf{H}/kT}. \tag{1.21}$$

The density operator is the most general representation of an isolated quantum state. Because we had assumed that the states Ψ_n in (1.14) are

mutually exclusive, which means these vectors are mutually orthogonal, $(\Psi_n, \Psi_m) = \delta_{nm}$, (1.14) shows that they are the eigenvectors and the p_n are the eigenvalues of $\boldsymbol{\rho}$. Since $p_n \geq 0$ for all n, the density operator is *positive semi-definite,* and since the probabilities p_n have to add up to unity, $\sum_n p_n = 1$, its *trace*, i.e., the sum of its eigenvalues, equals unity:

$$\mathrm{tr}\boldsymbol{\rho} = 1. \tag{1.22}$$

Furthermore we have $\boldsymbol{\rho}^2 = \sum_n p_n^2 |n\rangle\langle n|$, so that $\langle \boldsymbol{\rho} - \boldsymbol{\rho}^2 \rangle = \sum_n (p_n - p_n^2)|\langle n|\,\rangle|^2 \geq 0$, because for all n, $0 \leq p_n \leq 1$, and therefore $p_n - p_n^2 > 0$ unless $p_n = 0$ or 1. Only if a single $p_n = 1$ and all others vanish can we have $\langle \boldsymbol{\rho} - \boldsymbol{\rho}^2 \rangle = 0$. This is the special case in which the system can be represented by a ray spanned by the vector $\Psi \stackrel{\text{def}}{=} |a\rangle$, and the density operator degenerates to a single projection on this ray: $\boldsymbol{\rho} = |a\rangle\langle a|$. *The density operator of such a state has the special property of being idempotent:* $\boldsymbol{\rho}^2 = \boldsymbol{\rho}$ and hence being a projection (on the ray representing the non-mixed, isolated state).[12] The trace of $\boldsymbol{\rho}^2$ is a number between 0 and 1: $0 < \mathrm{tr}\boldsymbol{\rho}^2 = \sum_n p_n^2 \leq \sum_n p_n = 1$, where the equality holds if and only if one of the $p_n = 1$ while all the others vanish, which means that $\boldsymbol{\rho}$ denotes a non-mixed state. Therefore, for any given system, $\mathrm{tr}\boldsymbol{\rho}^2$ may be taken to be a relative measure of the "degree of mixing" or "coherence" of its state. How small the minimum value of $\mathrm{tr}\boldsymbol{\rho}^2$ is when the state is completely incoherent, however, depends on the number of states available to the system: if there are n states, the minimum of $\mathrm{tr}\boldsymbol{\rho}^2$ is $1/n$.[13]

1.4 Representations

The spectrum of every Hermitian operator on a Hilbert space $\mathfrak{H}$ spans the space: this is the mathematical expression of the physical requirement that the measurement of an observable on any system must yield *some* result. Suppose, to begin with, that the Hermitian operator $\mathbf{A}$ corresponding to the observable A has a discrete point spectrum only, and that to each of its eigenvalues A there corresponds a one-dimensional eigenspace spanned by Ψ_A, i.e., there is no degeneracy; then every $\Psi \in \mathfrak{H}$ can be expanded on the basis of these mutually orthogonal vectors in the form (1.2). The sequence of coefficients given by (1.3) as $c_A = (\Psi_A, \Psi) = \langle A|\,\rangle$ therefore uniquely represents the vector Ψ; the set of numbers $\{\langle A|\,\rangle\}$, where A runs over the spectrum of $\mathbf{A}$, is the $\mathbf{A}$-*representation* of $|\,\rangle$. If there is degeneracy, i.e., if some, or all, of the eigenspaces are multidimensional, then more than one

[12] I am using the term "non-mixed" rather than "pure" because the state is not as well defined as nature permits; it is isolated and has "no hair." States represented by a density operator never have "hair."

[13] **Prove this as an exercise**.

physical variable is needed to specify each possible state of the system and a complete set of commuting observables has to be used to label the basis uniquely.

A very convenient way of expressing the completeness (1.2) together with (1.3) is in the formula [assuming $\langle A|A\rangle = 1$ for all A; see (B.16) in Appendix B]

$$\sum_A |A\rangle\langle A| = \mathbb{1}, \tag{1.23}$$

where $\mathbb{1}$ stands for the unit operator on $\mathfrak{H}$. The meaning of (1.23) becomes clear when both sides act on an arbitrary vector $|\,\rangle \in \mathfrak{H}$, in which case it reads

$$|\,\rangle = \sum_A |A\rangle\langle A|\,\rangle,$$

combining (1.2) with (1.3). The use of (1.23) becomes particularly convenient when we want to express an arbitrary inner product of two vectors $\Psi_B = |B\rangle$ and $\Psi_C = |C\rangle$ in the **A**-representation:

$$(\Psi_C, \Psi_B) = \langle C|B\rangle = \langle C|\mathbb{1}|B\rangle = \sum_A \langle C|A\rangle\langle A|B\rangle = \sum_A \langle A|C\rangle^*\langle A|B\rangle. \tag{1.24}$$

The two quantities appearing on the right-hand side are simply the **A**-representations of the vectors $|B\rangle$ and $|C\rangle$. As a special case of (1.24) we have, for $B = C$,

$$(\Psi_B, \Psi_B) = \| \Psi_B \|^2 = \langle B|B\rangle = \sum_A |\langle A|B\rangle|^2,$$

or more generally,

$$\| \Psi \| = \sqrt{\sum_A |\langle A|\,\rangle|^2},$$

expressing the norm of an arbitrary state vector in terms of its **A**-representation.

Similarly for an arbitrary operator **X**:

$$(\Psi_C, \mathbf{X}\Psi_B) = \langle C|\mathbf{X}|B\rangle = \langle C|\mathbb{1}\mathbf{X}\mathbb{1}|B\rangle = \sum_{AA'} \langle C|A\rangle\langle A|\mathbf{X}|A'\rangle\langle A'|B\rangle,$$

which expresses the left-hand side completely in terms of the set of numbers $\langle A|\mathbf{X}|A'\rangle$ and the **A**-representatives of $|B\rangle$ and $|C\rangle$. In other words, in the **A**-representation the abstract operator is replaced by a *matrix,* whose elements are given by $X_{AA'} = \langle A|\mathbf{X}|A'\rangle$. Since operator multiplication now simply becomes matrix multiplication,

$$\langle A|\mathbf{XY}|A'\rangle = \langle A|\mathbf{X}\mathbb{1}\mathbf{Y}|A'\rangle = \sum_{A''} \langle A|\mathbf{X}|A''\rangle\langle A''|\mathbf{Y}|A'\rangle,$$

all abstract calculations in $\mathfrak{H}$ can be replaced by ordinary numerical matrix calculations.

These things become a little stickier if the Hermitian operator $\mathbf{A}$ has a continuous spectrum, in which case the quasi-eigenvectors are not in $\mathfrak{H}$. (See Appendix B.3.1 for details.) The Dirac notation, however, is designed to allow us to manipulate all the formulas given above just as for discrete spectra, simply replacing sums by integrals: Eq. (1.23) is relaced by (B.23) in the Appendix, and with the "normalization" (B.24) we obtain the generalized Fourier integral

$$|\,\rangle = \int dA\, |A\rangle\langle A|\,\rangle,$$

in which $|A\rangle$ plays a role analogous to the exponential $\exp(ikx)$ and the integral extends over the continuous spectrum of $\mathbf{A}$.

The "matrix" representing the operator $\mathbf{X}$ becomes the kernel $X(A', A'')$ of an *integral operator*, since by (B.23)

$$\begin{aligned}\langle A|\mathbf{X}|A'\rangle &= \langle A|\mathbb{1}\mathbf{X}\mathbb{1}|A'\rangle = \int dA''\, dA'''\, \langle A|A''\rangle\langle A''|\mathbf{X}|A'''\rangle\langle A'''|A'\rangle \\ &\stackrel{\text{def}}{=} \int dA''\, dA'''\, \langle A|A''\rangle X(A', A'')\langle A'''|A'\rangle.\end{aligned}$$

As we shall see, however, in some instances they become equivalent to differential operators.

Suppose that $|A_1\rangle$ is an eigenvector of the operator $\mathbf{A}$ with the discrete eigenvalue A_1. Then its $\mathbf{A}$-representation $\{\langle A|A_1\rangle\}$ is a sequence with $\langle A_1|A_1\rangle = 1$ and zeros everywhere else. (This is on the assumption that there is no degeneracy; if there is, there may be several nonzero entries in the sequence.) In the $\mathbf{A}$-representation, the operator $\mathbf{A}$ acting on an arbitrary vector $|\,\rangle$ becomes simply the number A multiplying the entry $\langle A|\,\rangle$ in the sequence,

$$\langle A|\mathbf{A}|\,\rangle = A\langle A|\,\rangle,$$

which may be viewed as expressing the fact that the matrix representing the operator $\mathbf{A}$ in the $\mathbf{A}$-representation is *diagonal,* with the eigenvalues of $\mathbf{A}$ on the diagonal, because $\langle A'|\mathbf{A}|A''\rangle = A''\langle A'|A''\rangle = A''\delta_{A'A''}$. One therefore refers to the problem of finding all the eigenvectors of an observable also as the problem of diagonalizing the matrix representing it on an arbitrary basis. (We shall discuss this in more detail a little later; see also Appendix B.)

For continuous spectra, these things are formally the same, even though there are no actual eigenvectors. The $\mathbf{A}$-representation of a quasi-eigenvector $|A_1\rangle$ is the Dirac delta function $\delta(A - A_1)$, a "function" that differs effectively from zero only at the point $A = A_1$. As for the integral kernel representing $\mathbf{A}$, it becomes a multiple of the delta function: $\langle A'|\mathbf{A}|A''\rangle = A'\delta(A' - A'')$. Here are some specific examples of particular importance.

1.4.1 The configuration representation

Let us begin with the simple case of a single point particle, a system of three degrees of freedom. The three parameters needed to identify the location of the particle are denoted by $\vec{q}$, the vector whose components are equal to its three Cartesian coordinates in Euclidean space. However, the vector $\vec{q}$, as a dynamical variable, does not denote a point in physical space but a point in *configuration space*. For two particles, a system of 6 degrees of freedom, the 6 needed dynamical variables are the Cartesian coordinates $\vec{q}_1$ and $\vec{q}_2$ which identify the locations of particles #1 and #2 in physical space and which make up a vector in the six-dimensional configuration space. Only for a single particle are the dimensions of the physical space equal to those of the configuration space; even in that case, however, the two spaces are distinct and must not be confused.

So the dynamical variables of a single particle in its configuration space are the three components of $\vec{q}$, and the corresponding mutually commuting (see Section 2.1.3) Hermitian operators will be denoted by $\vec{\mathbf{q}}$. We are now going to construct their spectra.

Let the eigenvector or quasi-eigenvector of the three components of $\vec{\mathbf{q}}$ with the eigenvalues or quasi-eigenvalues $\vec{q}$ be $|\vec{q}\rangle$, so that $\vec{\mathbf{q}}|\vec{q}\rangle = \vec{q}|\vec{q}\rangle$. Using the canonical commutation relation (1.12), we then find for any small real ϵ, to first order in ϵ, that

$$\mathbf{q}_l \left[\mathbb{1} - i(\epsilon/\hbar)\mathbf{p}_l\right] |q_l\rangle = [\mathbf{q}_l - i(\epsilon/\hbar)\mathbf{p}_l\mathbf{q}_l + \epsilon]|q_l\rangle = (q_l + \epsilon)\left[\mathbb{1} - i(\epsilon/\hbar)\mathbf{p}_l\right] |q_l\rangle,$$

which implies that $[\mathbb{1} - i(\epsilon/\hbar)\mathbf{p}_l]|q_l\rangle$ is an eigenvector or quasi-eigenvector of $\mathbf{q}_l$ with the eigenvalue or quasi-eigenvalue $q_l + \epsilon$,

$$\left[\mathbb{1} - i(\epsilon/\hbar)\mathbf{p}_l\right] |q_l\rangle = |q_l + \epsilon\rangle. \tag{1.25}$$

This equation, for the derivation of which we used nothing but the commutation relations (1.12), allows us to draw three important conclusions. The first is that the spectrum of all three components of $\vec{\mathbf{q}}$ must be continuous: if q_l is a quasi-eigenvalue of the l-component $\mathbf{q}_l$ of the particle-position operator, then so is $q_l + \epsilon$ for any small real ϵ. Furthermore, this spectrum must fill the real line from $-\infty$ to $+\infty$, because the only way in which (1.25) could break down is for $[\mathbb{1} - i(\epsilon/\hbar)\mathbf{p}_l]|q_l\rangle$ to vanish. But this would require that the Hermitian operator $\epsilon\mathbf{p}_l/\hbar$ have the quasi-eigenvalue $-i$, which is impossible. Hence the spectrum of $\mathbf{q}_l$ has to be unbounded above and below.

Since $\vec{\mathbf{q}}$ is Hermitian and its three components commute, the simultaneous quasi-eigenvectors $|\vec{q}\rangle$ of its three components form a quasi-basis, so that any vector $|\,\rangle \in \mathfrak{H}$ can be expanded as in a Fourier integral:

$$|\,\rangle = \int d^3q\, |\vec{q}\rangle\langle\vec{q}\,|\,\rangle, \tag{1.26}$$

an equation that we can write symbolically as

$$\int d^3q\, |\vec{q}\rangle\langle\vec{q}| = \mathbb{1}. \tag{1.27}$$

Since $\langle\vec{q}|\vec{\mathbf{q}}|\,\rangle = \vec{q}\langle|\,\rangle$, the operator $\vec{\mathbf{q}}$ in this representation is simply the numerical vector $\vec{q}$.

The second conclusion we can draw from (1.25) is that, since a rearrangement of it reads

$$i\hbar\frac{|q_l+\epsilon\rangle - |q_l\rangle}{\epsilon} = \mathbf{p}_l|q_l\rangle,$$

we find in the limit as $\epsilon \to 0$, $\mathbf{p}_l|q_l\rangle = i\hbar\frac{\partial}{\partial q_l}|q_l\rangle$, or

$$\vec{\mathbf{p}}|\vec{q}\rangle = i\hbar\nabla_q|\vec{q}\rangle.$$

Therefore we obtain for matrix elements $\langle\,|\vec{\mathbf{p}}|\vec{q}\rangle = \langle\,|i\hbar\nabla_q|\vec{q}\rangle = i\hbar\nabla_q\langle\,|\vec{q}\rangle$; but $\langle\,|\vec{q}\rangle$ is the complex conjugate of the configuration representation of the vector $|\,\rangle$, that is, of the *configuration-space wave function*

$$\psi(\vec{q}) \stackrel{\text{def}}{=} \langle\vec{q}|\,\rangle$$

representing the state $|\,\rangle$. We have therefore found that in the configuration representation, the momentum operator is given by $\vec{\mathbf{p}} = -i\hbar\nabla_q$,

$$\vec{\mathbf{p}}\psi(\vec{q}) = -i\hbar\nabla_q\psi(\vec{q}), \tag{1.28}$$

and the expectation value of the momentum is calculated by means of the formula

$$\langle\vec{\mathbf{p}}\rangle = \int d^3q\, \psi(\vec{q})^*(-i\hbar\nabla_q)\psi(\vec{q}),$$

assuming that $\int d^3q|\psi(\vec{q})|^2 = 1$.

The third conclusion we can draw from (1.25) is to extend it to all finite shifts a of q by setting $\epsilon = a/n$, applying (1.25) n times, and letting $n \to \infty$. This leads to the result

$$|\vec{q}+\vec{a}\rangle = \lim_{n\to\infty}[\mathbb{1} - i(\vec{\mathbf{p}}\cdot\vec{a}/n\hbar)\vec{\mathbf{p}}]^n\,|\vec{q}\rangle = \exp(-i\vec{\mathbf{p}}\cdot\vec{\mathbf{a}}/\hbar)|\vec{q}\rangle, \tag{1.29}$$

which identifies the momentum operator $\vec{\mathbf{p}}$ as the *generator of translations.* Note that a power series expansion of the exponential in (1.29) and the use of (1.28) yields the Maclaurin-series of the left-hand side.

In the configuration representation, the Hilbert space $\mathfrak{H}$ of a one-particle system is the space of square-integrable functions of three real variables (also called $L^2(\mathbb{R}^3)$), and the state vector of the system is represented by a square-integrable function $\psi(\vec{q})$ of the particle's position $\vec{q}$ in its configuration space, with an inner product defined by

$$(\Phi,\Psi) = \int d^3q\, \phi^*(\vec{q})\psi(\vec{q}). \tag{1.30}$$

The physical interpretation of the wave function $\psi(\vec{q})$ is that of a probability-density amplitude, in the sense that the probability of finding the particle in the infinitesimal neighborhood d^3q of the point $\vec{q}$ in its configuration space is given by $|\psi(\vec{q})|^2 d^3q$. It is also customary to define the *probability-current density* of a particle of mass M in its configuration space by

$$\vec{j}(\vec{q}) \stackrel{\text{def}}{=} -\frac{i\hbar}{2M}[\psi^*(\vec{q})\nabla\psi(\vec{q}) - (\nabla\psi^*(\vec{q}))\psi(\vec{q})], \tag{1.31}$$

an expression which makes good physical sense in that the operator $-i\hbar\nabla/M$ has the significance of the velocity $\vec{\mathbf{p}}/M$, and the two terms are needed to make the current density real. The *density matrix* in the configuration representation is given in terms of the density operator by

$$\boldsymbol{\rho}(\vec{q}, \vec{q}') = \langle \vec{q}|\boldsymbol{\rho}|\vec{q}'\rangle = \sum\!\!\!\!\!\!\int_n p_n \psi_n(\vec{q})\psi_n^*(\vec{q}'). \tag{1.32}$$

It is extremely important to keep in mind that the wave function ψ is not a function on physical space but on the *configuration space* of a particle; ψ must certainly not be thought of as a condition of space. [The configuration representation is also often referred to as "coordinate representation"; this name, however, is misleading, because the "coordinate-space" wave function is *not* defined on physical space but on the configuration space of particles; for two particles, it is a function of *six* variables (see below) and thus lives on a six-dimensional space, the configuration space of the two particles. In order to emphasize this point, we denote the particle coordinates as $\vec{q}$ rather than $\vec{r}$ in this chapter.] Without a clear recognition of the distinction between physical space and configuration space one is easily led into puzzling paradoxes, particularly in connection with the description of measurements. Remember that after a system in a pure state Ψ is subjected to a measurement of an observable A with the outcome A, its new state is an eigenvector Ψ_A of $\mathbf{A}$ with the eigenvalue A, i.e., its state is instantly changed from Ψ to Ψ_A. If this basic assumption of quantum physics is described in terms of "coordinate space" wave functions, it leads to the notorious "collapse of the wave function," which has been endlessly discussed over the years because it appears to be an instantaneous action-at-a-distance effect characteristic of the quantum world. In reality it should be recognized as the inevitable effect of the probabilistic nature of quantum physics. *Whenever a probabilistic prediction is confronted with an individual experimental result, the probability distribution instantaneously "collapses" to a single point.* If you are playing roulette and bet your last \$10 on #17, when the ball comes to rest, your state instantly changes from one in which the probability of owning \$360 is 1/36, either to a state in which you definitely own \$360 or to a state in which you are broke. This becomes puzzling only if wave functions or probability densities are thought to be analogous to electric fields, defined as conditions of physical space, so that the collapse appears to be a long-range physical effect.

For n particles these results are easily generalized. The configuration space is now $3n$-dimensional, and $\mathfrak{H}$ is the space of square-integrable functions of $3n$ variables $\vec{q}_k$, $k = 1, \ldots, n$, while the momentum of the k^{th} particle is represented by the differential operator $\vec{\mathbf{p}}_k = -i\hbar\nabla_{q_k}$, and $|\psi(\vec{q}_1, \ldots, \vec{q}_n)|^2 d^3q_1 \cdots d^3q_n$ is the probability for the particles to be found in the volume element $d^3q_1 \cdots d^3q_n$ near the point $\vec{q}_1, \ldots, \vec{q}_n$ of their joint configuration space. The probability density of one of these particles is obtained from this by integrating $|\psi(\vec{q}_1, \ldots, \vec{q}_n)|^2$ over the coordinates of the others. More generally, if the n-particle system is in a pure state with the (normalized) configuration wave function $\psi(\vec{q}_1, \ldots \vec{q}_n)$, then particle #1, considered by itself and isolated, is in a mixed state and its density matrix is given by

$$\boldsymbol{\rho}(\vec{q}_1, \vec{q}_1') = \int d^3q_2 \cdots d^3q_n \, \psi(\vec{q}_1, \vec{q}_2, \ldots)\psi^*(\vec{q}_1', \vec{q}_2, \ldots). \tag{1.33}$$

Since the Hilbert space of an n-particle system is the tensor product of n Hilbert spaces of single-particle systems, every n-particle wave function can be expanded in a series of products of single-particle wave functions in the form

$$\psi(\vec{q}_1, \ldots, \vec{q}_n) = \sum_{k_1, \ldots, k_n} a_{k_1, \ldots, k_n} \psi_{k_1}(\vec{q}_1) \cdots \psi_{k_n}(\vec{q}_n), \tag{1.34}$$

if the functions $\psi_k(\vec{q})$ form a basis in the one-particle Hilbert space.[14]

1.4.2 The momentum representation

Since all our results concerning the configuration representation were based on the commutation relation (1.12), which, apart from a sign, is symmetrical between $\vec{\mathbf{p}}$ and $\vec{\mathbf{q}}$, we can draw similar conclusions for the momentum representation. Thus we find that the spectra of the three components of $\vec{\mathbf{p}}$ must generally be continuous, running from $-\infty$ to $+\infty$, and a one-particle system's Hilbert space can be represented by the space of square-integrable functions $\widehat{\psi}(\vec{p}) \stackrel{\text{def}}{=} \langle \vec{p} | \, \rangle$ representing the states $\Psi = | \, \rangle$, with the inner product

$$(\Phi, \Psi) = \int d^3p \, \widehat{\phi}^*(\vec{p})\widehat{\psi}(\vec{p}). \tag{1.35}$$

The momentum operator is here represented simply by the numerical vector $\vec{p}$ multiplying $\widehat{\psi}(\vec{p})$, and the operator $\vec{\mathbf{q}}$, operating on a momentum-space wave function, is represented by the differential operator

$$\vec{\mathbf{q}}\widehat{\psi}(\vec{p}) = i\hbar\nabla_p \widehat{\psi}(\vec{p}),$$

[14] **As an exercise, prove that $\Psi^{\mathrm{I}} \otimes \Psi^{\mathrm{II}}$ is mapped into the product $\psi^{\mathrm{I}}\psi^{\mathrm{II}}$ of the corresponding wave functions.**

so that the expectation value of $\vec{q}$ is given by

$$\langle \vec{\mathbf{q}} \rangle = \int d^3p\, \widehat{\psi}(\vec{p})^* (i\hbar \nabla_p) \widehat{\psi}(\vec{p}).$$

The physical interpretation of the momentum-space wave function is analogous to that in configuration space: $|\widehat{\psi}(\vec{p})|^2 d^3p$ is the probability for the particle to have its momentum in the neighborhood d^3p of $\vec{p}$.

Again, these things are easily generalized to n particles. The Hilbert space $\mathfrak{H}$ is the space of square-integrable functions of $3n$ variables $\vec{p}_k,\ k = 1, \ldots, n$, and the position of the k^{th} particle in its configuration space is represented by the differential operator $\vec{\mathbf{q}}_k = i\hbar \nabla_{p_k}$. The physical interpretation of $\widehat{\psi}(\vec{p}_1, \ldots, \vec{p}_n)$ is that of a probability-density amplitude: the probability that the momenta of the particles lie in the neighborhood $d^3p_1 \cdots d^3p_n$ of the point $\vec{p}_1, \ldots, \vec{p}_n$ is $|\widehat{\psi}(\vec{p}_1, \ldots, \vec{p}_n)|^2 d^3p_1 \cdots d^3p_n$.

There is, however, an important exception to the continuity of the spectrum of the momentum operator, if the particle is confined to a finite spatial region such as a box. Let us, for simplicity, take a one-dimensional situation and assume the "box" has length b. Such a case is handled most simply by repeating the "box" periodically, setting $|q + b\rangle = e^{i\varphi}|q\rangle$, where φ is a constant. As a result we obtain by (1.29)

$$\langle p|q + b\rangle = e^{i\varphi}\langle p|q\rangle = \langle p|e^{-ib\mathbf{p}/\hbar}|q\rangle = e^{-ibp/\hbar}\langle p|q\rangle,$$

from which we may conclude that $bp/\hbar = 2\pi n - \varphi$; the spectrum of $\mathbf{p}$ must therefore consist of the points $p_n = \frac{2\pi\hbar}{b} n + \text{const.}$, $n = 0, \pm 1, \pm 2, \ldots$. Thus confining the particle to a "box" has discretized the momentum spectrum. The same result, of course, holds in three dimensions: if the coordinate axes are chosen along the sides of the box with orthogonal edges of lengths $b_i,\ i = 1, 2, 3$, the i^{th} component of the momentum $\mathbf{p}_i$ has the eigenvalues $p_{i,n_i} = \frac{2\pi\hbar}{b_i} n_i + \text{const.}, \quad n_i = 0, \pm 1, \pm 2, \ldots$. The integral in (B.23) then has to be replaced by a sum, quite analogous to the fact that, whereas Fourier integrals are used to analyze functions defined on an infinite region, the appropriate tool for functions on finite intervals is the Fourier series. (In that case, too, any extension of the series outside the interval of definition is necessarily periodic.)

1.4.3 The number representation

Just as (1.12) was the essential tool for determining the spectrum and representations of $\vec{\mathbf{q}}$ and $\vec{\mathbf{p}}$, so to get at the properties of the field operator Ψ, we shall use the commutation relations

$$[\Psi(\vec{r}), \Psi^\dagger(\vec{r}')] = \delta^3(\vec{r} - \vec{r}'), \quad [\Psi(\vec{r}), \Psi(\vec{r}')] = 0, \tag{1.36}$$

which will be justified, along with (1.12), in Chapter 2 in the context of dynamics. (We do not explicitly indicate the fact that the field operators generally depend on the time; they are all taken at the same time.)

First of all, it will facilitate the physical interpretation to perform a Fourier transformation,

$$\Psi(\vec{r}) = (2\pi)^{-3/2} \int d^3k\, e^{i\vec{k}\cdot\vec{r}} \widehat{\Psi}(\vec{k}), \tag{1.37}$$

so that

$$\widehat{\Psi}(\vec{k}) = (2\pi)^{-3/2} \int d^3r\, e^{-i\vec{k}\cdot\vec{r}} \Psi(\vec{r}), \tag{1.38}$$

which casts the commutation relations into

$$[\widehat{\Psi}(\vec{k}), \widehat{\Psi}^\dagger(\vec{k}')] = \delta^3(\vec{k} - \vec{k}'), \quad [\widehat{\Psi}(\vec{k}), \widehat{\Psi}(\vec{k}')] = 0. \tag{1.39}$$

It will also be helpful for the moment to confine the system to a large box, thereby replacing the Fourier integral by a Fourier series, so that $\widehat{\Psi}(\vec{k})$ is replaced by $\widehat{\Psi}_i$ and (1.39) takes the form

$$[\widehat{\Psi}_i, \widehat{\Psi}_j^\dagger] = \delta_{ij}, \quad [\widehat{\Psi}_i, \widehat{\Psi}_j] = 0. \tag{1.40}$$

Now consider the Hermitian operator

$$\mathsf{N}_i \stackrel{\text{def}}{=} \widehat{\Psi}_i^\dagger \widehat{\Psi}_i, \tag{1.41}$$

which, according to (1.40), satisfies the commutation relations

$$[\mathsf{N}_i, \widehat{\Psi}_i] = -\widehat{\Psi}_i, \quad [\mathsf{N}_i, \widehat{\Psi}_i^\dagger] = \widehat{\Psi}_i^\dagger, \tag{1.42}$$

and let $|n_i\rangle$ be an eigenvector of N_i with the eigenvalue n_i. It follows from (1.42) that

$$\mathsf{N}_i \widehat{\Psi}_i |n_i\rangle = (\widehat{\Psi}_i \mathsf{N}_i - \widehat{\Psi}_i)|n_i\rangle = (n_i - 1)\widehat{\Psi}_i |n_i\rangle,$$

which implies that $\widehat{\Psi}_i|n_i\rangle$ is an eigenvector of N_i with the eigenvalue $n_i - 1$. Therefore, whenever n_i is an eigenvalue of N_i, then so is $n_i - 1$, except when $\widehat{\Psi}_i|n_i\rangle = 0$. But the spectrum of the operator N_i is non-negative, because for all $\Psi \in \mathfrak{H}$, $(\Psi, \mathsf{N_i}\Psi) = (\Psi, \widehat{\Psi}_i^\dagger \widehat{\Psi}_i \Psi) = (\widehat{\Psi}_i \Psi, \widehat{\Psi}_i \Psi) > 0$, except when $\widehat{\Psi}_i \Psi = 0$, in which case $\mathsf{N}_i \Psi = 0$. Consequently, the downward sequence, in integral steps, of eigenvalues of N_i must terminate before it becomes negative, and that can happen only when $\widehat{\Psi}_i \Psi = 0$, which also means that $\mathsf{N}_i \Psi = 0$. We therefore conclude that the eigenvalues of N_i must be the non-negative integers $0, 1, \ldots$. Similarly we find from (1.42) that

$$\mathsf{N}_i \widehat{\Psi}_i^\dagger |n_i\rangle = (\widehat{\Psi}_i^\dagger \mathsf{N}_i + \widehat{\Psi}_i^\dagger)|n_i\rangle = (n_i + 1)\widehat{\Psi}_i^\dagger |n_i\rangle.$$

This sequence can terminate at the upper end only at some value of n_i for which $\widehat{\Psi}_i^\dagger |n_i\rangle = 0$, which would imply that $0 = \langle n_i | \widehat{\Psi}_i \widehat{\Psi}_i^\dagger |n_i\rangle = \langle n_i | \mathsf{N}_i + \mathbb{1} | n_i \rangle = (n_i + 1)\langle n_i | n_i \rangle$, and therefore $|n_i\rangle = 0$. It follows that the sequence

cannot terminate and the eigenvalues of N_i are *all* the non-negative integers.

As for the normalization, if $|n_i\rangle$ and $|n_i+1\rangle$ are normalized, then $c|n_i+1\rangle = \widehat{\Psi}_i^\dagger|n_i\rangle$, and hence $|c|^2 = \langle n_i|\widehat{\Psi}_i\widehat{\Psi}_i^\dagger|n_i\rangle = \langle n_i|\mathsf{N}_i + \mathbb{1}|n_i\rangle = n_i + 1$; hence we may choose $c = \sqrt{n_i+1}$, so that

$$\widehat{\Psi}_i^\dagger|n_i\rangle = \sqrt{n_i+1}|n_i+1\rangle. \tag{1.43}$$

We then have $|1_i\rangle = \widehat{\Psi}_i^\dagger|0\rangle$, $|2_i\rangle = \frac{1}{\sqrt{2}}\widehat{\Psi}_i^\dagger|1_i\rangle = \frac{1}{\sqrt{2}}(\widehat{\Psi}_i^\dagger)^2|0\rangle$, $|3_i\rangle = \frac{1}{\sqrt{3}}\widehat{\Psi}_i^\dagger|2_i\rangle = \frac{1}{\sqrt{3!}}\widehat{\Psi}_i^\dagger|0\rangle$, and

$$|n_i\rangle = \frac{1}{\sqrt{n_i!}}(\widehat{\Psi}_i^\dagger)^{n_i}|0\rangle, \tag{1.44}$$

where $|0\rangle$ is determined by

$$\widehat{\Psi}_i|0\rangle = 0 \tag{1.45}$$

for all i, and we assume $|0\rangle$ to be normalized. We will tentatively interpret this result (subject to later verification that this makes sense as far as the energy is concerned) by saying that $|n_i\rangle$ is the state that has n quanta of wave vectors $\vec{k}_i$ (which, as we shall see in Chapter 2, means they have momenta $\hbar\vec{k}_i$) and no others. The state $|0\rangle$ is the *vacuum state;* it has no particles. Since all the N_i commute with one another,[15]

$$[\mathsf{N}_i, \mathsf{N}_j] = 0, \tag{1.46}$$

there is a complete set of common eigenfunctions.

A general one-quantum state is now given by $\Psi_1 = |\,\rangle_1 = \sum_i a_i\widehat{\Psi}_i^\dagger|0\rangle$, so that the probability of finding the particle to have momentum $\hbar\vec{k}_i$ is, by (1.40) and (1.45),

$$|\langle 1_i|\sum_j a_j\widehat{\Psi}_j^\dagger|0\rangle|^2 = |\langle 0|\sum_j a_j\widehat{\Psi}_i\widehat{\Psi}_j^\dagger|0\rangle|^2 = |a_i|^2.$$

If we abandon the Fourier series, going back to physical space, the general one-particle state is given by

$$\Psi_1 = |\,\rangle_1 = \int d^3r\, f(\vec{r})\Psi^\dagger(\vec{r})|0\rangle, \tag{1.47}$$

and the probability of finding the quantum in the region d^3q in its configuration space is correspondingly given by $|f(\vec{q})|^2d^3q$. By (1.36), the inner product with another one-particle state $\Phi_1 = \int d^3q\, g(\vec{q})\Psi(\vec{q})|0\rangle$, is given by

$$(\Phi_1, \Psi_1) = \int d^3q\, d^3q'\, g^*(\vec{q}')f(\vec{q})\langle 0|\Psi(\vec{q}')\Psi^\dagger(\vec{q})|0\rangle = \int d^3q\, g^*(\vec{q})f(\vec{q}).$$

[15] **Prove this as an exercise.**

Hence we may conclude that the function $f(\vec{q})$ is nothing but the configuration-space wave function of the one-particle state $\Psi_1 = |\,\rangle_1$, and we have

$$|\vec{q}\rangle = \Psi^\dagger(\vec{q})|0\rangle, \tag{1.48}$$

as well as

$$\psi(\vec{q}) = \langle\vec{q}|\,\rangle_1 = \langle 0|\Psi(\vec{q})|\,\rangle_1. \tag{1.49}$$

For a two-particle state the configuration wave function is obtained in the same manner,

$$\psi(\vec{q}_1, \vec{q}_2) = \langle 0|\Psi(\vec{q}_1)\Psi(\vec{q}_2)|\,\rangle_2. \tag{1.50}$$

Notice that it follows from (1.36) that the right-hand side of this equation is symmetric under an exchange of $\vec{q}_1$ and $\vec{q}_2$, and so must therefore be the left-hand side. In fact, the variables $\vec{q}_1$ and $\vec{q}_2$ do not refer to the coordinates of two different particles at all; they refer, instead, to the two *locations in configuration space that are occupied* by quanta. (This is why the number representation is also referred to as the *occupation-number representation.*) Thus the number representation, obtained from the field, automatically has built into it one of the basic differences between classical and quantum particles, namely, their *indistinguishability,* which has profound consequences for their statistics: instead of being subject to Maxwell-Boltzmann statistics, they obey those of Bose and Einstein (see Chapter 9). You should notice that whereas the field operator $\Psi(\vec{r})$ lives on physical space—it denotes a condition of space—at the point of converting its action on the vacuum state, as in (1.50), to a wave function, the argument of the field operator changes to denoting a point in configuration space occupied by a particle; that's why the variables in (1.50) were changed from $\vec{r}$ to $\vec{q}$.

There is, however, an alternative to the commutation relation (1.36) for the field, and that is

$$\{\Psi(\vec{r}), \Psi^\dagger(\vec{r}')\} = \delta(\vec{r} - \vec{r}'), \quad \{\Psi(\vec{r}), \Psi(\vec{r}')\} = 0, \tag{1.51}$$

where $\{\mathbf{A}, \mathbf{B}\} \stackrel{\text{def}}{=} \mathbf{AB} + \mathbf{BA}$ is the *anti-commutator,* or for the Fourier-transformed operators in the discrete form,

$$\{\widehat{\Psi}_i, \widehat{\Psi}_j^\dagger\} = \delta_{ij}, \quad \{\widehat{\Psi}_i, \widehat{\Psi}_j\} = 0, \tag{1.52}$$

the second equation of which implies that $\widehat{\Psi}_i^2 = 0$. Again defining $\mathsf{N}_i = \widehat{\Psi}_i^\dagger\widehat{\Psi}_i$, we now obtain $\mathsf{N}_i\widehat{\Psi}_i = 0$, $\widehat{\Psi}_i\mathsf{N}_i = \widehat{\Psi}_i$, and $\mathsf{N}_i^2 = \mathsf{N}_i$. The last equation implies that $\mathsf{N_i}$ has only two eigenvalues, 0 and 1 (because the eigenvalues have to obey the same equation, $n_i^2 = n_i$), while we get from the first $\widehat{\Psi}_i|0\rangle = 0$ and from the Hermitian adjoint of the first, $\mathsf{N_i}\widehat{\Psi}_i^\dagger|0\rangle = \widehat{\Psi}_i^\dagger|0\rangle$, which implies that $\widehat{\Psi}_i^\dagger|0\rangle = c|1_i\rangle$. Assuming both $|0\rangle$ and $|1_i\rangle$ to be normalized, we find furthermore that $|c|^2 = \langle 0|\widehat{\Psi}_i\widehat{\Psi}_i^\dagger|0\rangle = \langle 0|0\rangle = 1$. So the two normalized eigenvectors of N_i are defined by

$$\widehat{\Psi}_i|0\rangle = 0, \qquad |1_i\rangle = \widehat{\Psi}_i^\dagger|0\rangle, \tag{1.53}$$

while $\widehat{\Psi}_i^\dagger|1_i\rangle = 0$. Again, all the number operators commute, i.e., (1.46) holds, and there is a complete set of common eigenvectors.

Returning to space coordinates, we may proceed as before to define one-particle configuration wave functions by

$$\psi(\vec{q}) = \langle\vec{q}|\,\rangle = \langle 0|\Psi(\vec{q})|\,\rangle_1 \tag{1.54}$$

and two-particle wave functions by

$$\psi(\vec{q}_1, \vec{q}_2) = \langle 0|\Psi(\vec{q}_2)\Psi(\vec{q}_1)|\,\rangle_2. \tag{1.55}$$

Now, however, this wave function is *anti-symmetric* in its two arguments rather than symmetric; furthermore, because $\psi(\vec{q}_1, \vec{q}_1) = 0$, the two particles satisfy the *exclusion principle.* In the aggregate, they therefore obey Fermi-Dirac rather than Bose-Einstein statistics. We shall return to a more detailed discussion of multi-particle systems and their properties in Chapter 9.

What have we learned from the introduction of the number representation? Whereas earlier we had started by assuming the existence of point particles and described systems by means of the positions and momenta of these particles as dynamical variables, the new point of departure was the quantum field. The operator properties of this field then generated quanta or particles on its own, and the description of physical systems ended up exactly as before. It may therefore be deemed a matter of choice whether to regard the particles as the fundamental entities, or the field, or both, as was the classical approach. The quantum field description, however, automatically produces particles with quantum properties such as indistinguishability and the two alternatives of symmetric and anti-symmetric wave functions, properties that would otherwise have to be artificially introduced into the formalism as separate postulates. The field should therefore be considered the more fundamental starting point for the quantum description of nature, with particles arising as a result of the commutation rules. Indeed the entire notion of the existence of stable objects called particles is a consequence of the properties of the quantum field.

1.5 Transformation Theory

Suppose that we have a representation $\{a_n = \langle A_n|\,\rangle,\ n = 1, 2, \ldots,\}$ of a given vector $|\,\rangle$ on an orthonormal basis $|A_n\rangle$, i.e., $|\,\rangle = \sum_n a_n|A_n\rangle$, so that all states of a system are represented in terms of their probability amplitudes for finding the results A_n upon measurement of the dynamical variable A. In order to understand the system, we have to be able to relate this given representation to others, based on measurements of other variables B. For instance, B may be the same physical parameter as A,

but expressed with respect to another coordinate frame, or, to allow us to describe the dynamics of the system, B may be the same variable A at a later time, or else it may be quite unrelated to A.

To represent $|\,\rangle$ on another basis of orthonormal vectors $|B_n\rangle$, i.e., to use the probability amplitudes for the measurement outcomes of another variable B rather than A, takes a *unitary transformation*, which we find by expanding each new basis vector on the old basis:

$$|B_n\rangle = \sum_m |A_m\rangle U^{(12)}_{mn}, \tag{1.56}$$

where

$$U^{(12)}_{mn} \stackrel{\text{def}}{=} \langle A_m|B_n\rangle. \tag{1.57}$$

That this transformation is unitary is easily checked by using (1.23):

$$(U^{(12)}U^{(12)\dagger})_{mn} = \sum_k \langle A_m|B_k\rangle\langle B_k|A_n\rangle = \langle A_m|A_n\rangle = \delta_{mn},$$

and similarly for the matrix $U^{(12)\dagger}U^{(12)}$.[16]

Another way of looking at this basis transformation (see also Appendix B) is to regard $\mathsf{U}^{(12)}$ as the operator that maps every basis vector of the first basis onto its counterpart in the second,

$$\mathsf{U}^{(12)}|A_n\rangle \stackrel{\text{def}}{=} |B_n\rangle. \tag{1.58}$$

As (1.57) shows, the operator so defined has $U^{(12)}_{mn}$ as its matrix elements in the basis $\{|A_n\rangle\}$:

$$U^{(12)}_{mn} = \langle A_m|\mathsf{U}^{(12)}|A_n\rangle.$$

The same transformation then allows us to express the matrix $X^{(2)}_{nm}$ representing the operator $\mathbf{X}$ on the second basis in terms of its representation $X^{(1)}_{nm}$ on the first. With $X^{(2)}_{nm} = \langle B_n|\mathbf{X}|B_m\rangle$, inserting (1.56) yields

$$X^{(2)}_{nm} = \sum_{kl} U^{(12)*}_{kn} X^{(1)}_{kl} U^{(12)}_{lm} = \sum_{kl} (U^{(12)\dagger})_{nk} X^{(1)}_{kl} U^{(12)}_{lm}. \tag{1.59}$$

Since $U^{(12)\dagger} = U^{(12)-1} = U^{(21)}$, this can also be written in the form

$$X^{(2)}_{nm} = \sum_{kl} U^{(21)}_{nk} X^{(1)}_{kl} U^{(12)}_{lm}. \tag{1.60}$$

Alternatively, this transformation may be viewed as using (1.58) and writing

$$X^{(2)}_{nm} = \langle A_n|\mathsf{U}^{(12)\dagger}\mathbf{X}\mathsf{U}^{(12)}|A_m\rangle, \tag{1.61}$$

[16] Remember that in order for an operator U to be unitary it is necessary and sufficient that both $\mathsf{U}\mathsf{U}^\dagger = \mathbf{1}$ and $\mathsf{U}^\dagger\mathsf{U} = \mathbf{1}$. Similarly for an infinite-dimensional matrix.

which means that $X^{(2)}_{nm}$ is the matrix, *on the old basis*, of the unitarily transformed operator $\mathsf{U}^{(12)\dagger}\mathbf{X}\mathsf{U}^{(12)}$. Such tranformations are called *canonical*, and we shall return to them later. If the new basis is an orthonormal set of eigenvectors of $\mathbf{X}$, $\{|B_n\rangle = |X_n\rangle\}$, then the eigenvectors of the transformed operator $\mathsf{U}^\dagger\mathbf{X}\mathsf{U}$, which has the same eigenvalues as $\mathbf{X}$, are $\{|A_n\rangle = \mathsf{U}^\dagger|X_n\rangle\}$, and the transformed operator is diagonal on the old basis.

These transformations work analogously for continuous spectra of Hermitian operators and their quasi-eigenvectors satisfying (B.23) and (B.24). Supposing that another set satisfies

$$\langle B|B'\rangle = \delta(B - B'), \quad \text{and} \quad \int dB\, |B\rangle\langle B| = \mathbb{1}, \tag{1.62}$$

we define an integral kernel[17] or *transformation function* $U^{(12)}(A, B) \stackrel{\text{def}}{=} \langle A|B\rangle$, so that

$$|B\rangle = \int dA\, |A\rangle U^{(12)}(A, B). \tag{1.63}$$

Again, this kernel is unitary in the sense that

$$\int dB\, U^{(12)}(A, B)U^{(12)*}(A', B) = \delta(A - A'),$$

$$\int dA\, U^{(12)*}(A, B)U^{(12)}(A, B') = \delta(B - B'),$$

and it also serves to transform the integral kernel of an operator $\mathbf{X}$, given in the $\mathbf{A}$-representation, into the corresponding integral operator in the $\mathbf{B}$-representation:

$$\begin{aligned} X^{(2)}(B, B') &= \langle B|\mathbf{X}|B'\rangle = \int dA\, dA'\, U^{(12)*}(A, B)\langle A|\mathbf{X}|A'\rangle U^{(12)}(A', B') \\ &= \int dA\, dA'\, U^{(12)*}(A, B)X^{(1)}(A, A')U^{(12)}(A', B'). \end{aligned} \tag{1.64}$$

Suppose now that we are given a Hermitian operator $\mathbf{B}$ and its representation on a basis $\{|A_n\rangle\}$, namely, $B^{(1)}_{nm} = \langle A_n|\mathbf{B}|A_m\rangle$. On the basis of its own orthonormal eigenvectors, the matrix representing $\mathbf{B}$ is diagonal: $\langle B_n|\mathbf{B}|B_m\rangle = B_n\delta_{nm}$. Therefore, the unitary transformation that takes us from the A-representation to the B-representation *diagonalizes* the Hermitian matrix $B^{(1)}_{nm}$:

$$\sum_{kl}(U^{(12)\dagger})_{nk}B^{(1)}_{kl}U^{(12)}_{lm} = B^{(2)}_{nm} = B_n\delta_{mn}. \tag{1.65}$$

[17]Strictly speaking, it is not an integral kernel in the mathematical sense; it usually has a component that is a Dirac delta-function; in other words, it is a distribution.

Thus, for a Hermitian operator given as a matrix in some representation, the problem of finding its eigenvalues and eigenvectors is equivalent to finding the unitary transformation that diagonalizes the matrix, as we already noted earlier.

In view of the formal similarity of the expressions of these transformations, independently of whether the matrix elements are between vectors or quasi-vectors, it is most convenient always to write the transformation functions in the form $\langle A|B\rangle$, irrespective of whether the values of A or B run over discrete or continuous sets. The summations and/or integrations will then be indicated by $\sum\!\!\!\!\!\!\int_A$ with a subscript.

Finally, note the **composition law** obeyed by the transformation functions. If we perform a transformation from one basis (or continuous quasi-basis) to a second, followed by a transformation to a third, the total transformation is the product, in a matrix sense, of the two, as in

$$\sum\!\!\!\!\!\!\int_A \langle B|A\rangle\langle A|C\rangle = \langle B|C\rangle, \tag{1.66}$$

which follows from (B.23) in the Appendix.

Let us take a specific case and construct the transformation function that leads from the configuration representation to the momentum representation of a one-particle system, that is, the function $\langle\vec{p}|\vec{q}\rangle$. In that case we have

$$\langle\vec{p}|\vec{\mathbf{p}}|\vec{q}\rangle = i\hbar\nabla_q\langle\vec{p}|\vec{q}\rangle = \vec{p}\langle\vec{p}|\vec{q}\rangle.$$

This differential equation for $\langle\vec{p}|\vec{q}\rangle$ is easily solved, with the result

$$\langle\vec{p}|\vec{q}\rangle = (2\pi\hbar)^{-3/2}e^{-i\vec{p}\cdot\vec{q}/\hbar}, \tag{1.67}$$

the constant $(2\pi\hbar)^{-3/2}$ being chosen to make the transformation unitary, so that both quasi-bases are δ-normalized. Consequently, we have

$$|\vec{q}\rangle = (2\pi\hbar)^{-3/2}\int d^3p\, e^{-i\vec{q}\cdot\vec{p}/\hbar}|\vec{p}\rangle, \qquad |\vec{p}\rangle = (2\pi\hbar)^{-3/2}\int d^3q\, e^{i\vec{q}\cdot\vec{p}/\hbar}|\vec{q}\rangle, \tag{1.68}$$

which shows that the momentum-space wave function and the configuration-space wave function are simply one another's Fourier transforms:

$$\widehat{\psi}(\vec{p}) = \langle\vec{p}\,|\,\rangle = (2\pi\hbar)^{-3/2}\int d^3q\, e^{-i\vec{q}\cdot\vec{p}/\hbar}\langle\vec{q}\,|\,\rangle = (2\pi\hbar)^{-3/2}\int d^3q\, e^{-i\vec{q}\cdot\vec{p}/\hbar}\psi(\vec{q}), \tag{1.69}$$

$$\psi(\vec{q}) = \langle\vec{q}\,|\,\rangle = (2\pi\hbar)^{-3/2}\int d^3p\, e^{i\vec{q}\cdot\vec{p}/\hbar}\langle\vec{p}\,|\,\rangle = (2\pi\hbar)^{-3/2}\int d^3p\, e^{i\vec{q}\cdot\vec{p}/\hbar}\widehat{\psi}(\vec{p}). \tag{1.70}$$

The configuration representations of the quasi-eigenfunctions of the momentum are nothing but the transformation functions $\langle\vec{q}|\vec{p}\rangle$, which are the complex conjugates of (1.67):

$$\psi_{\vec{p}}(\vec{q}) = \langle\vec{q}|\vec{p}\rangle = (2\pi\hbar)^{-3/2}e^{i\vec{p}\cdot\vec{q}/\hbar}, \tag{1.71}$$

and these are appropriately δ-normalized. Since these wave functions are not square integrable and thus not in the Hilbert space L^2, they do not represent *states* of a particle in the ordinary sense. The fact that their norms are infinite might be interpreted as an indication that they may be regarded as representing a collection of infinitely many particles: infinitely long and wide *beams* of particles with the momentum $\vec{p}$, which is consistent with the fact that (1.71) implies a uniform current density, as defined by (1.31), of $\vec{p}/[M(2\pi\hbar)^3]$.

For particles confined to a line, a one-dimensional space, these equations become

$$\widehat{\psi}(p) = \langle p|\,\rangle = (2\pi\hbar)^{-1/2}\int dq\, e^{-iqp/\hbar}\langle q|\,\rangle = (2\pi\hbar)^{-1/2}\int dq\, e^{-iqp/\hbar}\psi(q), \tag{1.72}$$

$$\psi(q) = \langle q|\,\rangle = (2\pi\hbar)^{-1/2}\int dp\, e^{iqp/\hbar}\langle p|\,\rangle = (2\pi\hbar)^{-1/2}\int dp\, e^{iqp/\hbar}\widehat{\psi}(p), \tag{1.73}$$

and

$$\psi_p(q) = \langle q|p\rangle = (2\pi\hbar)^{-1/2}e^{ipq/\hbar}. \tag{1.74}$$

Analogous relations hold for n-particle systems:

$$\begin{aligned}\langle\vec{p}_1,\ldots,\vec{p}_n|\,\rangle &= (2\pi\hbar)^{-3n/2}\int d^3q_1\cdots d^3q_n\, e^{-i\sum_j \vec{q}_j\cdot\vec{p}_j/\hbar}\langle\vec{q}_1,\ldots,\vec{q}_n|\,\rangle,\\ \langle\vec{q}_1,\ldots,\vec{q}_n|\,\rangle &= (2\pi\hbar)^{-3n/2}\int d^3p_1\cdots d^3p_n\;\; e^{i\sum_j \vec{q}_j\cdot\vec{p}_j/\hbar}\langle\vec{p}_1,\ldots,\vec{p}_n|\,\rangle.\end{aligned}$$

1.6 Problems and Exercises

1. Show how (1.14) implies that p_n are the eigenvalues of the density operator.

2. Calculate the Fourier transform of the function $e^{-\alpha|x|}\cos(\beta x)$, where $\alpha > 0$, and of the function $\frac{1}{\alpha^2+x^2}$.

3. Calculate the three-dimensional Fourier transform of the function $e^{-\alpha r}\cos(\beta r)$, where $\alpha > 0$.

4. A 1-ounce rifle bullet takes 0.5 sec to reach its target. Regarding the bullet as a mass point, and neglecting effects of air resistance and earth motion, find the order of magnitude of the minimal spread of successive shots at the target under optimal conditions of aiming and firing.

5. Suppose you try to balance an ice pick on its point, setting it up under optimal conditions. What is the longest time T you can expect it to stand more or less straight? If, performing this experiment, you manage to keep the pick standing for a longer time than T, have you thereby disproved quantum mechanics?

6. Consider a one-dimensional system near stable equilibrium. If you try to set it up at rest in its equilibrium position, what is its minimum energy on account of Heisenberg's uncertainty principle?

7. A wheel on a frictionless axle is made to rotate freely with a given angular momentum. We wish to predict the angular position of a mark on the wheel after a time t. If the wheel is set in motion under experimentally optimal conditions, how long does it take until the position of the mark is entirely uncertain?

8. Two independent physical systems can each exist in three eigenstates of the operator $\mathbf{C}$, Ψ_i, $i = 1, 2, 3$ with the eigenvalues C_i, i.e., $\mathbf{C}\Psi_i = C_i\Psi_i$. They are combined into a total system S whose state is

$$\Psi = \Psi_1^{(a)} \otimes \Psi_3^{(b)} + \frac{1}{3}\Psi_3^{(a)} \otimes \Psi_1^{(b)} + \frac{1}{5}\Psi_3^{(a)} \otimes \Psi_2^{(b)}.$$

Calculate the probability that a measurement on system a that ignores system b will yield the result C_3. Exhibit the density matrices for system a alone and for system b alone.

2
Quantum Basics: Dynamics

2.1 Time Dependence and Commutation Rules

Now that we have at our disposal the basic tools for the quantum representation of physical systems, we have to turn to the description of their development in the course of time, that is, their dynamics.[1] The basic idea is to adopt for the given system a complete set of commuting observables, which will be assumed schematically to be a single Hermitian operator $\mathbf{X}$, and let this operator change in the course of time so as to describe the development of the system. We will assume the function $\mathbf{X}(t)$ to be such that the *possible* outcomes of experiments do not depend on t, which implies that the spectrum of $\mathbf{X}(t)$ is time-independent. This does not mean, of course, that the actual results of measuring the observable X corresponding to $\mathbf{X}$ on a given system do not change in the course of time; it is precisely this change that we want to determine. In this way of looking at the quantum dynamics, called the **Heisenberg picture,** all the time dependence is carried by the dynamical variables, and the vector representing the state of a developing system remains unaltered. However, if we fix our attention on a given quasi-eigenstate of the operator $\mathbf{X}(t)$ with a specific quasi-eigenvalue X, that quasi-statevector will depend on t because $\mathbf{X}$ does.

So let $|X, t\rangle$ be a quasi-eigenstate of $\mathbf{X}(t)$ with the quasi-eigenvalue X. These quasi-eigenstates form a quasi-basis in $\mathfrak{H}$, and we focus on the trans-

[1] For a different approach, known as *path integrals*, to generating the quantum-mechanical equations of motion, see [Feynman].

formation function $\langle X_2, t_2|X_1, t_1\rangle$, where X_1 and X_2 are quasi-eigenvalues of the operators $\mathbf{X_1}(t_1)$ and $\mathbf{X}_2(t_2)$; the latter may have developed from $\mathbf{X}_1(t_1)$ in the course of time, but not necessarily according to the physically correct equations of motion. We are looking for a principle that allows us to determine what the correct equations are. As an essential tool for this search we will subject the development of the system from the time t_1 to the time t_2, as well as possibly the endpoint variables, such as t_1, t_2, $\mathbf{X}_1(t_1)$, and $\mathbf{X}_2(t_2)$, to infinitesimal variations. The corresponding change of the transformation function can then be used to define an infinitesimal operator $\delta\mathsf{W}_{21}$ with the dimension of an action so that

$$\delta\langle X_2, t_2|X_1, t_1\rangle \stackrel{\text{def}}{=} \frac{i}{\hbar}\langle X_2, t_2|\delta\mathsf{W}_{21}|X_1, t_1\rangle.$$

The composition law (1.66),

$$\langle 3|1\rangle = \sum\!\!\!\!\!\!\int_{X_2} \langle 3|X_2, t_2\rangle\langle X_2, t_2|1\rangle,$$

which holds for any arbitrary intervening time t_2 and where we have written for simplicity $|X_1, t_1\rangle \stackrel{\text{def}}{=} |1\rangle$, etc., therefore leads to

$$\begin{aligned}
\delta\langle 3|1\rangle &= \sum\!\!\!\!\!\!\int_{X_2} [\delta\langle 3|2\rangle\langle 2|1\rangle + \langle 3|2\rangle\delta\langle 2|1\rangle] \\
&= \frac{i}{\hbar}\sum\!\!\!\!\!\!\int_{X_2} [\langle 3|\delta\mathsf{W}_{32}|2\rangle\langle 2|1\rangle + \langle 3|2\rangle\langle 2|\delta\mathsf{W}_{21}|1\rangle] \\
&= \frac{i}{\hbar}[\langle 3|\delta\mathsf{W}_{32}|1\rangle + \langle 3|\delta\mathsf{W}_{21}|1\rangle] = \frac{i}{\hbar}\langle 3|\delta\mathsf{W}_{32} + \delta\mathsf{W}_{21}|1\rangle,
\end{aligned}$$

from which we conclude that

$$\delta\mathsf{W}_{31} = \delta\mathsf{W}_{32} + \delta\mathsf{W}_{21}.$$

It follows that $\delta\mathsf{W}_{21} = -\delta\mathsf{W}_{12}$, since $\delta\mathsf{W}_{11} = 0$. Furthermore, since $\langle 2|1\rangle = \langle 1|2\rangle^*$, we have $\delta\langle 2|1\rangle = \frac{i}{\hbar}\langle 2|\delta\mathsf{W}_{21}|1\rangle = \delta\langle 1|2\rangle^* = -\frac{i}{\hbar}\langle 1|\delta\mathsf{W}_{12}|2\rangle^* = -\frac{i}{\hbar}\langle 2|\delta\mathsf{W}_{12}^\dagger|1\rangle = \frac{i}{\hbar}\langle 2|\delta\mathsf{W}_{21}^\dagger|1\rangle$, which implies that

$$\delta\mathsf{W}_{21} = \delta\mathsf{W}_{21}^\dagger.$$

2.1.1 The principle of stationary action

We now make the crucial assumption that if $\mathbf{X_1}(t_1)$ turned into $\mathbf{X_2}(t_2)$ according to some given time-development and $\delta\mathsf{W}$ is an infinitesimal variation *in the neighborhood of a "trajectory" compatible with the correct equations of motion and commutation relations,* then $\delta\mathsf{W}$ is the first variation of an operator W that satisfies the corresponding relations

$$\mathsf{W}_{12}^\dagger = \mathsf{W}_{12}, \qquad \mathsf{W}_{31} = \mathsf{W}_{32} + \mathsf{W}_{21}.$$

Cutting the interval from t_1 to t_2 into n subintervals then leads to $\mathsf{W}_{21} = \sum_i \mathsf{W}_{i\,i-1}$ and in the limit as $n \to \infty$, to

$$\mathsf{W}_{21} = \int_{t_1}^{t_2} dt\, \mathsf{L}[t], \tag{2.1}$$

for some $\mathsf{L}[t]$, which is called the *Lagrangian* operator. In order for W, the *action integral,* to be Hermitian, L has to be Hermitian too.

Suppose, then, that the system is described by a number of independent dynamical variables which we schematically call q_i. (These may be the coordinates of point particles, but they need not be.) For the quantum description they are Hermitian operators on $\mathcal{H}$, and we assume that the Lagrangian is a function only of these q_i and their first time derivatives, so that $\mathsf{L}[t] = \mathsf{L}(\mathsf{q}_1, \ldots, \mathsf{q}_i, \ldots; \dot{\mathsf{q}}_1, \ldots, \dot{\mathsf{q}}_i, \ldots)$, implying that the time-dependence of L comes entirely from the fact that the qs and $\dot{\mathsf{q}}$s depend on the time. All of this is quite analogous to the usual assumptions in classical mechanics. Here, however, we have to worry about the order of the qs and $\dot{\mathsf{q}}$s, which cannot be assumed to commute. Terms in the Lagrangian that have both qs and $\dot{\mathsf{q}}$s as factors, always assumed to be polynomials, will have to be symmetrically arranged in all possible orders, as a result of which L is Hermitian. For example, a term $\mathsf{q}^2\dot{\mathsf{q}}$ in a polynomial will have to appear as $\frac{1}{3}[\mathsf{q}\dot{\mathsf{q}}\mathsf{q} + \dot{\mathsf{q}}\mathsf{q}^2 + \mathsf{q}^2\dot{\mathsf{q}}]$; similarly for q_i and q_j, etc.

The kinds of changes in the basis or quasi-basis vectors $|X\rangle$ we are concerned with are always such as to preserve their orthonormal character, which means that the new are connected to the old by a canonical transformation: the new set is $|X\rangle_{\text{new}} \stackrel{\text{def}}{=} \mathsf{U}|X\rangle_{\text{old}}$. Such an infinitesimal unitary transformation is necessarily of the form[2] $\mathsf{U} = \mathbb{1} + \frac{i}{\hbar}\mathsf{F}$, so that for small changes $\delta|X\rangle \stackrel{\text{def}}{=} |X\rangle_{\text{new}} - |X\rangle_{\text{old}} = \frac{i}{\hbar}\mathsf{F}|X\rangle$, where the "small operator" F is Hermitian and the factor of $1/\hbar$, which gives F the dimension of an action, is inserted as a matter of convenience. The new operator $\mathbf{X}_{\text{new}}$ that has $|X\rangle_{\text{new}}$ as its eigenvectors or quasi-eigenvectors is then given by $\mathbf{X}_{\text{new}} = \mathsf{U}\mathbf{X}_{\text{old}}\mathsf{U}^\dagger$, which for infinitesimal changes means that

$$\hat{\delta}\mathbf{X} \stackrel{\text{def}}{=} \mathbf{X}_{\text{new}} - \mathbf{X}_{\text{old}} = \frac{i}{\hbar}[\mathsf{F}, \mathbf{X}]. \tag{2.2}$$

You should keep in mind that this change $\hat{\delta}$ in the operator $\mathbf{X}$ is of a very specific kind: the changed operator $\mathbf{X}_{\text{new}}$ is such that $|X\rangle_{\text{new}} = (\mathbb{1} + \frac{i}{\hbar}\mathsf{F})|X\rangle_{\text{old}}$ is an eigenvector or quasi-eigenvector of $\mathbf{X}_{\text{new}}$ with the same eigenvalue X as that of $|X\rangle_{\text{old}}$, with $\mathbf{X}_{\text{old}}|X\rangle_{\text{old}} = X|X\rangle_{\text{old}}$. The operator F is called the *generator* of the unitary transformation.

[2] **Why?**

Now any small change in $\langle 2|1\rangle$ must be due to small changes in the vectors $|1\rangle$ and $|2\rangle$. Hence we must have

$$\delta\langle 2|1\rangle = \frac{i}{\hbar}\langle 2|\mathsf{F}_1 - \mathsf{F}_2|1\rangle = \frac{i}{\hbar}\langle 2|\delta\mathsf{W}_{21}|1\rangle,$$

and therefore, since the choice of $\mathbf{X}_1$ and $\mathbf{X}_2$ is arbitrary,

$$\delta\mathsf{W}_{21} = \mathsf{F}_1 - \mathsf{F}_2, \tag{2.3}$$

which is the *operator principle of stationary action,*[3] the quantum generalization of Hamilton's principle. It says that *in a neighborhood of any physically allowed time development, i.e., one that follows the correct equations of motion, the first variation of the action integral (2.1) depends only on changes at the initial and final times; it does not depend on variations in the intervening history.* If the "endpoints" are kept fixed, the first variation of the action integral vanishes: W is stationary. Just as in classical mechanics, this principle will generate the equations of motion that determine the "physically allowed time development" from a given Lagrangian.

2.1.2 The Lagrangian equations of motion

Let us then use (2.1) and the principle (2.3) to generate the equations of motion; we have

$$\delta\mathsf{W}_{21} = \delta\int_{t_1}^{t_2} dt\,\mathsf{L} = \int_{t_1}^{t_2} dt\,\delta_0\mathsf{L} + \mathsf{L}[t_2]\delta t_2 - \mathsf{L}[t_1]\delta t_1, \tag{2.4}$$

where δ_0 is a variation at fixed time. On the assumption that L depends only on the qs and the $\dot{\mathsf{q}}$s (which, at this point, are all regarded as independent variables), this variation is given by

$$\delta_0\mathsf{L} = \sum_i [\![\frac{\partial\mathsf{L}}{\partial\mathsf{q}_i}\delta_0\mathsf{q}_i + \frac{\partial\mathsf{L}}{\partial\dot{\mathsf{q}}_i}\delta_0\dot{\mathsf{q}}_i]\!].$$

We have to be careful with the order in which the operators appear in this expression: the variations must always appear in the same position as the varied quantities, and they must be symmetrized. This is what the double brackets are meant to indicate. (For example, $\delta_0\mathsf{q}^3 = \delta_0\mathsf{q}\,\mathsf{q}^2 + \mathsf{q}\delta_0\mathsf{q}\,\mathsf{q} + \mathsf{q}^2\delta_0\mathsf{q}$.) Furthermore, in order for the varied $\dot{\mathsf{q}}_i$ near the physically correct "trajectory" to be the velocity corresponding to the varied q_i, the first variation $\delta_0\dot{\mathsf{q}}$ must be equal to $\frac{d}{dt}\delta_0\mathsf{q}$. We therefore have

$$\delta_0\mathsf{L} = \frac{d}{dt}\sum_i [\![\frac{\partial\mathsf{L}}{\partial\dot{\mathsf{q}}_i}\delta_0\mathsf{q}_i]\!] + \sum_i [\![\left[\frac{\partial\mathsf{L}}{\partial\mathsf{q}_i} - \frac{d}{dt}\frac{\partial\mathsf{L}}{\partial\dot{\mathsf{q}}_i}\right]\delta_0\mathsf{q}_i]\!],$$

[3]This method of generating it is due to Julian Schwinger.

and consequently

$$\begin{aligned}\delta \mathsf{W}_{21} &= \left[\sum_i [\![\frac{\partial \mathsf{L}}{\partial \dot{\mathsf{q}}_i} \delta_0 \mathsf{q}_i]\!] + \mathsf{L}[t]\delta t\right]_{t_1}^{t_2} \\ &\quad + \sum_i \int_{t_1}^{t_2} dt\, [\![\left[\frac{\partial \mathsf{L}}{\partial \mathsf{q}_i} - \frac{d}{dt}\frac{\partial \mathsf{L}}{\partial \dot{\mathsf{q}}_i}\right] \delta_0 \mathsf{q}_i]\!] = \mathsf{F}_1 - \mathsf{F}_2\end{aligned} \tag{2.5}$$

according to (2.3). The integral here depends on the history of the system and must therefore vanish in a neighborhood of physically allowed motions for arbitrary functions $\delta_0 \mathsf{q}_i(t)$, restricted only by their commutation relations. If we assume that the $\delta_0 \mathsf{q}_i(t)$ all commute with the operators in L, it follows that the bracket in the integrand of (2.6) must vanish:

$$\frac{\partial \mathsf{L}}{\partial \mathsf{q}_i} = \frac{d}{dt}\frac{\partial \mathsf{L}}{\partial \dot{\mathsf{q}}_i}, \qquad i = 1, 2, \ldots, n. \tag{2.6}$$

These are the Euler-Lagrange equations resulting from the principle of stationary action, the Lagrangian equations of motion, just as in classical mechanics.

2.1.3 The canonical commutation relations

For *actual* physical systems, i.e., those satisfying the equations of motion (2.6), for which, therefore, the last term in (2.6) vanishes, we can conclude that

$$\mathsf{F} = -\sum_i \frac{\partial \mathsf{L}}{\partial \dot{\mathsf{q}}_i}\delta_0 \mathsf{q}_i - \mathsf{L}[t]\delta t = -\sum_i \mathfrak{p}_i \delta_0 \mathfrak{q}_i - \mathsf{L}[t]\delta t, \tag{2.7}$$

where the canonical momenta have been defined by

$$\mathfrak{p}_i \stackrel{\text{def}}{=} \frac{\partial \mathsf{L}}{\partial \dot{\mathfrak{q}}_i}, \qquad i = 1, 2, \ldots, \tag{2.8}$$

and these equations have to be solved for the $\dot{\mathfrak{q}}_i$, thus expressing them in terms the $\mathfrak{p}_j$ and the $\mathfrak{q}_k$. Henceforth, the $\mathfrak{p}_j$ and the $\mathfrak{q}_k$ are to be treated as independent variables. This is what in classical mechanics is called a Legendre transformation; it is designed to lead to the Hamiltonian equations of motion, which are first-order differential equations in the time, whereas the Lagrangian equations are generally of second order.

We begin by keeping the time fixed, setting $\delta t = 0$, so that

$$\mathsf{F} = -\sum_i \mathfrak{p}_i \delta_0 \mathfrak{q}_i, \tag{2.9}$$

and we take $\mathbf{X}$ to be $\mathfrak{q}_j$, so that (2.2) and (2.9) say that

$$\hat{\delta}\mathfrak{q}_j = \frac{i}{\hbar}[\mathfrak{q}_j, \sum_k \mathfrak{p}_k \delta_0 \mathfrak{q}_k].$$

Whatever the precise relation between $\delta_0\mathfrak{q}_j$ and $\hat{\delta}\mathfrak{q}_j$ is going to turn out to be, since the different degrees of freedom are meant to be independent of one another, we may conclude from this that for $k \neq j$, $[\mathfrak{q}_j, \mathfrak{p}_k] = 0$, and we are left with $\hat{\delta}\mathfrak{q}_j = \frac{i}{\hbar}[\mathfrak{q}_j, \mathfrak{p}_j]\delta_0\mathfrak{q}_j$. (Remember we had assumed that the $\delta_0\mathfrak{q}_i$ commute with all the operators in L.) Now it is an easy matter of calculation[4] to find that up to linear terms in the variations,

$$(\mathfrak{q}_j - \delta_0\mathfrak{q}_j)(\mathbb{1} - \frac{i}{\hbar}\mathfrak{p}_j\delta_0\mathfrak{q}_j)|q_j\rangle = (q_j - \hat{\delta}\mathfrak{q}_j - \delta_0\mathfrak{q}_j)(\mathbb{1} - \frac{i}{\hbar}\mathfrak{p}_j\delta_0\mathfrak{q}_j)|q_j\rangle, \quad (2.10)$$

which implies that in order for $|q_j\rangle_{\text{new}} = (\mathbb{1} - \frac{i}{\hbar}\mathfrak{p}_j\delta_0\mathfrak{q}_j)|q_j\rangle$ to be an eigenvector of $\mathfrak{q}_{j\text{new}} \stackrel{\text{def}}{=} \mathfrak{q}_j - \delta_0\mathfrak{q}_j$ with the eigenvalue q_j, we must have $\delta_0\mathfrak{q}_j = -\hat{\delta}\mathfrak{q}_j$. It therefore follows that $[\mathfrak{q}_j, \mathfrak{p}_j] = i\hbar$, or

$$[\mathfrak{q}_j, \mathfrak{p}_k] = i\hbar\delta_{jk}. \quad (2.11)$$

Similarly, taking $\mathbf{X}$ to be $\mathfrak{p}_j$, (2.2) and (2.9) give us $\hat{\delta}\mathfrak{p}_j = -\delta_0\mathfrak{p}_j = \frac{i}{\hbar}[\mathfrak{p}_j, \sum_k \mathfrak{p}_k\delta_0\mathfrak{q}_k]$. However, since the $\mathfrak{q}_j$ and the $\mathfrak{p}_k$ are to be treated as independent variables, so that variations of the $\mathfrak{q}$s do not cause variations of the $\mathfrak{p}$s, it follows that for all j and all k,

$$[\mathfrak{p}_j, \mathfrak{p}_k] = 0. \quad (2.12)$$

At this point, note that if the Lagrangian is changed by a total time derivative,

$$\bar{\mathsf{L}} = \mathsf{L} + \frac{d}{dt}\mathsf{W}, \quad (2.13)$$

the Lagrangian equations and hence the dynamics are unchanged. We then have, by (2.1),

$$\bar{\mathsf{W}}_{21} = \mathsf{W}_{21} + \mathsf{W}_2 - \mathsf{W}_1, \quad (2.14)$$

and by (2.3),

$$\bar{\mathsf{F}} = \mathsf{F} - \delta\mathsf{W}. \quad (2.15)$$

So, in order to find the commutation relations between the $\mathfrak{q}$s, we change the Lagrangian by a total time derivative as in (2.13), choosing $\mathsf{W} = -\sum_j \mathfrak{p}_j\mathfrak{q}_j$. According to (2.15) this gives us a new generating function

$$\bar{\mathsf{F}} = \sum_j \delta_0\mathfrak{p}_j\mathfrak{q}_j,$$

in which we now assume that the $\delta_0\mathfrak{p}_j$ commute with all the operators, which in turn leads to $\hat{\delta}\mathfrak{q}_k = -\delta_0\mathfrak{q}_k = \frac{i}{\hbar}\sum_j \delta_0\mathfrak{p}_j[\mathfrak{q}_j, \mathfrak{q}_k]$, from which we conclude by the same argument as before that we must have for all j and all k,

$$[\mathfrak{q}_j, \mathfrak{q}_k] = 0 \quad (2.16)$$

[4] **Do it.**

because of the assumed independence of the $\mathfrak{p}$s and the $\mathfrak{q}$s.

Equations (2.11), (2.12), and (2.16) are the complete set of canonical commutation relations (often referred to as the *ccr*s), obtained on the assumption that the independent variations $\delta_0\mathfrak{q}_j$ and $\delta_0\mathfrak{p}_i$ commute with all the $\mathfrak{q}$s and $\mathfrak{p}$s. Note that, except for the factor of $i\hbar$ in (2.11), Eqs. (2.11), (2.12), and (2.16) are identical to the corresponding Poisson bracket equations in classical mechanics. Dirac originally employed this analogy to lay the ccrs down as simple postulates.

2.1.4 The Hamiltonian equations of motion

The next step is to make an infinitesimal shift δt in the time, accompanied by a backward shift of the $\mathfrak{q}$s, $\delta_0\mathfrak{q}_i = -\dot{\mathfrak{q}}_i\delta t$ to compensate for the change that $\mathfrak{q}_i$ would ordinarily undergo during δt, thus keeping it fixed. (These $\delta_0\mathfrak{q}_i$ of course do not commute with all the operators in L and have to be moved to the right by means of the ccrs.) According to (2.7), the concomitant generator is given by

$$\mathsf{F} = \mathbf{H}\delta t, \tag{2.17}$$

where the Hermitian operator

$$\mathbf{H} \stackrel{\text{def}}{=} \sum_i [\![\mathfrak{p}_i\dot{\mathfrak{q}}_i]\!] - \mathsf{L} \tag{2.18}$$

is defined to be the Hamiltonian of the system, and it is understood to be expressed as a function of the $\mathfrak{q}$s and the $\mathfrak{p}$s rather than the $\mathfrak{q}$s and the $\dot{\mathfrak{q}}$s.

The first consequence of (2.17) is to yield the Hamiltonian equations of motion of all dynamical variables. Insertion of (2.17) in (2.2) gives

$$\frac{d\mathbf{X}}{dt} = \frac{i}{\hbar}[\mathbf{H}, \mathbf{X}], \tag{2.19}$$

on the assumption that the operator $\mathbf{X}$ does not depend on the time explicitly; if $\mathbf{X}$ has an additional explicit time dependence, (2.19) has to be replaced by

$$\frac{d\mathbf{X}}{dt} = \frac{\partial\mathbf{X}}{\partial t} + \frac{i}{\hbar}[\mathbf{H}, \mathbf{X}], \tag{2.20}$$

called *Heisenberg's equation of motion*. Equation (2.19) is easily solved, giving us, in the Heisenberg picture, the time dependence of all (not intrinsically time-dependent) dynamical variables of a system whose Hamiltonian is $\mathbf{H}$:

$$\mathbf{X}(t) = e^{\frac{i}{\hbar}\mathbf{H}t}\mathbf{X}(0)e^{-\frac{i}{\hbar}\mathbf{H}t}. \tag{2.21}$$

The time-development operator $e^{\frac{i}{\hbar}\mathbf{H}t}$ has to be regarded at this point as no more than a symbolic expression. However, an energy basis of eigenvectors and/or quasi-eigenvectors of $\mathbf{H}$, together with (B.23), may be used to

expand it in the form

$$e^{\frac{i}{\hbar}\mathbf{H}t} = \sum\!\!\!\!\!\!\int_E e^{\frac{i}{\hbar}Et}|E\rangle\langle E|. \tag{2.22}$$

Other mathematical tools for the construction of the unitary time-development operator $e^{\frac{i}{\hbar}\mathbf{H}t}$ will be discussed later.

The equations of motion can now be written in the Hamiltonian form. The commutation relations (2.11), (2.12), and (2.16) imply[5] that for any differentiable function $\mathsf{f}(\mathsf{q}_1, \ldots; \mathsf{p}_1, \ldots)$ of the ps and qs,

$$[\mathsf{p}_j, \mathsf{f}] = -i\hbar\frac{\partial \mathsf{f}}{\partial \mathsf{q}_j}, \qquad [\mathsf{q}_j, \mathsf{f}] = i\hbar\frac{\partial \mathsf{f}}{\partial \mathsf{p}_j}. \tag{2.23}$$

Equation (2.19) therefore leads to the equations

$$\dot{\mathsf{p}}_j = -\frac{\partial \mathbf{H}}{\partial \mathsf{q}_j}, \qquad \dot{\mathsf{q}}_j = \frac{\partial \mathbf{H}}{\partial \mathsf{p}_j}, \qquad j = 1, 2, \ldots, n, \tag{2.24}$$

just as in classical physics.

What is more, equation (2.17) can be used to obtain the conservation laws. First of all, (2.19) tells us that *any quantity that is not explicitly time-dependent and that commutes with the Hamiltonian is independent of the time*: it is conserved. Specifically, of course, the Hamiltonian, whose eigenvalues are the possible energy values of the system, is constant: energy is conserved. Since this result is a consequence of the assumption that the Lagrangian (and hence also the Hamiltonian) does not depend explicitly on the time, it holds for all systems that are *invariant under time translations.* The conservation laws are always consequences of invariance properties of a system. (Recall Noether's theorem in classical mechanics.) We will shortly find more of them for specific kinds of systems.

2.1.5 The Schrödinger picture

The Heisenberg picture had been defined so as to keep the state vector of a developing physical system fixed, while the operators corresponding to dynamical variables change in the course of time, analogous to the classical description. However, this is not the only way of proceeding. Since ultimately all physically observable results are expressed as expectation values or probabilites, which are the squared magnitudes of matrix elements of observables, the quantities whose time-development really matters are all of the form $(\Psi, \mathbf{X}\Phi)$. According to (2.20), such quantities are subject to the equation

$$\frac{d}{dt}(\Psi, \mathbf{X}\Phi) = (\Psi, \frac{\partial \mathbf{X}}{\partial t}\Phi) + (\Psi, \frac{i}{\hbar}[\mathbf{H}, \mathbf{X}]\Phi). \tag{2.25}$$

---

[5] **It will be one of your homework problems to prove this.**

Because $\mathbf{H}$ is Hermitian, this same equation can also be obtained by assuming the state vectors Ψ and Φ to vary with time according to the equation

$$i\hbar\frac{d}{dt}\Psi = \mathbf{H}\Psi, \tag{2.26}$$

while $\mathbf{X}$ depends on the time only through its explicit time dependence; in other words, unless it depends explicitly on the time, the operator $\mathbf{X}$ does not vary.[6] This is called the *Schrödinger picture.* Here the dynamical time dependence is carried by the state vectors, which satisfy the *Schrödinger equation* (2.26), whereas those dynamical variables that do not vary explicitly with t are time-independent. The two pictures are entirely equivalent and lead to the same physical predictions. However, for actual calculations, the Schrödinger picture is used almost exclusively and is of far greater utility. Just as (2.19) was solved symbolically by (2.21), so we can solve (2.26) to express the time development of the state vector of a system whose Hamiltonian is $\mathbf{H}$ in the Schrödinger picture,

$$\Psi(t) = e^{-\frac{i}{\hbar}\mathbf{H}(t-t')}\Psi(t'). \tag{2.27}$$

If the state of a system is described by a density operator rather than a state vector, the Schrödinger equation of motion is obtained from (1.14) and the fact that the states Ψ_n satisfy (2.26). As a result, the equation of motion of the density operator in the Schrödinger picture is given by

$$i\hbar\frac{d}{dt}\boldsymbol{\rho} = [\mathbf{H}, \boldsymbol{\rho}]. \tag{2.28}$$

Since $\mathrm{tr}\mathbf{AB} = \mathrm{tr}\mathbf{BA}$ for any two operators $\mathbf{A}$ and $\mathbf{B}$, we have both $d\mathrm{tr}\boldsymbol{\rho}/dt = 0$ and[7] $d\mathrm{tr}\boldsymbol{\rho}^2/dt = 0$, which implies that not only does the density operator retain its characteristic property $\mathrm{tr}\boldsymbol{\rho} = 1$ in the course of time, but the "degree of coherence" of a system remains unchanged; if it was in a pure state at one time, for example, it will remain that way.

By contrast, we saw in the last chapter that a measurement converts a pure state into a mixed one (or else, instantaneously into an eigenstate of the measured observable—the infamous "collapse of the wave function"), which is sometimes ascribed to the disturbance any measurement, no matter how carefully performed, produces in the measured system. As a result of this disturbance, the probability of finding a system in the state Ψ_B at the time t_2, given that it was in the state Ψ_A at the time $t_1 < t_2$, depends upon whether or not a measurement was performed on it at some intermediate time $t_1 < t_3 < t_2$, even if the outcome of this measurement is ignored, a conclusion that is often depicted as one of the weird aspects

[6] **Verify that (2.26) leads to (2.25).**
[7] **Show this.**

of quantum mechanics. Such an effect, however, exists for any probabilistic theory, including classical statistical mechanics; it is not an intrinsic quantum phenomenon.[8]

The operator symbolically indicated in (2.27), which solves the equation of motion (2.26) for $t > t'$, is called the *retarded propagator*. It is defined by

$$\mathcal{G}^+(t) = \begin{cases} e^{-i\mathbf{H}t/\hbar} & t > 0 \\ 0 & t < 0 \end{cases} . \tag{2.29}$$

This propagator allows us to express the state vector $\Psi(t)$ in terms of the state vector $\Psi(t')$ at any earlier time $t' < t$ by (2.27), or

$$\Psi(t) = \mathcal{G}^+(t - t')\Psi(t'). \tag{2.30}$$

You will easily verify that $\mathcal{G}^+$ satisfies the equation

$$\left(i\hbar\frac{d}{dt} - \mathbf{H}\right)\mathcal{G}^+(t) = \mathbb{1}i\delta(t), \tag{2.31}$$

and furthermore, $\lim_{t\to 0+}\mathcal{G}^+(t) = \mathbb{1}$; as a result, the vector $\Psi(t)$ satisfies the time-dependent Schrödinger equation for $t > t'$ and approaches $\Psi(t')$ as $t \to t'$.[9] Later we shall discuss various ways of constructing this propagator, which plays a central role in quantum dynamics.

The time dependence of the configuration wave function $\psi(\vec{q}_1, \ldots; t)$, which is the inner product of an eigenstate of the $\vec{q}$s and the state Ψ of a system, can be expressed either in the Heisenberg or in the Schrödinger picture. If the Heisenberg operators and states are denoted by $\underline{\vec{q}}$ and $\underline{\Psi}$, respectively, we have

$$\underline{\vec{\mathbf{q}}}(t)\underline{\Psi}_{\vec{q}} = \vec{q}\underline{\Psi}_{\vec{q}}$$

so that

$$\vec{q}\underline{\Psi}_{\vec{q}} = e^{\frac{i}{\hbar}\mathbf{H}t}\underline{\vec{\mathbf{q}}}(0)e^{-\frac{i}{\hbar}\mathbf{H}t}\underline{\Psi}_{\vec{q}}$$

or

$$\underline{\vec{\mathbf{q}}}(0)e^{-\frac{i}{\hbar}\mathbf{H}t}\underline{\Psi}_{\vec{q}} = \vec{q}e^{-\frac{i}{\hbar}\mathbf{H}t}\underline{\Psi}_{\vec{q}}.$$

---

[8] After coarse-graining, the probability $P(G_2|G_1)$ of finding a system that was in grain G_1 at t_1 in grain G_2 at the time t_2 is $P(G_2|G_1) = \mu[G_2 \cap \varphi_{21}(G_1)]/\mu(G_1)$ if $\varphi_{21}(G_1)$ is where the Hamiltonian flow takes G_1 in the time from t_1 to t_2, and $\mu(G)$ is the volume of G in phase space. This formula is based on the assumption that the initial grain G_1 is uniformly filled. The flow $\varphi(G_1)$, on the other hand, will fill any one grain at a later time t only sparsely and generally nonuniformly. Therefore, if a measurement is performed at an intermediate time t_3, the new probability of finding the system in G_2, given the result of that measurement, is in a sense "renormalized," and even if the result of the intermediate measurement is ignored and we sum over all the possible grains the system can visit on the way from G_1 to G_2, the answer will generally differ from that without intermediate measurement. For more on this, see [Newton 00], pp. 139–147.

[9] The limits here have to be specified more precisely as having the meaning that $\| \Psi(t) - \Psi(t') \| \to 0$ and $\| \mathcal{G}^+(t) - \mathbf{1} \| \to 0$.

But $\underline{\vec{\mathbf{q}}}(0)$ and $e^{-\frac{i}{\hbar}\mathbf{H}t}\underline{\Psi}_{\vec{q}}$ are equal to the Schrödinger operators and states $\vec{\mathbf{q}}$ and $\Psi_{\vec{q}}(t)$, respectively. Thus we find

$$\psi(\vec{q},t) = (\underline{\Psi}_{\vec{q}}(t), \underline{\Psi}) = (\Psi_{\vec{q}}, e^{-\frac{i}{\hbar}\mathbf{H}t}\underline{\Psi}) = (\Psi_{\vec{q}}, \Psi(t)),$$

which means either picture may be regarded as underlying the time-dependent configuration wave function. The same argument holds for momentum-space wave functions and any other representation. Therefore, the time dependence of a wave function is governed by the Schrödinger equation in both pictures.

2.2 Systems of Particles

2.2.1 Linear and angular momentum

For a system of point particles, whose locations in configuration space are denoted by $\vec{\mathbf{q}}_i$, we can obtain a second conservation law from (2.2) and (2.9) by subjecting the system to a rigid spatial translation $\delta\vec{\mathbf{q}}_i = \delta\vec{r}$. Such a translation is generated by $\mathsf{F} = \sum_i \vec{\mathbf{p}}_i \cdot \delta\vec{r} = \vec{\mathbf{P}} \cdot \delta\vec{r}$, where $\vec{\mathbf{P}} \stackrel{\text{def}}{=} \sum_i \vec{\mathbf{p}}_i$ is the *total momentum.* If the system is invariant under such a rigid translation, the generator F must commute with the Hamiltonian so that its energy-eigenstates remain unchanged, and it follows that *momentum is conserved:* $\vec{\mathbf{P}}(t_1) = \vec{\mathbf{P}}(t_2)$. Similarly, a third conservation law is obtained by *rotating* the entire system by the infinitesimal angle ε about a given axis $\widehat{n}$ (in a right-handed screw sense), which implies that $\delta\vec{\mathbf{q}}_i = \varepsilon\widehat{n} \times \vec{\mathbf{q}}_i$. This leads to the generator $\mathsf{F} = \varepsilon \sum_i \vec{\mathbf{p}}_i \cdot \widehat{n} \times \vec{\mathbf{q}}_i = \varepsilon\widehat{n} \cdot \sum_i \vec{\mathbf{q}}_i \times \vec{\mathbf{p}}_i = \varepsilon\widehat{n} \cdot \vec{\mathbf{L}}$, where $\vec{\mathbf{L}} \stackrel{\text{def}}{=} \sum_i \vec{\mathbf{L}}_i$ is the total *angular momentum* and $\vec{\mathbf{L}}_i \stackrel{\text{def}}{=} \vec{\mathbf{q}}_i \times \vec{\mathbf{p}}_i$ are the angular momenta of the individual particles. It follows that *if the system is invariant under the rotations $\varepsilon\widehat{n}$, then the corresponding component of the angular momentum is conserved, and if it is invariant under all rotations, all components of the angular momentum are conserved.* Note, however, that the three components of the angular momentum do not commute, so they cannot all be used to label a state. From the commutation relations (2.11), (2.12), and (2.16) we easily find[10] that

$$[\mathbf{L}_1, \mathbf{L}_2] = i\hbar\mathbf{L}_3 \tag{2.32}$$

and its cyclical permutations, which can be written in the vector form

$$\vec{\mathbf{L}} \times \vec{\mathbf{L}} = i\hbar\vec{\mathbf{L}}. \tag{2.33}$$

(We shall return to further discussion of the angular momentum in Chapter 5.)

[10] **Verify it.**

2.2.2 The equations of motion

Let us apply our general results to specific kinds of particle systems, beginning with one consisting of a single point particle of mass M, subject to a conservative force, with the aim of obtaining the equation of motion. As in classical mechanics, the Lagrangian is given by

$$\mathsf{L} = \frac{1}{2}M\dot{\vec{\mathbf{q}}}^2 - V(\vec{\mathbf{q}}),$$

where V is the potential energy. According to (2.8) and (2.18), the momentum is defined as $\vec{\mathbf{p}} = M\dot{\vec{\mathbf{q}}}$ and the Hamiltonian is

$$\mathbf{H} = \frac{\vec{\mathbf{p}}^2}{2M} + V(\vec{\mathbf{q}}). \tag{2.34}$$

Taking expectation values of (2.24), and remembering that in the Heisenberg picture the state vector representing a system is independent of the time, therefore immediately leads to the equations

$$\frac{d}{dt}\langle\vec{\mathbf{p}}\rangle = -\langle\nabla V(\vec{\mathbf{q}})\rangle, \qquad \frac{d}{dt}\langle\vec{\mathbf{q}}\rangle = \langle\frac{\vec{\mathbf{p}}}{M}\rangle, \tag{2.35}$$

which are obviously the quantum analogues of familiar equations of classical mechanics, implying furthermore that

$$M\frac{d^2}{dt^2}\langle\vec{\mathbf{q}}\rangle = -\langle\nabla V(\vec{\mathbf{q}})\rangle. \tag{2.36}$$

Equation (2.36), the quantum version of Newton's equation of motion, is known as *Ehrenfest's theorem.*

If, for a system subject to the Hamiltonian (2.34), Eq.(2.19) is applied to the operator $\vec{\mathbf{q}}\cdot\vec{\mathbf{p}}$ and (2.23) is used, the result is

$$\frac{d}{dt}\vec{\mathbf{q}}\cdot\vec{\mathbf{p}} = -\frac{i}{\hbar}[\vec{\mathbf{q}}\cdot\vec{\mathbf{p}}, \mathbf{H}] = \frac{\vec{\mathbf{p}}^2}{M} - \vec{\mathbf{q}}\cdot\nabla V.$$

In a steady state, i.e., in an eigenstate of the Hamiltonian, however, we have $\langle[\vec{\mathbf{q}}\cdot\vec{\mathbf{p}}, \mathbf{H}]\rangle = 0$. It therefore follows that

$$\langle\frac{\vec{\mathbf{p}}^2}{M} - \vec{\mathbf{q}}\cdot\nabla V\rangle = 0,$$

or, in terms of the kinetic energy $\mathbf{T} = \vec{\mathbf{p}}^2/2M$ of the particle,

$$\langle\mathbf{T}\rangle = \frac{1}{2}\langle\vec{\mathbf{q}}\cdot\nabla V\rangle. \tag{2.37}$$

In classical mechanics, $\frac{1}{2}\vec{q}\cdot\nabla V$ is known as the *virial of Clausius*; thus (2.37) is the quantum version of the *virial theorem.*

In view of (2.34), the Schrödinger equation (2.26) for a single-particle system in the Schrödinger picture, reads,

$$i\hbar\frac{d\Psi}{dt} = \left[\frac{\vec{\mathbf{p}}^2}{2M} + V(\vec{\mathbf{q}})\right]\Psi. \tag{2.38}$$

This equation takes its most useful form in the configuration representation, in which the state of the system is represented by a wave function $\psi(\vec{q},t)$, and, according to (1.28), the momentum operator $\vec{\mathbf{p}}$ by $-i\hbar\nabla_q$, so that the Schrödinger equation becomes the partial differential equation

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{q},t) = \left[-\frac{\hbar^2}{2M}\nabla_q^2 + V(\vec{q})\right]\psi(\vec{q},t). \tag{2.39}$$

Recall that the physical interpretation of $\rho(\vec{q},t) \stackrel{\text{def}}{=} |\psi(\vec{q},t)|^2$ is the probability density of finding the particle at the time t in the vicinity of $\vec{q}$ in configuration space; the Schrödinger equation (2.39) therefore allows us to define a "conservation of probability" in terms of the current density defined by (1.31). Provided the potential function V is *real*, it is a simple matter to verify[11] by means of (2.39) that ρ and j satisfy the *continuity equation*

$$\nabla\cdot j + \frac{\partial}{\partial t}\rho = 0. \tag{2.40}$$

These things are readily generalized to n particles at the positions $\vec{\mathbf{q}}^{(j)}$, for which the Lagrangian is given by

$$\mathsf{L} = \sum_j \frac{1}{2}M_j\dot{\vec{\mathbf{q}}}^{(j)2} - V(\vec{\mathbf{q}}^{(1)},\ldots),$$

the momenta by $\vec{\mathbf{p}}^{(j)} = M_j\dot{\vec{\mathbf{q}}}^{(j)}$, and the Hamiltonian by

$$\mathbf{H} = \sum_j \frac{\vec{\mathbf{p}}^{(j)2}}{2M_j} + V(\vec{\mathbf{q}}^{(1)},\ldots). \tag{2.41}$$

The Schrödinger equation in the configuration representation therefore reads

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{q}^{(1)},\ldots,t) = \left[-\sum_j \frac{\hbar^2}{2M_j}\nabla_{q^{(j)}}^2 + V(\vec{q}^{(1)},\ldots)\right]\psi(\vec{q}^{(1)},\ldots,t). \tag{2.42}$$

[11] **Do it as an exercise.**

2.2.3 The time-independent Schrödinger equation

For most applications it is extremely useful to separate the partial differential equations (2.39) and (2.42) as far as the time is concerned, that is, to make the *ansatz* that the solution ψ of (2.39) can be written as a product $\psi(\vec{q},t) = \psi(\vec{q})f(t)$, and similarly for (2.42). It follows that $f(t)$ has to satisfy the equation[12]

$$i\hbar\frac{df}{dt} = Ef,$$

where E is a constant, and this is easily solved: apart from a constant factor, $f(t) = e^{-iEt/\hbar}$. The function $\psi(\vec{q})$, therefore, must satisfy the **time-independent Schrödinger equation**

$$\left[-\frac{\hbar^2}{2M}\nabla_q^2 + V(\vec{q})\right]\psi_E(\vec{q}) = E\psi_E(\vec{q}) \tag{2.43}$$

in the one-particle case, and

$$\left[-\sum_j \frac{\hbar^2}{2M_j}\nabla_{q^{(j)}}^2 + V(\vec{q}^{(1)},\ldots)\right]\psi_E(\vec{q}^{(1)},\ldots) = E\psi_E(\vec{q}^{(1)},\ldots) \tag{2.44}$$

in the n-particle case. In terms of the Hamiltonian differential operators

$$\mathbf{H} \stackrel{\text{def}}{=} \left[-\frac{\hbar^2}{2M}\nabla_q^2 + V(\vec{q})\right]$$

or

$$\mathbf{H} \stackrel{\text{def}}{=} \left[-\sum_j \frac{\hbar^2}{2M_j}\nabla_{q^{(j)}}^2 + V(\vec{q}^{(1)},\ldots)\right],$$

which are the configuration representations of (2.34) and (2.41), respectively, these equations can be written as

$$\mathbf{H}\psi_E = E\psi_E, \tag{2.45}$$

which shows that ψ_E has to be an eigenstate or quasi-eigenstate of $\mathbf{H}$ with the eigenvalue or quasi-eigenvalue E. The physical significance of the separation constant E is therefore that it is the *energy* of the system.

Since $\mathbf{H}$ is Hermitian, its spectrum is complete. If this spectrum forms a discrete set of points, any arbitrary square-integrable function of $\vec{q}$ can be expanded on the basis of the eigenfunctions of $\mathbf{H}$,

$$\psi(\vec{q}) = \sum_n c_n\psi_{E_n}(\vec{q}),$$

[12]This is based on the assumption that the potential V is independent of the time; otherwise the separation is generally not possible.

so that $\mathbf{H}\psi(\vec{q}) = \sum_n E_n c_n \psi_{E_n}(\vec{q})$, from which we may conclude that the general solution of (2.39) is given by

$$\psi(\vec{q}, t) = \sum_n c_n e^{-iE_n t/\hbar} \psi_{E_n}(\vec{q}), \tag{2.46}$$

and analogously for n-particle systems.

In many instances, however, the spectrum of $\mathbf{H}$ is, at least in part, continuous, with quasi-eigenfunctions $\psi_E(\vec{q})$. In such cases the discrete sum is replaced by an integral, so that any square-integrable function of $\vec{q}$ can be written as a generalized Fourier integral,

$$\psi(\vec{q}) = \int dE\, c(E) \psi_E(\vec{q}),$$

in which the integral ranges over the continuous spectrum. The general solution of (2.39) is then given by

$$\psi(\vec{q}, t) = \int dE\, c(E) e^{-iEt/\hbar} \psi_E(\vec{q}), \tag{2.47}$$

and analogously for n-particle systems. If the spectrum is partly discrete and partly continuous, there will be be both a sum and an integral.

The solutions (2.46) and (2.47) of the Schrödinger equation allow us to solve the initial-value problem of (2.39) or (2.42). If $f(\vec{q}) \stackrel{\text{def}}{=} \psi(\vec{q}, 0)$ is given, we expand it on the basis of the orthonormal set $\{\psi_{E_n}\}$,

$$f(\vec{q}) = \sum_n c_n \psi_{E_n}(\vec{q})$$

where the coefficients c_n are obtained by

$$c_n = \int d^3q\, \psi^*_{E_n}(\vec{q}) f(\vec{q}). \tag{2.48}$$

The solution of (2.39) that has $f(\vec{q})$ as its initial value is then (2.46), with the c_n of (2.48). For a continuous spectrum we assume the quasi-eigenfunctions to be normalized so that

$$\int d^3q\, \psi^*_E(\vec{q}) \psi_{E'}(\vec{q}) = \delta(E - E'),$$

and calculate the function $c(E)$ in (2.47) from $f(\vec{q})$ in a similar manner as in (2.48):

$$c(E) = \int d^3q\, \psi^*_E(\vec{q}) f(\vec{q}). \tag{2.49}$$

Analogous relations hold for n-particle systems and for spectra that are partly discrete and partly continuous.

2.2.4 Example: Gaussian wave packets

Gaussian wave functions form particularly useful examples for which many of the needed mathematical manipulations can be carried out explicitly, with results that often have general qualitative validity. Suppose that the normalized momentum-space wave function (in one dimension) is given by

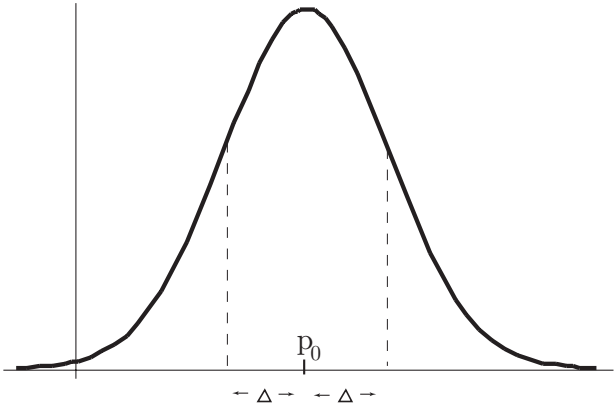

FIGURE 2.1. The probability density of a Gaussian wave function centered at p_0. The dashed lines show $p_0 \pm \Delta$.

$$\widehat{\psi}(p) = \Delta^{-1/2}(2\pi)^{-1/4} e^{-(p-p_0)^2/(4\Delta^2)}. \tag{2.50}$$

The probability density in momentum space described by this function has the bell shape shown in Figure 2.1, centered at $p = p_0$ and decreasing to ~ 0.6 of its peak value at $p = p_0 \pm \Delta$. The variance (dispersion), defined by (1.10), of the momentum in this state is easily calculated[13] to be

$$\Delta p = \Delta. \tag{2.51}$$

The corresponding configuration wave function is found by the one-dimensional version of (1.67) to be

$$\psi(q) = \frac{1}{\sqrt{2\pi\hbar}} \int dp\, e^{ipq/\hbar} \widehat{\psi}(p) = \Delta^{1/2} \hbar^{-1/2} (2/\pi)^{1/4} e^{-q^2\Delta^2/\hbar^2} e^{iqp_0/\hbar}. \tag{2.52}$$

Note that this too is a Gaussian! Furthermore, if the "half-width" of the momentum-space Gaussian is Δ, then that of the configuration-space Gaussian is $\bar{\Delta} = \hbar/(2\Delta)$, and the variance of the particle position is accordingly

$$\Delta q = \hbar/(2\Delta) = \bar{\Delta}, \tag{2.53}$$

[13] **Do it as an exercise.**

so that we find

$$\Delta p \Delta q = \Delta \bar{\Delta} = \frac{1}{2}\hbar. \tag{2.54}$$

Comparison with (1.13) shows the important result that *Gaussian wave packets minimize the uncertainty product.*

There is a simple trick to evaluate the Gaussian integral

$$I = \int_{-\infty}^{\infty} dx\, e^{-cx^2},$$

namely,

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdy\, e^{-c(x^2+y^2)} = 2\pi \int_0^{\infty} rdr\, e^{-cr^2} = \pi \int_0^{\infty} dr^2\, e^{-cr^2} = \frac{\pi}{c};$$

therefore

$$\int_{-\infty}^{\infty} dx\, e^{-cx^2} = \sqrt{\frac{\pi}{c}}. \tag{2.55}$$

Moreover,

$$\int_{-\infty}^{\infty} dx\, x^2 e^{-cx^2} = -\frac{\partial}{\partial c}\int_{-\infty}^{\infty} dx\, e^{-cx^2} = \frac{\sqrt{\pi}}{2c^{3/2}}. \tag{2.56}$$

Integrals of the form

$$J = \int_{-\infty}^{\infty} dx\, e^{-cx^2 - ax},$$

are evaluated by completing the square, so that $cx^2 + ax = c(x + \frac{a}{2c})^2 - \frac{a^2}{4c}$, and shifting the variable of integration, setting $y = x + \frac{a}{2c}$, with the result

$$J = \sqrt{\pi/c}\exp(a^2/4c).$$

Let us next look at what happens to a freely developing Gaussian wave packet in the course of time. Since its time development follows the free Schrödinger equation (2.38) with $V = 0$, so that its Hamiltonian is given by $\mathbf{H} = \mathbf{p}^2/2M$, the momentum-space wave function $\widehat{\psi}(p,t)$ is simply

$$\widehat{\psi}(p,t) = \widehat{\psi}(p,0)e^{-i(p^2/2M\hbar)t},$$

and this has to be inserted in (2.52) in place of $\widehat{\psi}(p)$,

$$\begin{aligned}\psi(q,t) &= \frac{1}{\sqrt{2\pi\hbar}}\int dp\, \widehat{\psi}(p,t)e^{ipq/\hbar} \\ &= \frac{(2\pi)^{-1/4}}{\sqrt{2\pi\hbar\Delta}}\int dp\, e^{-(p-p_0)^2/4\Delta^2 - ip^2 t/(2M\hbar) + ipq/\hbar},\end{aligned} \tag{2.57}$$

an integral that I will let you do as an exercise.[14] The most interesting part of the result is that the factor $e^{-q^2\Delta^2/\hbar^2}$ in (2.52) is replaced by

$$\exp\left[-\frac{\Delta^2}{\hbar^2}\frac{(q - \frac{p_0}{M}t)^2}{1 + \frac{4t^2\Delta^4}{M^2\hbar^2}}\right],$$

[14] **Do it.**

which means that two things happen in the course of time: 1) the center of the packet moves with velocity p_0/M, and 2) while the packet retains the same width in momentum space, its width in configuration space spreads,

$$\bar{\Delta} \to \sqrt{\bar{\Delta}^2 + \left(\frac{t\hbar}{2M\bar{\Delta}}\right)^2}. \tag{2.58}$$

That the center of the packet moves with the velocity p_0/M is, of course, easily physically understood. Its spreading in configuration space—and that the narrower it is, the faster it spreads—can be physically explained by the fact that an uncertainty of $\bar{\Delta}$ in a particle's position implies an uncertainty of at least $\hbar/\bar{\Delta}$ in its momentum; this is the physical origin of the spreading.

Equation (2.57) expresses the configuration wave function as a superposition of plane waves traveling with a phase velocity equal to $p/2M$, because when q increases by $pt/2M$, the phase of the integrand remains constant in time. However, the *center* of the packet, as we have seen, moves with velocity p/M. That these results hold not only for Gaussian wave packets but are of much more general validity can be seen by applying the *stationary-phase argument* to integrals of the form

$$\psi(q,t) = \frac{1}{\sqrt{2\pi\hbar}} \int dp\, f(p) e^{-i\omega t + ipq/\hbar}, \tag{2.59}$$

where $|f(p)|$ is assumed to have a maximum at $p = p_0$. As $|q|$ or $|t|$ tend to ∞, the integrand on the right oscillates rapidly as a function of p and the integral will generally tend to zero,[15] except when q is allowed to grow linearly along with t. In that case, the limit of the integral need not vanish: its value will come predominantly from the neighborhood of the value (or values) of p for which the phase of the integrand is stationary, preventing it from rapidly oscillating; these values of p are determined by the equation

$$\frac{d}{dp}\left[pq/\hbar - \omega(p)t + \varphi\right] = 0,$$

where φ is the phase of the function $f(p)$ in (2.59). If this equation is satisfied at a value of p for which $|f(p)|$ is small, the resulting $|\psi|$ will still be relatively small, whereas $|\psi|$ will be largest when q and t advance together at a rate at which the phase is stationary at the point p_0 at which $|f(p)|$ has its peak, i.e., when $q = t\hbar\frac{d}{dp}\omega|_{p=p_0} - \hbar\frac{d}{dp}\varphi|_{p=p_0}$, resulting in the *group velocity*

$$v_{\text{group}} \stackrel{\text{def}}{=} \hbar d\omega(p)/dp\,|_{p=p_0}\,. \tag{2.60}$$

In this particular case we have $\omega(p) = p^2/(2M\hbar)$ and therefore $v_{\text{group}} = p_0/M$, which is equal to the velocity of the particle at the center of the

[15] This theorem, known as the *Riemann-Lebesgue lemma*, strictly holds if $|f(p)|$ is integrable.

packet: the group velocity of the wave packet is equal to the most probable classical velocity of the particle, a result that we might have expected on the basis of the correspondence principle. At the same time we see that if the phase φ of the weight function $f(p)$ depends of p, the center of the packet is shifted backwards by the amount $\hbar \frac{d}{dp}\varphi|_{p=p_0}$, so that its arrival is retarded by the *time delay*

$$\tau_D \stackrel{\text{def}}{=} \left.\frac{\hbar M}{p_0}\frac{d\varphi}{dp}\right|_{p=p_0}. \tag{2.61}$$

As for the spreading of the packet with time, the stationary-phase argument shows that what matters for large times is only the shape of the function $|f(p)|$ near its peak. Therefore, let us expand $\log|f(p)|$ in a Taylor series around $p = p_0$, keeping only the first term: $\log|f(p)| = -a(p-p_0)^2 + \ldots$, and define $\Delta^2 \stackrel{\text{def}}{=} 1/4a$. Then the integral in (2.59) has the same form as the one in (2.57), and we obtain the same result (2.58) for the spreading of the configuration-space packet.

Equation (2.52) illustrates another important point: to measure the position of a particle with an accuracy Δ does not uniquely determine its state to within a constant factor. The last factor in the expression on the right-hand side of (2.52) can be determined only by a simultaneous measurement of the particle's momentum, even though the latter is uncertain. This is a point emphasized in Chapter 1.

2.3 Fields

2.3.1 The matter field

In contrast to particle systems, where the dynamical variables are observables and hence Hermitian, the field is generally *not* a Hermitian operator, and both Ψ and $\Psi^\dagger$ enter into the Lagrangian in such a way that L is Hermitian. Rather than regarding the Hermitian and skew-Hermitian[16] parts of Ψ as the independent variables, it will be convenient to take the independent variables to be Ψ and $\Psi^\dagger$, which is mathematically equivalent.[17]

In addition to depending on both Ψ and $\Psi^\dagger$, the Lagrangian will now also contain $\nabla\Psi$ and $\nabla\Psi^\dagger$; it will generally be of the form

$$\mathsf{L} = \int d^3x\, \mathcal{L}(\Psi, \Psi^\dagger, \nabla\Psi, \nabla\Psi^\dagger, \dot{\Psi}, \dot{\Psi}^\dagger), \tag{2.62}$$

[16] A skew-Hermitian operator $\mathbf{K}$ is such that $\mathbf{K}^\dagger = -\mathbf{K}$; the Hermitian part of an operator $\mathbf{K}$ is $\frac{1}{2}(\mathbf{K} + \mathbf{K}^\dagger)$, and the skew-Hermitian part is $\frac{1}{2}(\mathbf{K} - \mathbf{K}^\dagger)$.

[17] If $z = x + iy$, then $\frac{\partial}{\partial z} = \frac{\partial}{\partial x} - i\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z^*} = \frac{\partial}{\partial x} + i\frac{\partial}{\partial y}$. Similarly for operators and their Hermitian and skew-Hermitian parts.

where $\mathcal{L}$ is called the *Lagrangian density*, so that

$$\begin{aligned}\delta_0 \mathsf{L} &= \int d^3x \,[\![\frac{\partial \mathcal{L}}{\partial \Psi}\delta_0\Psi + \frac{\partial \mathcal{L}}{\partial \Psi^\dagger}\delta_0\Psi^\dagger + \frac{\partial \mathcal{L}}{\partial \nabla\Psi}\cdot\delta_0\nabla\Psi \\ &\quad + \frac{\partial \mathcal{L}}{\partial \nabla\Psi^\dagger}\cdot\delta_0\nabla\Psi^\dagger + \frac{\partial \mathcal{L}}{\partial \dot\Psi}\delta_0\dot\Psi + \frac{\partial \mathcal{L}}{\partial \dot\Psi^\dagger}\delta_0\dot\Psi^\dagger]\!].\end{aligned}$$

Again, we have to take $\delta_0\dot\Psi = \frac{\partial}{\partial t}\delta_0\Psi$ and $\delta_0\dot\Psi^\dagger = \frac{\partial}{\partial t}\delta_0\Psi^\dagger$ as well as $\delta_0\nabla\Psi = \nabla\delta_0\Psi$ and $\delta_0\nabla\Psi^\dagger = \nabla\delta_0\Psi^\dagger$, and we perform integrations by parts, both in the time-integral in W and in the space-integral in L, assuming that the boundary terms at spatial infinity vanish. In place of (2.4) we then obtain

$$\begin{aligned}\delta \mathsf{W}_{21} &= \int_{t_1}^{t_2} dt \int d^3x \,[\![\left[\frac{\partial \mathcal{L}}{\partial \Psi} - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\Psi} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot\Psi}\right]\delta_0\Psi \\ &\quad + \left[\frac{\partial \mathcal{L}}{\partial \Psi^\dagger} - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\Psi^\dagger} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot\Psi^\dagger}\right]\delta_0\Psi^\dagger]\!] \\ &\quad + \left[\int d^3x [\![\frac{\partial \mathcal{L}}{\partial \dot\Psi}\delta_0\Psi + \frac{\partial \mathcal{L}}{\partial \dot\Psi^\dagger}\delta_0\Psi^\dagger]\!] + \mathsf{L}\delta t\right]_{t_1}^{t_2}. \qquad (2.63)\end{aligned}$$

Assuming that the $\delta_0\Psi$ and $\delta_0\Psi^\dagger$ commute with all the factors in $\mathcal{L}$, we can therefore conclude that the Lagrangian equations for Ψ and $\Psi^\dagger$ read

$$\frac{\partial \mathcal{L}}{\partial \Psi} - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\Psi} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot\Psi} = 0, \qquad (2.64)$$

$$\frac{\partial \mathcal{L}}{\partial \Psi^\dagger} - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\Psi^\dagger} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot\Psi^\dagger} = 0. \qquad (2.65)$$

At this point we add the constraint that $\Psi^\dagger$ is the Hermitian conjugate of Ψ, which makes (2.65) redundant, since it is nothing but the Hermitian conjugate of (2.64). The canonically conjugate field Π and the Hamiltonian are defined analogous to (2.8) and (2.18),

$$\Pi \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}}{\partial \dot\Psi}, \qquad \mathbf{H} \stackrel{\text{def}}{=} \int d^3x \,[\Pi\dot\Psi + \dot\Psi^\dagger\Pi^\dagger] - \mathsf{L}, \qquad (2.66)$$

while the translation generator is obtained from (2.9) and (2.8),

$$\mathsf{F} = -\frac{i\hbar}{2}(\Psi^\dagger\nabla\Psi - \nabla\Psi^\dagger\,\Psi)\cdot\delta\vec r,$$

so that the total momentum

$$\vec{\mathbf{P}} \stackrel{\text{def}}{=} -\frac{i\hbar}{2}\int d^3r\,(\Psi^\dagger\nabla\Psi - \nabla\Psi^\dagger\,\Psi) \qquad (2.67)$$

is conserved if the system is translationally invariant.

To be specific, here is the Lagrangian density actually to be used for the matter field:

$$\mathcal{L} = \frac{i\hbar}{2}\Psi^\dagger\dot{\Psi} - \frac{i\hbar}{2}\dot{\Psi}^\dagger\Psi - \frac{\hbar^2}{2M}\nabla\Psi^\dagger\cdot\nabla\Psi - V\Psi^\dagger\Psi, \tag{2.68}$$

where V is a given function of space. The Lagrangian equation of motion (2.65) that follows from this reads

$$i\hbar\dot{\Psi} = -\frac{\hbar^2}{2M}\nabla^2\Psi + V\Psi, \tag{2.69}$$

which is the Schrödinger equation, applied not to the wave function, as in (2.39), but to the operator function $\Psi(\vec{r},t)$.

The replacement of the wave function ψ in (2.39) by the field operator Ψ is often referred to as "second quantization," because the replacement of the numerical $\vec{q}$s and $\vec{p}$s by the operators $\vec{\mathbf{q}}$ and $\vec{\mathbf{p}}$ is regarded as the first "quantization" step from classical mechanics, and the replacement of the numerical wave functions ψ by the field operators Ψ as the second step. However, though historically correct, this view should be regarded as an anachronism. Because of the argument given in Chapter 1, there is good reason to consider (2.69) as the more basic equation, in that the field Ψ leads to the existence of particles whose wave functions satisfy the Schrödinger equation (2.39) because the field satisfies Eq. (2.69), as follows from (1.54).

According to (2.66), the canonically conjugate field is given by $\Pi \stackrel{\text{def}}{=} \frac{i\hbar}{2}\Psi^\dagger$, and the Hamiltonian is found to be[18]

$$\mathbf{H} = \int d^3x \left[\frac{\hbar^2}{2M}\nabla\Psi^\dagger\cdot\nabla\Psi + V\Psi^\dagger\Psi\right], \tag{2.70}$$

which can also be written, after an integration by parts, assuming that $\Psi(\vec{r})$ vanishes sufficiently rapidly at infinity so that boundary terms do not appear,

$$\mathbf{H} = \int d^3x\,\Psi^\dagger\left[-\frac{\hbar^2}{2M}\nabla^2 + V\right]\Psi. \tag{2.71}$$

As for the canonical commutation relations, we arrive at the analogues of (2.11), (2.12), and (2.16) by analogous arguments.[19] For this purpose it is best to proceed as in Section 1.4.3, perform a Fourier transform as in (1.37) and replace the Fourier integral by a series, so that the functions $\widehat{\Psi}(\vec{k})$ and

[18] **Show it.**

[19] The derivation of (2.11) relied on the hermiticity of the dynamical variables. Therefore the Lagrangian has to be re-expressed in terms of the Hermitian and skew-Hermitian parts of Ψ and $\Psi^\dagger$, as in (2.78) below, and the result of the calculation has to be re-expressed in terms of Ψ and $\Psi^\dagger$.

$\widehat{\Pi}(\vec{k})$ are replaced by sequences $\widehat{\Psi}_j$ and $\widehat{\Pi}_j$, respectively. The results are the commutation relations

$$[\widehat{\Psi}_i, \widehat{\Psi}_j] = 0, \qquad [\widehat{\Psi}_i, \widehat{\Psi}_j^\dagger] = \delta_{ij}, \tag{2.72}$$

or, without the discretization,

$$[\widehat{\Psi}(\vec{k}), \widehat{\Psi}(\vec{k}')] = 0, \qquad [\widehat{\Psi}(\vec{k}), \widehat{\Psi}^\dagger(\vec{k}')] = \delta(\vec{k} - \vec{k}'), \tag{2.73}$$

and

$$[\Psi(\vec{r}), \Psi(\vec{r}')] = 0, \qquad [\Psi(\vec{r}), \Psi^\dagger(\vec{r}')] = \delta(\vec{r} - \vec{r}'). \tag{2.74}$$

The appearance of Dirac delta functions in these relations shows that the fields cannot really be pointwise defined functions with values that are operators on the Hilbert space $\mathfrak{H}$, as we will nevertheless continue to pretend as long as it causes no mathematical trouble.

It will be useful to expand the operator function $\Psi(\vec{r}, t)$ that solves (2.69) on the basis of the complete set of eigenfunctions of the Schrödinger equation (2.43),

$$\left[-\frac{\hbar^2}{2M}\nabla^2 + V\right] f_E(\vec{r}) = E f_E(\vec{r}),$$

with the normalization

$$\int d^3r\, f_E^*(\vec{r}) f_{E'}(\vec{r}) = \delta_{EE'},$$

or, for continuous quasi-eigenvalues, with the Dirac $\delta(E - E')$ on the right. The expansion

$$\Psi(\vec{r}) = \sum\!\!\!\!\!\!\int_E f_E(\vec{r}) \Psi_E, \tag{2.75}$$

which implies, by (2.74), that

$$[\Psi_E, \Psi_{E'}] = 0, \qquad [\Psi_E, \Psi_{E'}^\dagger] = \delta_{EE'}, \tag{2.76}$$

(or with a Dirac $\delta(E - E')$ for continuous E) allows us to transform (2.71) into the form

$$\mathbf{H} = \sum\!\!\!\!\!\!\int_E \mathbf{H}_E, \qquad \mathbf{H}_E \stackrel{\text{def}}{=} E \Psi_E^\dagger \Psi_E = E\, \mathsf{N}_E, \tag{2.77}$$

where $\mathsf{N}_E \stackrel{\text{def}}{=} \Psi_E^\dagger \Psi_E$, as in (1.41). As we already saw in (1.46), the operators N_E all commute with one another. Therefore there is not only a complete set of common eigenvectors, but all the N_E commute with the Hamiltonian $\mathbf{H}$, which means they are constants of the motion: their eigenvalues are conserved.

Remarkably enough, the Hamiltonian $\mathbf{H}_E$ is, apart from a shift by a constant, identical to that of a *simple harmonic oscillator.*

Field quanta and the harmonic oscillator. Split Ψ_E into its Hermitian and skew-Hermitian parts, which we will call, by definition, $\sqrt{\frac{\omega}{2\hbar}}\mathsf{q}$ and $\sqrt{\frac{1}{2\hbar\omega}}i\mathsf{p}$:

$$\Psi_E \stackrel{\text{def}}{=} \sqrt{\frac{\omega}{2\hbar}}\left[\mathsf{q} + i\frac{\mathsf{p}}{\omega}\right], \tag{2.78}$$

where q and p are Hermitian and $\omega \stackrel{\text{def}}{=} E/\hbar$. We then find from (2.76) that $[\mathsf{q}, \mathsf{p}] = i\hbar$ and from (2.77),

$$\mathbf{H}_E = \frac{1}{2}(\mathsf{p}^2 + \omega^2\mathsf{q}^2) - \frac{1}{2}\hbar\omega, \tag{2.79}$$

which, apart from a shift by the constant $-\frac{1}{2}\hbar\omega$, is the Hamiltonian of a simple harmonic oscillator of unit mass and frequency ω. Together with the results of our discussion in Section 1.4.3 we have therefore learned that *the allowed energy levels of a simple harmonic oscillator of frequency ω are $E_n = (n + 1/2)\hbar\omega$, with $n = 0, 1, \cdots$*, as follows from (2.77) and the fact that the eigenvalues of N are the non-negative integers: the energy levels are equally spaced, each a distance $\hbar\omega$ apart from the next, and the ground state has the energy $\frac{1}{2}\hbar\omega$.

Because the energy-eigenvalues of a simple harmonic oscillator are equally spaced in steps of $\hbar\omega$, the field gives rise to quanta, each of energy $\hbar\omega$. If it were not for that equal spacing, the infinitely many eigenvalues of the matter-field Hamiltonian *for each fixed value of E* in (2.77) could not be interpreted as simply adding another particle of that same energy E. [This has nothing to do with the spacing between the eigenvalues E of (2.43). We are adding here particles all with the same energies E.] It follows directly from (1.54) and (2.69) that the configuration-space wave function of these particles satisfies the Schrödinger equation, and similarly for multiparticle systems.

Anti-commutation relations

The results we have obtained are based on the assumption that the variations $\delta_0\Psi$ commute with all the operators in the Lagrangian. If, instead, these variations are assumed to anti-commute with the other operators, one readily finds all the commutators in (2.72), (2.73), and (2.74) have to be replaced by anti-commutators:

$$\{\widehat{\Psi}_i, \widehat{\Psi}_j\} = 0, \qquad \{\widehat{\Psi}_i, \widehat{\Psi}_j^\dagger\} = \delta_{ij}, \tag{2.80}$$

or without the discretization,

$$\{\widehat{\Psi}(\vec{k}, t), \widehat{\Psi}(\vec{k}', t)\} = 0, \qquad \{\widehat{\Psi}(\vec{k}, t), \widehat{\Psi}^\dagger(\vec{k}', t)\} = \delta(\vec{k} - \vec{k}'), \tag{2.81}$$

and

$$\{\Psi(\vec{r}, t), \Psi(\vec{r}', t)\} = 0, \qquad \{\Psi(\vec{r}, t), \Psi^\dagger(\vec{r}', t)\} = \delta(\vec{r} - \vec{r}'). \tag{2.82}$$

As we have already seen in Chapter 1, these anti-commutation relations give rise to *anti-symmetric* multiparticle states obeying the Pauli exclusion principle.[20]

2.3.2 Infinitely many particles as a field

A collection of infintely many elastically coupled particles making up a solid can be treated as equivalent to a quantum field, so that their collective harmonic vibrations give rise to quanta called *phonons.* Let us consider a very simple one-dimensional model, in which the particles are all of equal mass and are coupled to one another with the same elastic constant (Fig.2.2). The Lagrangian of such a system is given by

$$\mathsf{L} = \sum_{n=-\infty}^{\infty} [\frac{1}{2} M \dot{\mathbf{q}}_n^2 - \frac{1}{2}\kappa(\mathbf{q}_{n+1} - \mathbf{q}_n)^2] \tag{2.83}$$

if the position operators of the particles in their configuration space are

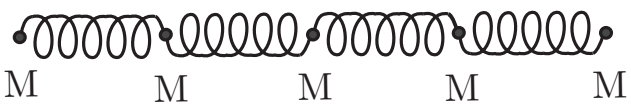

FIGURE 2.2. Equal masses coupled by equal oscillators.

denoted by $\mathbf{q}_n$. The introduction of the function

$$\Psi(k) \stackrel{\text{def}}{=} \sum_{-\infty}^{\infty} \mathbf{q}_n e^{i\pi n k}, \tag{2.84}$$

which implies that

$$\mathbf{q}_n = \frac{1}{2}\int_{-1}^{1} dk\, \Psi(k) e^{-i\pi n k} \tag{2.85}$$

and therefore

$$\mathbf{q}_{n+1} - \mathbf{q}_n = -i\int_{-1}^{1} dk\, \Psi(k) \sin(\frac{1}{2}\pi k) e^{-i\pi(n+\frac{1}{2})k},$$

[20]Note, however, that the spin-statistics connection, to be discussed again later, requires that scalar and vector fields, giving rise to spin-0 and spin-1 particles, respectively, obey commutation relations, whereas spinor fields, which give rise to spin-1/2 particles, obey anti-commutation relations. Therefore, the anti-commutation relations (2.80), (2.81), and (2.82), though written for a scalar field, are possible only if the field Ψ is a spinor field, which has two components, so that the right-hand sides of the second equations in (2.80), (2.81), and (2.82) should be multiplied by the unit matrix in spin space.

leads to[21]

$$\mathsf{L} = \int_{-1}^{1} dk\, \mathcal{L}(k) = \int_{-1}^{1} dk\, [\frac{1}{4} M \dot{\Psi}^\dagger(k) \dot{\Psi}(k) - \kappa \sin^2(\frac{1}{2}\pi k) \Psi^\dagger(k) \Psi(k)]. \tag{2.86}$$

The Lagrangian equation of motion obtained from this reads

$$\frac{1}{4} M \ddot{\Psi}(k) = -\kappa \sin^2(\frac{1}{2}\pi k) \Psi(k),$$

the solution of which is

$$\Psi(k,t) = e^{-i\omega_k t} \Psi(k,0), \tag{2.87}$$

where

$$\omega_k = 2\sqrt{\kappa/M}\, \sin(\frac{1}{2}\pi |k|). \tag{2.88}$$

(We shall see that the choice of the sign in the exponential in (2.87) is needed for consistency.[22]) The conjugate momentum operator turns out to be

$$\Pi(k) \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}(k)}{\partial \dot{\Psi}(k)} = \frac{1}{4} M \dot{\Psi}^\dagger(k) = \frac{i}{4} M \omega_k \Psi^\dagger(k),$$

so that we obtain the commutation relations

$$i\hbar \delta(k - k') = [\Psi(k), \Pi(k')] = \frac{i}{4} M \omega_k [\Psi(k), \Psi^\dagger(k')],$$

or

$$[\Psi(k), \Psi^\dagger(k')] = \frac{4\hbar}{M\omega_k} \delta(k - k'). \tag{2.89}$$

The number operator for the k-quanta of the field is therefore

$$\mathsf{N}(k) \stackrel{\text{def}}{=} \frac{M\omega_k}{4\hbar} \Psi^\dagger(k) \Psi(k)$$

and the field Hamiltonian is given by

$$\mathbf{H} = \int_{-1}^{1} dk\, \kappa \sin^2(\frac{1}{2}\pi k) \Psi^\dagger(k) \Psi(k) = \int_{-1}^{1} dk\, \hbar\omega_k \mathsf{N}(k), \tag{2.90}$$

which leads to the Hamiltonian equation of motion[23]

$$i\hbar \dot{\Psi}(k) = [\Psi(k), \mathbf{H}] = \hbar\omega_k \Psi(k),$$

whose solution is (2.87).

The string of harmonically coupled point masses described by the Lagrangian (2.83) therefore, for each wave number k, quantum-mechanically produces quanta whose energy is given by $\hbar\omega_k = 2\hbar\sqrt{\kappa/M}\, \sin(\frac{1}{2}\pi|k|)$ as a function of k. These quanta are called *phonons.*

---

[21] **Show this.**

[22] **Try the other sign and see what happens.**

[23] **Check this.**

2.3.3 The electromagnetic field

The Lagrangian densitity $\mathcal{L}$ for the free electromagnetic field is taken to be

$$\mathcal{L} = \frac{1}{8\pi}(\vec{\mathbf{E}}^2 - \vec{\mathbf{B}}^2), \tag{2.91}$$

in terms of the electric field $\vec{\mathbf{E}}$ and the magnetic field $\vec{\mathbf{B}}$, which in turn are expressed in terms of the scalar and vector potentials by

$$\vec{\mathbf{E}} = -\frac{1}{c}\frac{\partial \vec{\mathbf{A}}}{\partial t} - \nabla\phi, \qquad \vec{\mathbf{B}} = \nabla \times \vec{\mathbf{A}}. \tag{2.92}$$

(Here c is the speed of light.) These potentials are regarded as the dynamical variables of the field, and $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ in (2.91) are to be expressed in terms of them. The Lagrangian equations of motion obtained from (2.91) via (2.64) are then found to be the Maxwell equations of the free electromagnetic field[24]

$$\nabla \cdot \vec{\mathbf{E}} = 0, \qquad \nabla \times \vec{\mathbf{B}} = \frac{1}{c}\frac{\partial \vec{\mathbf{E}}}{\partial t},$$

while the other two Maxwell equations

$$\nabla \cdot \vec{\mathbf{B}} = 0, \qquad \nabla \times \vec{\mathbf{E}} = -\frac{1}{c}\frac{\partial \vec{\mathbf{B}}}{\partial t},$$

are identities that follow directly from (2.92).

The canonically conjugate momenta are defined by (2.66). Since the Lagrangian is independent of $\dot{\phi}$, there is no momentum conjugate to ϕ, but the canonical momentum conjugate to the vector potential $\vec{\mathbf{A}}$ is the vector

$$\vec{\Pi} \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}}{\partial \dot{\vec{\mathbf{A}}}} = -\frac{1}{4\pi c}\vec{\mathbf{E}},$$

which leads to the Hamiltonian[25]

$$\mathbf{H}_{\text{elmag}} = \int d^3r \left[2\pi c^2 \vec{\Pi}^2 + \frac{1}{8\pi}\vec{\mathbf{B}}^2 + c\phi\nabla \cdot \vec{\Pi}\right],$$

where $\vec{\mathbf{B}}$ has to be expressed in terms of the vector potential by (2.92). Use of the Maxwell equations and the definition of $\vec{\Pi}$ makes the last term vanish, with the result that the energy operator, that is, the conserved Hamiltonian, is given in terms of the energy density $\mathsf{U}_{\text{elmag}}$ by

$$\mathbf{H}_{\text{elmag}} = \int d^3r\, \mathsf{U}_{\text{elmag}}, \qquad \mathsf{U}_{\text{elmag}} \stackrel{\text{def}}{=} \frac{1}{8\pi}(\vec{\mathbf{E}}^2 + \vec{\mathbf{B}}^2). \tag{2.93}$$

---

[24] **Show this as an exercise**. It requires some integrations by parts, in which the boundary terms at infinity are assumed to vanish.

[25] **Show this as an exercise, and also verify that the Hamiltonian equations of motion are, again, the source-free Maxwell equations.**

As in the case of the matter field, it is best to perform a spatial Fourier transform before turning to the canonical commutation relations. Thus we define

$$\vec{\mathbf{A}}(\vec{r},t)=\frac{1}{(2\pi)^{3/2}}\int d^3k\, e^{i\vec{k}\cdot\vec{r}}\widehat{\vec{\mathbf{A}}}(\vec{k},t),\quad \phi(\vec{r},t)=\frac{1}{(2\pi)^{3/2}}\int d^3k\, e^{i\vec{k}\cdot\vec{r}}\widehat{\phi}(\vec{k},t) \tag{2.94}$$

and express the Lagrangian directly in terms of the Fourier-transformed fields,

$$\mathsf{L}=\int d^3k\,\frac{1}{8\pi}(\widehat{\vec{\mathbf{E}}}\cdot\widehat{\vec{\mathbf{E}}}^{\dagger}-\widehat{\vec{\mathbf{B}}}\cdot\widehat{\vec{\mathbf{B}}}^{\dagger}), \tag{2.95}$$

where $\widehat{\vec{\mathbf{E}}}=-\frac{1}{c}\dot{\widehat{\vec{\mathbf{A}}}}-i\vec{k}\widehat{\phi}$ and $\widehat{\vec{\mathbf{B}}}=i\vec{k}\times\widehat{\vec{\mathbf{A}}}$. This leads to the Fourier-transformed Maxwell equations

$$\vec{k}\cdot\widehat{\vec{\mathbf{E}}}=0,\quad -\frac{1}{c}\dot{\widehat{\vec{\mathbf{E}}}}=i\vec{k}\times\widehat{\vec{\mathbf{B}}},\quad \vec{k}\cdot\widehat{\vec{\mathbf{B}}}=0,\quad \frac{1}{c}\dot{\widehat{\vec{\mathbf{B}}}}=i\vec{k}\times\widehat{\vec{\mathbf{E}}},$$

and the momentum canonically conjugate to $\widehat{\vec{\mathbf{A}}}$

$$\widehat{\vec{\Pi}}=-\frac{1}{4\pi c}\widehat{\vec{\mathbf{E}}}^{\dagger}.$$

The canonical commutation relations therefore are

$$[\widehat{\vec{\mathbf{A}}}_j(\vec{k}),\widehat{\vec{\Pi}}_l(\vec{k}')]=i\hbar\delta_{jl}\delta^3(\vec{k}-\vec{k}'),$$

or

$$[\widehat{\vec{\mathbf{A}}}_j(\vec{k}),\widehat{\vec{\mathbf{E}}}_l^{\dagger}(\vec{k}')]=-4\pi ic\hbar\delta_{jl}\delta^3(\vec{k}-\vec{k}'); \tag{2.96}$$

all other commutators vanish.

The most suitable gauge to adopt is the *radiation gauge,* also called the *Coulomb gauge,* in which[26] $\nabla\cdot\vec{\mathbf{A}}=0$ and the scalar potential, which depends on distant charges instantaneously, is set equal to zero for a pure radiation field. In that case, the Fourier transformed Maxwell equations for the vector potential simply read $\vec{k}\cdot\widehat{\vec{\mathbf{A}}}(\vec{k},t)=0$, so that $\widehat{\vec{\mathbf{A}}}(\vec{k},t)$ is transverse, i.e., orthogonal to $\vec{k}$, and satisfies the equation of motion

$$\frac{\partial^2}{\partial t^2}\widehat{\vec{\mathbf{A}}}+\omega^2\widehat{\vec{\mathbf{A}}}=0 \tag{2.97}$$

with $\omega\stackrel{\text{def}}{=}kc$.

[26]There are certain mathematical difficulties associated with this, which we shall ignore.

The fact that the field operators $\vec{\mathbf{A}}$, $\vec{\mathbf{E}}$, and $\vec{\mathbf{B}}$ are Hermitian has the immediate consequence that their Fourier transforms are subject to the constraint

$$\widehat{\vec{\mathbf{A}}}(-\vec{k},t) = \widehat{\vec{\mathbf{A}}}^{\,\dagger}(\vec{k},t), \tag{2.98}$$

and similarly for the electric and magnetic fields. In order to take this restriction into account it is most convenient to define the operators $\vec{\mathsf{b}}(\vec{k},t)$ by the equations

$$\widehat{\vec{\mathbf{E}}}(\vec{k},t) = \sqrt{2\pi\hbar\omega}[\vec{\mathsf{b}}(\vec{k},t) + \vec{\mathsf{b}}^{\dagger}(-\vec{k},t)], \tag{2.99}$$

$$\dot{\widehat{\vec{\mathbf{E}}}}(\vec{k},t) = -i\omega\sqrt{2\pi\hbar\omega}[\vec{\mathsf{b}}(\vec{k},t) - \vec{\mathsf{b}}^{\dagger}(-\vec{k},t)]. \tag{2.100}$$

In terms of the vector potential, this means that[27]

$$\vec{\mathsf{b}}(\vec{k},t) = \frac{\pi}{\sqrt{\hbar\omega}}[ik\widehat{\vec{\mathbf{A}}}(\vec{k},t) - \frac{1}{c}\dot{\widehat{\vec{\mathbf{A}}}}(\vec{k},t)]; \tag{2.101}$$

they are transverse, $\vec{k}\cdot\vec{\mathsf{b}}(\vec{k},t) = 0$, and they satisfy the simple equation of motion $\dot{\vec{\mathsf{b}}} = -i\omega\vec{\mathsf{b}}$, implying that they can be expressed in terms of time-indendent operators $\vec{\mathsf{a}}(\vec{k})$ by

$$\vec{\mathsf{b}}(\vec{k},t) = e^{-i\omega t}\vec{\mathsf{a}}(\vec{k}).$$

(Remember that $\omega = kc > 0$, so that the operators $\vec{\mathsf{b}}$ have positive frequency.) Since $\dot{\widehat{\vec{\mathbf{E}}}} = -\frac{1}{c}\ddot{\widehat{\vec{\mathbf{A}}}} = k^2 c\widehat{\vec{\mathbf{A}}}$, the vector potential is given in terms of $\vec{\mathsf{a}}$ by

$$\widehat{\vec{\mathbf{A}}}(\vec{k},t) = -ic\sqrt{\frac{2\pi\hbar}{\omega}}[\vec{\mathsf{a}}(\vec{k})e^{-i\omega t} - \vec{\mathsf{a}}^{\dagger}(-\vec{k})e^{i\omega t}]. \tag{2.102}$$

The canonical commutation relations for the operators $\vec{\mathsf{a}}$ are obtained by inserting (2.102) and (2.99) in (2.96), with the result

$$[\mathsf{a}_j(\vec{k}), \mathsf{a}_l^{\dagger}(\vec{k}')] = \delta_{jl}\delta^3(\vec{k}-\vec{k}'), \tag{2.103}$$

while all the other commutators vanish. Thus, again, we arrive at the commutation relations of simple harmonic oscillators, with the result that for each given $\vec{k}$ (best discretized for this purpose), and for each of the two plane polarizations perpendicular to $\vec{k}$ (remember that $\vec{k}\cdot\vec{\mathsf{a}}(\vec{k}) = 0$), there are infinitely many states forming a ladder with equidistant steps, $\vec{\mathsf{a}}^{\dagger}$ acting as the raising operator and $\vec{\mathsf{a}}$ as the lowering operator, and the ground state $|0\rangle$ defined by $\vec{\mathsf{a}}|0\rangle = 0$. The number operator, which counts the number of quanta in each state of a given $\vec{k}$ and plane polarization $\widehat{e}^{(j)} \perp \vec{k}$, is

$$\mathsf{N}_j(\vec{k}) = \mathsf{a}_j^{\dagger}(\vec{k})\mathsf{a}_j(\vec{k}). \tag{2.104}$$

[27] **Show this as an exercise**.

There remains the task of calculating the energy of this collection of oscillators.[28] Using (2.93), we find that[29]

$$\begin{aligned}\mathbf{H}_{\text{elmag}} &= \frac{1}{8\pi}\int d^3r\,(\vec{\mathbf{E}}^2+\mathbf{B}^2) = \frac{1}{8\pi}\int d^3k\,[\vec{\mathbf{E}}(\vec{k})\cdot\vec{\mathbf{E}}(-\vec{k})+\mathbf{B}(\vec{k})\cdot\mathbf{B}(-\vec{k})]\\ &= \int d^3k\,\hbar\omega\sum_j \mathsf{a}_j^\dagger(\vec{k})\mathsf{a}_j(\vec{k}) = \int d^3k\,\hbar\omega\sum_j \mathsf{N}_j(\vec{k}).\end{aligned}$$

Thus, the energy of a state $|n_{\vec{k},j}\rangle$ of n quanta of wave vector $\vec{k}$ and polarization j is

$$E = \langle n_{\vec{k},j}|\mathbf{H}_{\text{elmag}}|n_{\vec{k},j}\rangle = n_{\vec{k},j}\hbar\omega,$$

as befits a state of n photons, each of energy $\hbar\omega$. Similarly we can calculate the electromagnetic momentum $\vec{\mathbf{P}}$ of such a state by using either[30]

$$\vec{\mathbf{P}} = \frac{1}{4\pi c}\int d^3r\,\vec{\mathbf{E}}\times\vec{\mathbf{B}}, \tag{2.105}$$

or else (2.67), with the result[31]

$$\langle n_{\vec{k},j}|\vec{\mathbf{P}}|n_{\vec{k},j}\rangle = n_{\vec{k},j}\hbar\vec{k}.$$

Finally, we can find the angular momentum, using the operator form of the angular momentum of the electromagnetic field,[32]

$$\vec{\mathbf{M}} = \frac{1}{4\pi c}\int d^3r\,\vec{r}\times(\vec{\mathbf{E}}\times\vec{\mathbf{B}}), \tag{2.106}$$

and we obtain[33] for its l component

$$\langle n_{\vec{k},j}|\mathbf{M}_l|n_{\vec{k},j}\rangle = \left\langle n_{\vec{k},j}|\mathsf{a}_j^\dagger(\vec{k})[(i\nabla_{\vec{k}}\times\hbar\vec{k})_l+\mathsf{S}_l]\mathsf{a}_j(\vec{k})|n_{\vec{k},j}\right\rangle,$$

where the *spin operator* S is defined by its action on any vector $\vec{f}$,

$$\mathsf{S}_j\vec{f}\stackrel{\text{def}}{=} i\hbar\hat{e}^{(j)}\times\vec{f} \tag{2.107}$$

and the operator $\vec{\mathbf{L}}\stackrel{\text{def}}{=} i\nabla_{\vec{k}}\times\hbar\vec{k}$ is interpreted as an analogue of the orbital angular momentum.[34] Since[35] $\sum_j \mathsf{S}_j^2 = 2\hbar^2$, the intrinsic spin of the photon equals $1\hbar$, as $s(s+1)=2$ means $s=1$.(See Chapter 5 for further details.)

---

[28] From the physical point of view it is again preferable here to discretize the wave vectors and replace the Fourier integrals by Fourier series.

[29] **Verify this.**

[30] See [Jackson], p. 238.

[31] **Verify this.**

[32] See [Jackson], p. 333.

[33] **Show it.**

[34] **Verify that S satisfies the correct commutation relations (2.33).**

[35] **Verify this.**

The radiation field being transverse, $\vec{k}\cdot\vec{\mathsf{a}}(\vec{k})=0$, for a given direction of $\vec{k}$ there are only two linearly independent $\mathsf{a}_j(\vec{k})$ available. Supposing for the moment that a coordinate system is chosen in which $\vec{k}$ points in the positive z direction, we may call the two $\mathsf{a}_1 \stackrel{\text{def}}{=} \widehat{e}^{(1)}\cdot\vec{\mathsf{a}}$ and $\mathsf{a}_2 \stackrel{\text{def}}{=} \widehat{e}^{(2)}\cdot\vec{\mathsf{a}}$, corresponding to plane polarization in the x and y directions. Alternatively, however, one may choose the two opposite *circular polarizations* as fundamental and correspondingly adopt the complex unit vectors

$$\widehat{e}^{(\pm)} \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}(\widehat{e}^{(2)} \mp i\widehat{e}^{(1)}),$$

which are eigenvectors of $\widehat{k}\cdot\mathsf{S}=\mathsf{S}_3$ with the eigenvalues $\pm 1\hbar$:[36]

$$\mathsf{S}_3\widehat{e}^{\pm} = \pm\hbar\widehat{e}^{\pm}.$$

These are called *helicity* eigenstates, and they lead to the operators

$$\mathsf{a}_\pm \stackrel{\text{def}}{=} \widehat{e}^{\pm}\cdot\mathsf{a} = \frac{1}{\sqrt{2}}(\mathsf{a}_2 \mp i\mathsf{a}_1) \tag{2.108}$$

and the commutation relations[37]

$$[\mathsf{a}_+(\vec{k}),\mathsf{a}_+^\dagger(\vec{k}')] = [\mathsf{a}_-(\vec{k}),\mathsf{a}_-^\dagger(\vec{k}')] = \delta^3(\vec{k}-\vec{k}'), \tag{2.109}$$

while all the other commutators vanish. The number operators for the two helicity states are then defined by

$$\mathsf{N}_\pm(\vec{k}) = \mathsf{a}_\pm^\dagger(\vec{k})\mathsf{a}_\pm(\vec{k}). \tag{2.110}$$

The advantage of using helicity, or circular polarization, rather than plane polarization for the labeling of the photon states is that, whereas the operators $\mathsf{a}_j(\vec{k})$ and $\mathsf{N}_j(\vec{k})$, $j=1,2$, depend on the choice of a coordinate system, the number operators $\mathsf{N}_\pm(\vec{k})$ do not. This is seen by performing a rotation of the reference frame around $\vec{k}$. You then find that if the rotation is by an angle ξ, the new operators $\mathsf{a}'_\pm(\vec{k})$ are given by[38]

$$\mathsf{a}'_\pm(\vec{k}) = e^{\pm i\xi}\mathsf{a}_\pm(\vec{k}),$$

and as a result, the $\mathsf{N}_\pm(\vec{k})$ remain unchanged, while their eigenvectors $|n_{\vec{k}\pm}\rangle$ undergo a phase change by $\pm\xi$.[39]

---

[36] **Show this.**

[37] **Check these.**

[38] **Show this.**

[39] Since the phase of $\mathsf{b}_\pm(\vec{k},t)$ varies in time as $-i\omega t$, the circular polarization can be regarded as a continual rotation of the plane of polarization about $\vec{k}$.

So here is a summary of our results: $\mathsf{a}^\dagger_\pm(\vec{k})$ is the creation operator of a photon of energy $\hbar\omega = \hbar ck > 0$, momentum $\hbar\vec{k}$, and helicity $\pm\hbar$, while $\mathsf{N}_\pm(\vec{k})$ is the corresponding number operator. A normalized state of $n_{\vec{k}\pm}$ photons of momentum $\hbar\vec{k}$ and helicity $\pm\hbar$ is, according to (1.41), given in terms of the normalized vacuum state $|0\rangle$ by

$$|n_{\vec{k}\pm}\rangle = \frac{1}{\sqrt{n_{\vec{k}\pm}!}}[\mathsf{a}^\dagger_\pm(\vec{k})]^{n_{\vec{k}\pm}}|0\rangle. \tag{2.111}$$

2.4 Canonical and Gauge Transformations

A canonical transformation is a unitary transformation of all the dynamical variables of a theory by the same operator U, so that, if the original operator was $\mathbf{D}$, the new one is given by

$$\mathbf{D}' = \mathsf{U}\mathbf{D}\mathsf{U}^{-1}. \tag{2.112}$$

Such a transformation leaves all the canonical commutation relations unchanged, since

$$\begin{aligned}[\mathbf{D}'_1, \mathbf{D}'_2] &= \mathsf{U}\mathbf{D}_1\mathsf{U}^{-1}\mathsf{U}\mathbf{D}_2\mathsf{U}^{-1} - \mathsf{U}\mathbf{D}_2\mathsf{U}^{-1}\mathsf{U}\mathbf{D}_1\mathsf{U}^{-1} \\ &= \mathsf{U}[\mathbf{D}_1, \mathbf{D}_2]\mathsf{U}^{-1} = [\mathbf{D}_1, \mathbf{D}_2]\end{aligned}$$

if $[\mathbf{D}_1, \mathbf{D}_2]$ is simply a number (i.e., a multiple of the unit operator). Thus the commutation relations between the $\mathbf{p}'$s and $\mathbf{q}'$s are the same as those between the $\mathbf{p}$s and $\mathbf{q}$s. Furthermore, since U is unitary, $\mathbf{D}'$ is Hermitian whenever $\mathbf{D}$ is, and the spectra—the possible results of measurements—of all observables are unchanged. Since the time development of all dynamical variables in the Heisenberg picture is of the form (2.21), these variables can be viewed as undergoing, in the course of time, a continuous group of canonical transformations with $\mathsf{U}(t) = e^{\frac{i}{\hbar}\mathbf{H}t}$.

A canonical transformation of the form

$$\mathsf{U} = e^{if(\mathbf{q},t)},$$

where f is a function of t and the $\mathbf{q}$s only, is called a *gauge transformation of the first kind.* It leaves the particle-position operators $\mathbf{q}_k$ unchanged, while the momentum operators are changed from $\mathbf{p}_k$ to

$$\mathbf{p}'_k = e^{if}\mathbf{p}_k e^{-if} = e^{if}[\mathbf{p}_k, e^{-if}] + \mathbf{p_k} = -\hbar\frac{\partial f}{\partial \mathbf{q}_k} + \mathbf{p}_k.$$

Consider, then, a system of particles of charge e. Recall that if the vector potential $\vec{A}(\vec{r}, t)$ is replaced by $\vec{A}'(\vec{r}, t) = \vec{A}(\vec{r}, t) + \nabla f(\vec{r}, t)$ and if at the

same time the scalar potential is changed from $\phi(\vec{r},t)$ to $\phi' = \phi - \frac{1}{c}\frac{\partial}{\partial t}f$, the electric and magnetic fields remains unchanged.[40] Such a change in the potentials is called a *gauge transformation of the second kind.* Therefore, if the canonical momentum operators for a particle of charge e always appear together with the vector potential in the combination $\mathbf{p}_k - \frac{e}{c}\mathbf{A}_k$, then a gauge transformation of the second kind changes

$$\mathbf{p}_k - \frac{e}{c}\mathbf{A}_k \longrightarrow \mathbf{p}_k - \frac{e}{c}\mathbf{A}_k - \frac{e}{c}\frac{\partial}{\partial \mathbf{q}_k}f,$$

which can be compensated by a gauge transformation of the first kind by setting $\mathsf{U} = \exp(\frac{ie}{\hbar c}f)$:

$$\mathbf{p}_k - \frac{e}{c}\mathbf{A}'_k = \mathsf{U}\left(\mathbf{p}_k - \frac{e}{c}\mathbf{A}_k\right)\mathsf{U}^{-1}. \tag{2.113}$$

This is why, just as in classical mechanics, the vector potential always has to appear together with the canonical momentum in the combination

$$\vec{\mathbf{p}} - \frac{e}{c}\vec{\mathbf{A}}, \tag{2.114}$$

called the "mechanical momentum." Moreover, if the Hamiltonian of a particle of charge e in the presence of electric and magnetic fields reads

$$\mathbf{H} = \frac{1}{2M}\left(\vec{\mathbf{p}} - \frac{e}{c}\vec{\mathbf{A}}\right)^2 + e\phi, \tag{2.115}$$

and similarly for multiparticle systems, then the Hamiltonian for the gauge-transformed potentials is given by

$$\mathbf{H}' = \frac{1}{2M}\left(\vec{\mathbf{p}} - \frac{e}{c}\vec{\mathbf{A}}'\right)^2 + e\phi' = \mathsf{U}\mathbf{H}\mathsf{U}^{-1} - i\hbar\mathsf{U}\frac{\partial}{\partial t}\mathsf{U}^{-1},$$

as a result of which $\Psi' \stackrel{\text{def}}{=} \mathsf{U}\Psi$ satisfies the Schrödinger equation $i\hbar\frac{\partial}{\partial t}\Psi' = \mathbf{H}'\Psi'$ if Ψ satisfies $i\hbar\frac{\partial}{\partial t}\Psi = \mathbf{H}\Psi$. Since

$$\frac{1}{2M}\left(\vec{\mathbf{p}} - \frac{e}{c}\vec{\mathbf{A}}\right)^2 = \frac{1}{2M}\vec{\mathbf{p}}^2 + \frac{e^2}{2Mc^2}\vec{\mathbf{A}}^2 - \frac{e}{2Mc}(\vec{\mathbf{p}}\cdot\vec{\mathbf{A}} + \vec{\mathbf{A}}\cdot\vec{\mathbf{p}}),$$

the interaction term in the Hamiltonian resulting from (2.115) is given by the last term, which is the equivalent of the classical $(1/c)\vec{A}\cdot\vec{j}$ for charged particles of velocity $\vec{p}/M$.

For the field, a gauge transformation of the first kind takes the form

$$\Psi(\vec{r},t) \longrightarrow \Psi'(\vec{r},t) = e^{if(\vec{r},t)}\Psi(\vec{r},t), \tag{2.116}$$

[40] **Show this, using (2.92).**

so that

$$\nabla\Psi' = e^{if}[\nabla\Psi + i(\nabla f)\Psi].$$

Therefore if ∇ appears in the form

$$\nabla + i\frac{e}{\hbar c}\vec{\mathbf{A}} \tag{2.117}$$

and whenever Ψ is subjected to the gauge transformation (2.116), the vector potential undergoes the gauge transformation of the second kind, $\vec{\mathbf{A}} \longrightarrow \vec{\mathbf{A}}' = \vec{\mathbf{A}} + \nabla g$, where $g = \frac{\hbar c}{e} f$, then

$$(\nabla - i\frac{e}{\hbar c}\vec{\mathbf{A}})\Psi \longrightarrow (\nabla - i\frac{e}{\hbar c}\vec{\mathbf{A}}')\Psi' = e^{if}(\nabla - i\frac{e}{\hbar c}\vec{\mathbf{A}})\Psi.$$

Similarly, the combination

$$i\hbar\frac{\partial}{\partial t} - e\phi \tag{2.118}$$

is such that

$$\left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\Psi \longrightarrow \left(i\hbar\frac{\partial}{\partial t} - e\phi'\right)\Psi' = e^{if}\left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\Psi.$$

Equation (2.68) shows that the Lagrangian remains invariant under such combined gauge transformations, from which it follows that the field equations are unchanged, and so are the canonical commutation relations.[41] The Schrödinger equation (2.69) for the (charged, that is, coupled to the electromagnetic field) matter field therefore has to contain the vector potential in the same manner as the Schrödinger equation for the state vector.

If, in the equations governing the field or the state vector of charged particles, the electromagnetic potentials enter just in the form given by (2.114), (2.117), and (2.118), and in no additional terms, the result is called *minimal coupling* to electromagnetism.

[41] This line of reasoning may be extended relativistically so that the postulate of gauge invariance (of the first kind) of the matter-field equations leads to the need for the existence of an electromagnetic vector potential, coupled in such a way that its gauge invariance (of the second kind) compensates the first, and which, in turn, has to satify the Maxwell equations. Even more generally, if the group of gauge transformations is taken to be non-abelian, the existence of other kinds of fields follows. All modern relativistic field theories are generated by arguments of this kind.

2.5 Problems and Exercises

1. Show that the commutation relations (2.11), (2.12), and (2.16) imply that for any differentiable function $f(\ldots, \mathbf{q}_i, \ldots; \ldots, \mathbf{p}_i, \ldots)$ of the $\mathbf{p}$s and $\mathbf{q}$s,
$$[\mathbf{p}_j, f] = -i\hbar \frac{\partial f}{\partial \mathbf{q}_j}, \qquad [\mathbf{q}_j, f] = i\hbar \frac{\partial f}{\partial \mathbf{p}_j}.$$
Hint: Begin by proving it for polynomials.

2. Show that if $f(\mathbf{q})$ is a differentiable function of $\mathbf{q}$, then
$$f(\mathbf{q})e^{-i\mathbf{p}c/\hbar} = e^{-i\mathbf{p}c/\hbar} f(\mathbf{q} + c).$$

3. Show that if $f(\mathbf{p})$ is a differentiable function of $\mathbf{p}$, then
$$f(\mathbf{p})e^{i\mathbf{q}c/\hbar} = e^{i\mathbf{q}c/\hbar} f(\mathbf{p} + c).$$

4. Prove that the parity operator commutes with the operator $\frac{d^2}{dx^2}$.

5. Show that the gaussian wave function (2.50) is properly normalized and calculate the momentum uncertainty as defined by (1.10), i.e., confirm (2.51).

6. Do the analogues of (2.50) to (2.54) in three dimensions.

7. Calculate the analogue of (2.58) in three dimensions.

8. Derive (2.105) by defining the momentum as the generator of translations of the whole system, using the generator (2.9).

9. Derive (2.106) by defining the angular momentum as the generator of rotations, using (2.9).

10. If an electron is initially localized within a region of diameter 10^{-4}cm, what is the size of the region in which it will be localized 1 sec later?

11. Suppose you want to minimize the uncertainty of an electron's position .01 sec after its initial position measurement. With what accuracy should you determine its inital location?

3

The Schrödinger Equation in One Dimension

According to Eq. (2.43), the time-independent Schrödinger equation for a one-particle system reads

$$\left[\frac{\vec{\mathbf{p}}^2}{2M} + V(\vec{\mathbf{q}})\right]\Psi_E = E\Psi_E, \tag{3.1}$$

assuming that the particle is subject to the potential $V(\vec{q})$, or if the particle is confined to a line,

$$\left[\frac{\mathbf{p}^2}{2M} + V(\mathbf{q})\right]\Psi_E = E\Psi_E. \tag{3.2}$$

3.1 A Free Particle

For a free particle in one dimension, subject to no forces or constraints, (3.2) becomes the simple equation

$$\frac{\mathbf{p}^2}{2M}\Psi = E\Psi, \tag{3.3}$$

so that E is a quasi-eigenvalue of the kinetic-energy operator $\frac{\mathbf{p}^2}{2M}$. Since we have already established that the spectrum of the momentum operator $\mathbf{p}$ of an unconfined particle is the entire real line, it follows that the spectrum of the kinetic energy operator consists of the positive real line: all non-negative numbers are in the continuous spectrum of the kinetic energy. In

the momentum representation, (3.3) becomes

$$\frac{\mathbf{p}^2}{2M}\widehat{\psi}(p) = E\widehat{\psi}(p), \tag{3.4}$$

and the Hermitian operator $\mathbf{p}^2$ can be replaced by the non-negative number p^2. This equation is easily solved for any given non-negative E; it has two linearly independent solutions,

$$\widehat{\psi}(\pm p', p) = \delta(p \mp p'), \quad p' = +\sqrt{2ME}, \tag{3.5}$$

neither of which is in the Hilbert space $L^2(\mathbb{R})$. Therefore every quasi-eigenvalue of the kinetic-energy operator is doubly degenerate. The physical reason for this, of course, is that for a given kinetic energy, the momentum of the particle can be positive or negative, pointing to the right or to the left. The quasi-eigenfunctions in (3.5) are normalized so that

$$\int_{-\infty}^{\infty} dp\, \widehat{\psi}^*(p', p)\widehat{\psi}(p'', p) = \delta(p' - p'').$$

Using (1.73) and (3.5), we find the configuration-space wave functions of a free particle of energy $E = p^2/2M \stackrel{\text{def}}{=} \hbar^2 k^2/2M$ to be

$$\psi(\pm k, q) = (2\pi\hbar)^{-1/2} e^{\pm ikq}, \tag{3.6}$$

a result we already obtained in the last chapter, in a somewhat different notation. It will be convenient, instead of insisting that $k = +\sqrt{2ME}/\hbar$, to allow the wave number k to be positive or negative, which facilitates the statement of the normalization of (3.6) to be

$$\int_{-\infty}^{\infty} dq\, \psi^*(k, q)\psi(k', q) = \delta(k - k')/\hbar = \delta(p - p').$$

That these functions form a complete set on $L^2(\mathbb{R})$ is expressed by the equation

$$\int_{-\infty}^{\infty} dp\, \psi^*(p/\hbar, q)\psi(p/\hbar, q') = \delta(q - q').$$

Since Eq. (3.4) is invariant under reflection, i.e., under the parity transformation, it can be expected to have odd and even solutions. Such solutions can be readily formed out of those given by (3.6), namely, $\sin kq$ and $\cos kq$. *The fact that the quasi-eigenvalues are degenerate allows us to form solutions such as (3.6) that violate the symmetry of the equation,* a feature of great physical significance possessed by many other systems: the symmetry of the equation can be *spontaneously broken.* Ultimately, the breaking of the symmetry for individual systems is determined by their initial or boundary conditions: in this particular instance, singling out motion to the left or right.

The wave functions (3.5) are, of course, not normalizable and cannot represent the state of "a free particle." As noted earlier, they may instead be regarded as describing *beams* of infinitely many (independent) free particles of energy $\hbar^2k^2/2M$. The way to recognize that the wave function $e^{i|k|q}$ describes a beam of particles moving to the *right* and $e^{-i|k|q}$ to the *left* is to remember that the corresponding time-dependent solutions of the Schrödinger equation are $e^{i(|k|q-\omega t)}$ and $e^{-i(|k|q+\omega t)}$, respectively; the phase of the first stays constant when q *increases* by $\omega t/|k| = t\hbar|k|/2M$ as t increases, while that of the second stays constant when q *decreases* by the same amount. A point of fixed phase of the first thus moves rightward and a point of fixed phase of the second, leftward; the flux density of the first is $\hbar|k|/M$; that of the second, $-\hbar|k|/M$.

3.2 A Particle With Potential Energy

In one dimension we will encounter the simplest instances of the most important phenomena described by the one-particle Schrödinger equation in three dimensions: scattering of particles, resonances with their time delays, bound states, continuous and band spectra. Let's look at some specific cases and then generalize.

3.2.1 The square potential barrier

We begin with a particle moving along the x-axis, encountering a potential barrier. (From now on we shall write simply x for the particle's coordinate, but you should keep in mind that x is really in *configuration space*, not in physical space.) Suppose the particle is free in region (i), where $x < -a/2$,

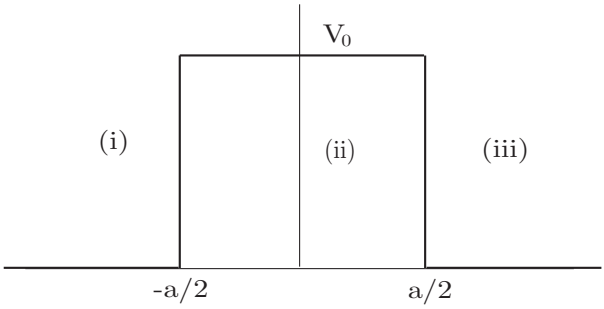

FIGURE 3.1. Square potential barrier of height V_0 and width a.

and in region (iii), where $x > a/2$, while in region (ii), where $-a/2 < x < a/2$, its potential energy has the constant positive value V_0 (Fig. 3.1). At the points $x = -a/2$ and $x = a/2$, but nowhere else, it encounters forces, the first pointing to the left, and the second to the right, so that in both

cases the force repels the particle away from the center of the potential region. Classically, such a particle, coming in from the left, say, is reflected back if its energy is less than V_0, and passes through the region, after being temporarily slowed down, if its energy is greater than V_0. This is not what happens quantum mechanically.

The initial condition that particles come in from the left means that, while in region (i) we can expect flux toward the right (the incoming beam, whose normalization is under our control) as well as to the left (the reflected beam), so that there we may write $\psi_l(x) = e^{ikx} + Re^{-ikx}$, with the constant R to be determined, in region (iii) there should be no flux toward the left: if there are any particles in this region, they should all move to the right. So the solution in region (iii) should be of the form $\psi_l(x) = Te^{ikx}$, where T is another constant to be determined. In region (ii), the equation reads $-\frac{\hbar^2}{2M}\psi_l'' = (E - V_0)\psi_l$, with the solution $\psi_l = be^{i\kappa x} + ce^{-i\kappa x}$, where

$$\kappa \stackrel{\text{def}}{=} \sqrt{2M(E - V_0)}/\hbar.$$

In order for ψ_l to solve the Schrödinger equation, a second-order differential equation, both ψ_l and ψ_l' have to be continuous at the point $x = a/2$, which leads to the two equations

$$Te^{ika/2} = be^{i\kappa a/2} + ce^{-i\kappa a/2}$$

and

$$kTe^{ika/2} = \kappa(be^{i\kappa a/2} - ce^{-i\kappa a/2}).$$

Similarly, the two equations of the continuity of ψ_l and of ψ_l' at $x = -a/2$ read

$$Re^{ika/2} + e^{-ika/2} = be^{-i\kappa a/2} + ce^{i\kappa a/2}$$

$$k\left[-Re^{ika/2} + e^{-ika/2}\right] = \kappa\left[be^{-i\kappa a/2} - ce^{i\kappa a/2}\right].$$

These four equations for the four unknowns, T, R, b, and c are easily solved, and we obtain

$$T = \frac{2k\kappa e^{-iak}}{2k\kappa\cos a\kappa - i(k^2 + \kappa^2)\sin a\kappa} \tag{3.7}$$

$$R = \frac{i\sin a\kappa\, e^{-iak}(\kappa^2 - k^2)}{2k\kappa\cos a\kappa - i(k^2 + \kappa^2)\sin a\kappa}. \tag{3.8}$$

(The two constants b and c are generally of little interest.)

You should, at this point, wonder why the wave function is required to be continuous and to have a continuous first derivative if all that quantum-mechanically matters is that it be locally square-integrable. The reason is the demand that the wave function also be in the domains of definition of the momentum $\mathbf{p}$ and the kinetic energy operator $\mathbf{p}^2/2M$. A step discontinuity would prevent the wave function from being in the domain of the

momentum operator (the derivative would produce a delta function, which is not locally square integrable), and a discontinuity of the first derivative would prevent it from being in that of the square of the momentum. Therefore, the quantum requirements based on Hilbert space considerations are identical to the demands of the classical theory of differential equations.

Consider the implications of the expressions (3.7) and (3.8). A wave packet formed out of a superposition of functions of the form $\psi_l(x)$ will, for large negative times, have its main contribution from the e^{ikx}-term in the region where $x < 0$, because that is the only term that can have a stationary phase; so this is the term that describes the *incoming* particle. For large positive times, on the other hand, the term e^{ikx} on the left cannot contribute, whereas both the term Re^{-ikx} on the left and the term Te^{ikx} on the right will; thus these terms describe the *outgoing* waves. If the solution ψ_l is denoted by $\psi_l(k, x)$, the label k means that *in the infinite past, the particle had the momentum $\hbar k$;* it does *not* mean that the particle's momentum remains $\hbar k$; this momentum, of course, is not a constant of the motion and is even undefined during part of the journey.

With the normalization chosen, the flux coming in from the left is given by $k\hbar/M$, that going out to the right is $|T|^2 k\hbar/M$, and the magnitude of that being reflected back to the left is $|R|^2 k\hbar/M$. Therefore, if by the *transmission coeffient* we mean the ratio of the magnitudes of the transmitted to the incoming fluxes, it is given by

$$|T|^2 = \frac{4E(E - V_0)}{4E(E - V_0) + V_0^2 \sin^2 a\kappa} , \tag{3.9}$$

and if the *reflection coefficient* is defined as the ratio of the magnitudes of the reflected to the incoming fluxes, it is

$$|R|^2 = \frac{V_0^2 \sin^2 a\kappa}{4E(E - V_0) + V_0^2 \sin^2 a\kappa} . \tag{3.10}$$

[See Figures 3.2 and 3.3 for plots of Eq. 3.10.] It is easy to check that

$$|R|^2 + |T|^2 = 1, \tag{3.11}$$

an important equation that expresses the *conservation of flux*, that is, the physical requirement that the reflected and transmitted fluxes must add up to the incoming flux: $\hbar k|R|^2/M + \hbar k|T|^2/M = \hbar k/M$.

The transmission and reflection coefficients for particles coming in from the right instead of from the left are obtained similarly, by defining a solution $\psi_r(x)$ that is of the form $\psi_r(x) = T_r e^{-ikx}$ on the left and of the form $\psi_r(x) = R_r e^{ikx} + e^{-ikx}$ on the right. The Wronskian $W(\psi_r, \psi_l) \stackrel{\text{def}}{=} \psi_r\psi_l' - \psi_r'\psi_l$ of these two solutions is a constant,[1] and its value for large

[1] **Why?**

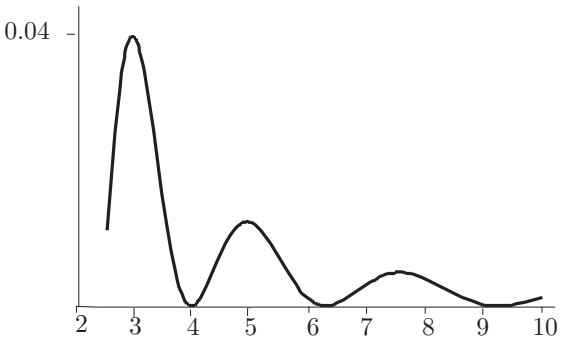

FIGURE 3.2. The reflection coefficient (3.10) plotted as a function of $x \stackrel{\text{def}}{=} E/V_0$ for $2MV_0a^2/\hbar^2 = 30$.

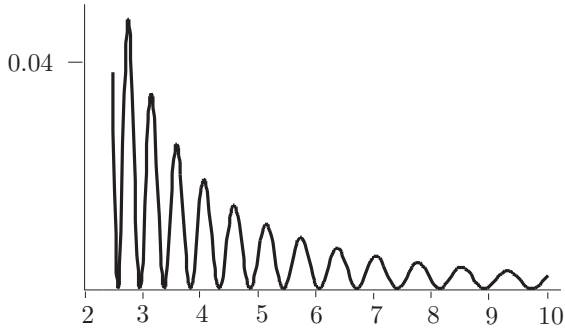

FIGURE 3.3. The reflection coefficient (3.10) plotted as a function of $x \stackrel{\text{def}}{=} E/V_0$ for $2MV_0a^2/\hbar^2 = 500$.

positive x is $2ikT$, while its value for large negative x is $2ikT_r$. We can therefore conclude that $T = T_r$; that is, *the transmission amplitudes for left and right incidence are always equal.* In this particular case we can say more: since the potential is invariant under reflection, so must be the values of R and R_r; hence both the formulas (3.7) and (3.8) hold for left-incidence as well as right-incidence.

These results have a number of interesting features. Notice the pronounced diffraction effects near the "edge," when $E \simeq V_0$: at those values of $E \geq V_0$ for which $a\kappa = n\pi,\ n = 0, 1, 2, \ldots$, there is full transmission and no reflection, the energy differences from one such transmission maximum to the next being $\Delta E = \pi^2\hbar^2/2Ma^2$. If the phase of T is defined to be $\varphi - ak$, i.e., $T = |T|e^{i\varphi - iak}$, (3.7) tells us that

$$\tan\varphi = \frac{k^2 + \kappa^2}{2k\kappa} \tan a\kappa,$$

and its derivative when $\sin a\kappa = 0$ is

$$\frac{d(\varphi - ak)}{dk} = \frac{a}{2\kappa^2}(k^2 - \kappa^2),$$

which implies, according to (2.61), that there is a *time delay* τ in the transmitted wave,

$$\tau = \frac{aM}{2\hbar k\kappa^2}(k^2 - \kappa^2) = \frac{aM^2}{\hbar^3 k\kappa^2}V_0. \tag{3.12}$$

This should be compared to the classical time delay caused by the slowing down of the particle as it crosses the barrier, where its velocity is $p/M = \hbar\kappa/M$ instead of its original $\hbar k/M$, as a result of which $\tau_{\rm cl} = aM/\hbar\kappa - aM/\hbar k = 2\tau\kappa/(k+\kappa)$. If $\kappa \ll k$, so that the particle skims over the top of a high potential barrier,

$$\tau = \frac{k}{2\kappa}\tau_{\rm cl} \gg \tau_{\rm cl}; \tag{3.13}$$

the quantum-mechanical time delay is then much larger than the classical one.

At certain energies two phenomena occur together: 1) the reflection coefficient vanishes—the barrier becomes transparent—and 2) the transmitted particle emerges after a lengthy delay—the particles linger near the barrier for a long time; the combination of these two effects is called a *resonance. A delay in the emerging flux is an important part of a resonance phenomenon,* not just the maximal transmission. Except at these special energies, there is always *some* reflection; however, since $|R|^2 \le V_0^2/(2E-V_0)^2 = 1/(2\frac{E}{V_0}-1)^2$, when $E \gg V_0$, there is essentially full transmission: at high energies, the barrier becomes practically invisible, as it is classically.

When the energy is below the top of the barrier, $E < V_0$, (3.9) and (3.10) still hold, but κ becomes imaginary and $|T|^2$ can be written in the form

$$|T|^2 = \frac{1}{1 + \frac{V_0^2}{4E(V_0-E)}\sinh^2\left(\frac{a}{\hbar}\sqrt{2M(V_0-E)}\right)}. \tag{3.14}$$

That this differs from zero shows that there is transmission even below the top of the barrier; the particles can penetrate it, even though, classically, they could not. This phenomenon is usually referred to as the *tunnel effect.* Far below the top, when $2Ma^2(V_0-E)/\hbar^2 \gg 1$, the tunneling probability can be written

$$|T|^2 \simeq 16\frac{E}{V_0}(1-\frac{E}{V_0})e^{-2a\sqrt{2M(V_0-E)}/\hbar}, \tag{3.15}$$

an expression that can be used as a rough approximation in more general settings. It allows us, for example, to estimate the lifetime of α-particles "trapped" inside a nucleus by its potential barrier. (This is left as a problem

for you to calculate.) At the same time as the amplitude of the transmitted wave function is reduced by transmission through the barrier and few particles manage to penetrate it, they are also *retarded* by the journey. As shown in Eq.(2.61), this effect can be calculated by means of the k-derivative of the phase of T, and we find that in the limit when (3.15) is valid, the wave packet center has been shifted back by the width a of the barrier.[2]

What are the consequences of rigidly *translating* the position of the potential? Setting $x = x' - \zeta$ in the Schrödinger equation has the effect of shifting the potential by ζ to the right. (The origin of the x-axis, i.e., the center of the barrier, is located at $x' = \zeta$.) In region (iii), the solution ψ_l, when multiplied by the constant $e^{ik\zeta}$, will now have the form $Te^{ikx'}$, and in region (i) it has the form $Re^{2ik\zeta}e^{-ikx'} + e^{ikx'}$. Therefore, T is unchanged and R has to be replaced by $Re^{2ik\zeta}$; neither $|T|^2$ nor $|R|^2$ is changed.

3.2.2 The potential well

The results we have obtained for $E > V_0 > 0$ are equally valid when $E > 0$ and $V_0 < 0$, so that the potential barrier becomes a *potential well* (Fig. 3.4), which, from both sides, exerts a force on the particle that attracts it towards the center. In that case, however, the energy spectrum will consist not only of the positive real line, but there will also be discrete eigenvalues. These are found as follows.

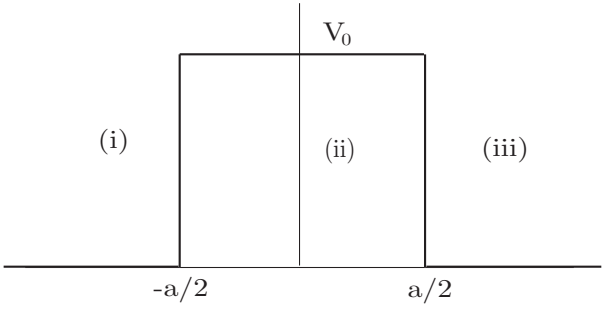

FIGURE 3.4. Potential well of depth V_0 and width a.

First of all, in order for a solution to be square integrable, it cannot be of the form $e^{\pm ikx}$ outside the well, with k real; instead we have to look for decreasing solutions of the form $e^{-|k|x}$ on the right and $e^{|k|x}$ on the left, with $E < 0$ and $|k| = \sqrt{-2ME}/\hbar$. The discrete eigenvalues therefore have to be negative. This makes very good physical sense: the potential energy at infinite distance being zero—which can be regarded as the definition of the zero point of the energy—a particle with positive energy can escape

[2]**Show this as an exercise.**

to infinity, while a particle with negative energy cannot. As a result, it is a general fact that whenever the potential vanishes at infinity (sufficiently rapidly), the continuous energy spectrum occupies the entire positive half-axis, while the discrete eigenvalues (if any) are negative.[3]

Since the potential is an even function of x, we are going to look for odd and even solutions. In region (iii), the solution must be of the form $be^{-|k|x}$ while in region (ii), the odd and even solutions are, respectively, $\psi = c\sin\kappa x$ and $\psi = c\cos\kappa x$, where now $\kappa^2 = -|k|^2 + 2M|V_0|/\hbar^2$. At $x = a/2$ we have to match the *logarithmic derivatives* of the solutions. (This is a convenient way of making sure that both the function and its derivative are continuous, and it eleminates an uninteresting constant factor.) For the odd solution, this requires

$$|k| = -\kappa\cot(\kappa a/2), \tag{3.16}$$

and for the even solution,

$$|k| = \kappa\tan(\kappa a/2). \tag{3.17}$$

Plotting the left- and right-hand sides of (3.16) as functions of κ shows the solutions as the intersections of the curve $f(\kappa) = \sqrt{2M|V_0|/\hbar^2 - \kappa^2}$ with the curves $g(\kappa) = -\kappa\cot(\kappa a/2)$. Since $f(\kappa) > 0$ up to the point $\kappa = \sqrt{2M|V_0|}/\hbar$ and $g(\kappa)$ does not turn positive until $\kappa > \pi/a$, there can be no intersection of f and g unless $|V_0| > \pi^2\hbar^2/2Ma^2$. (Fig. 3.5.) As $|V_0|$ increases, the number of intersections grows without limit.

For the even solutions, (3.17) shows that we are looking for the intersections of the curve $f(\kappa)$ with $h(\kappa) = \kappa\tan(\kappa a/2)$. Since $h(\kappa)$ starts out positive at the origin and increases to infinity at $\kappa = \pi/a$, the two curves will always intersect at least once, so that there is *always* at least one solution of (3.17).

The solutions we have found are the *bound states* of the particle in the presence of a square well potential, which produces an inward-directed kick at the two points $x = \pm a/2$ but no other forces. Whereas classically, such a particle will be confined to the interior if its energy is negative, quantum mechanicaly it has a finite probability of being found outside. (To calculate this probability will be one of your homework exercises.) The eigenvalues are those values of the energy at which the particle can remain in a steady state without escaping to infinity. This is because if at the time $t = 0$ the particle is in an eigenstate $\psi_E(x)$ of $\mathbf{H}$ with the eigenvalue E, then at the later time t its state is $\psi(x,t) = \psi_E(x)e^{-iEt/\hbar}$, so that the probability of finding the particle in the volume element dx is $|\psi(x,t)|^2dx = |\psi_E(x)|^2dx$, just as it was at $t = 0$. By contrast, for any state made up of a superposition of continuum quasi-eigenstates, such as $\int dk\, e^{-i\hbar k^2t/2M}f(k)\psi(k,x)$,

[3]There are instances in which the Schrödinger equation has positive discrete eigenvalues, which therefore are *embedded in the continuum*, but for the one-particle Schrödinger equation these are exceptional cases, which occur only for rather pathological potentials.

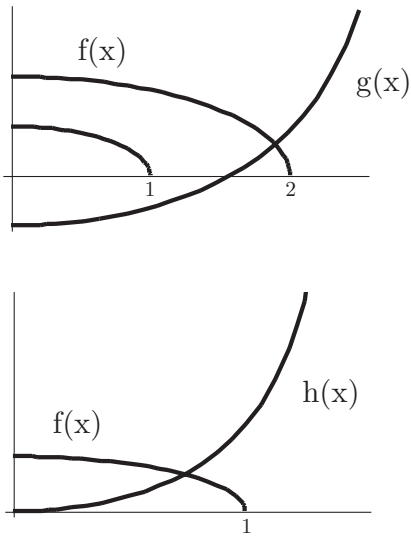

FIGURE 3.5. The upper figure shows the function $g(x) = -x \cot x$, the right-hand side of (3.16), and the function $f(x) = \sqrt{A^2 - x^2}$, its left-hand side, with $A = 1, 2$. The lower figure shows $h(x) = x \tan x$, the right-hand side of (3.17), and $f(x)$ with $A = 1$. Here $x \stackrel{\text{def}}{=} a\kappa/2$ and $A^2 \stackrel{\text{def}}{=} M|V_0|a^2/2\hbar^2$.

the probability density tends to zero as $t \to \pm\infty$: the probability of finding the particle in any given finite region in the infinity past or future vanishes. *This difference is the essential characteristic of bound states,* not just in one dimension, but in three dimensions as well. (On the other hand, the fact that, *no matter how weak the negative potential* $-|V_0|$ *is, there is always at least one bound state,* is a general feature of the Schrödinger equation in one dimension; *it does not hold in three dimensions.*)[4] The existence of such spatially self-confined states of certain specific energies is one of the most characteristic features of quantum mechanics as distinct from classical mechanics. As a consequence, two identical systems in the same bound state are completely indistinguishable and stable: two hydrogen atoms in isolation are truly identical stable systems.

As the potential well is made deeper and deeper it becomes convenient to set the origin of energy at the bottom of the well, which means setting $E = \hbar^2\kappa^2/2M$ and $k^2 = 2M(E - |V_0|)/\hbar^2 \to -\infty$. In the region outside the potential well, the wave function then vanishes identically in the limit, and we are simply left with the problem to solve the Schrödinger equation $\psi'' = -\kappa^2\psi$ for $-a/2 < x < a/2$, with the boundary conditions $\psi(-a/2) =$

[4]You should also note that the threshold potential strength at which the first odd bound state appears just at the energy $E = 0$, i.e., when $|V_0| = \pi^2\hbar^2/2Ma^2$, the reflection coefficient vanishes at $E = 0$.

$\psi(a/2) = 0$.[5] In agreement with (3.16) and (3.17) in the limit as $|V_0| \to \infty$, this problem has the odd solutions $\psi = \sin \kappa x$ with $\kappa = \sqrt{2ME}/\hbar = 2n\pi/a,\ n = 1, 2, \ldots$, and the even solutions $\psi = \cos \kappa x$ with $\kappa = (2n + 1)\pi/a,\ n = 0, 1, \ldots$. The ground state has the energy $E = (\hbar\pi/a)^2/2M$ and (in accordance with Sturm-Liouville theory) its wave function is *nodeless,* while the wave function of the n^{th} level has $n - 1$ nodes.

3.2.3 General transmission and reflection

The physical effects described for the cases of square potential wells and barriers are of general validity. Whenever the potential vanishes sufficiently rapidly as $|x| \to \infty$, it is useful to define two linearly independent solutions ψ_l and ψ_r of the Schrödinger equation for $E > 0$ by the following boundary conditions: the function $\psi_l(x)$ is defined by the requirements that as $x \to -\infty$

$$\psi_l'(x) + ik\psi_l(x) - 2ike^{ikx} \to 0, \tag{3.18}$$

while as $x \to +\infty$,

$$\psi_l'(x) - ik\psi_l(x) \to 0; \tag{3.19}$$

and $\psi_r(x)$ is defined by the requirements that as $x \to +\infty$,

$$\psi_r'(x) - ik\psi_r(x) + 2ike^{-ikx} \to 0, \tag{3.20}$$

while as $x \to -\infty$

$$\psi_r'(x) + ik\psi_r(x) \to 0. \tag{3.21}$$

It then follows that there exist three constants T, R_l, and R_r such that as $x \to -\infty$,

$$\psi_l(x) = e^{ikx} + R_l e^{-ikx} + O(x^{-1}), \qquad \psi_r(x) = Te^{-ikx} + O(x^{-1}), \tag{3.22}$$

and as $x \to +\infty$,

$$\psi_r(x) = e^{-ikx} + R_r e^{ikx} + O(x^{-1}), \qquad \psi_l(x) = Te^{ikx} + O(x^{-1}), \tag{3.23}$$

from which the ratio of the fluxes leads to the reflection and transmission coefficients $|R_l|^2$, $|R_r|^2$, and $|T|^2$. That the two transmission amplitudes are equal follows from the same Wronskian argument as before. Since the conservation of flux again leads to (3.11), it follows that $|R_l|^2 = |R_r|^2 = 1 - |T|^2$, even though the amplitudes R_l and R_r need not be equal.

---

[5] This eigenvalue problem is a special case of the Schrödinger equation on a finite interval, a standard *Sturm-Liouville* problem.

3.2.4 The double well

Suppose now there are two identical wells of width l and with infinitely high walls, located symmetrically about the origin, a distance $2a$ apart. In that case, the energy levels are the sums of any two of the eigenvalues for each individual well. However, there now is a twofold *degeneracy*, because the energy of a state in which there is a wave with n nodes in the left well and one with $m \neq n$ nodes in the right well is equal to that of the reflected state, corresponding to the classical fact that for a given energy, the particle could be either in the left or in the right well. (If the potential is "well behaved," rather than being infinite on a stretch of the real axis, as in this instance, there can be no degeneracy of eigenvalues in one dimension—though there is degeneracy of the quasi-eigenvalues in the continuous spectrum— because that would imply that *every* solution of this second-order ordinary differential equation with a given energy is square integrable.) Notice that, even though the Hamiltonian has reflection symmetry, so that parity is conserved, there are non-symmetric solutions (solutions that are not eigenstates of parity). This *spontaneous symmetry breaking* is possible only because of the degeneracy, which allows the formation of asymmetric superpositions of two reflection-symmetric solutions.

Let us next modify this double well and give it the form shown in Figure 3.6, two wells separated by a barrier of width $2a$ and height V_0. The even

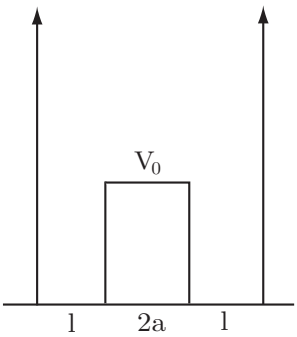

FIGURE 3.6. A double potential well with infinitely high walls.

solutions below the barrier top must have the form $\psi = \cosh \kappa x$ in the barrier ($\kappa = \sqrt{2M(V_0 - E)}/\hbar$), and $\psi = c \sin[k(x - l - a)]$ in the right-hand well. Matching logarithmic derivatives at the right barrier wall leads to the equation

$$\kappa \tanh \kappa a = -k \cot kl. \tag{3.24}$$

The odd solution, on the other hand, should have the form $\psi = \sinh \kappa x$ in the barrier, which leads to the equation

$$\kappa \coth \kappa a = -k \cot kl. \tag{3.25}$$

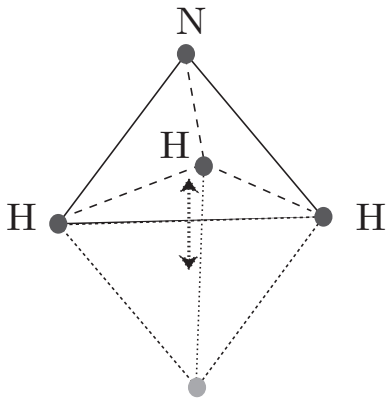

FIGURE 3.7. The oscillating ammonia molecule.

As $V_0 \to \infty$, both (3.24) and (3.25) lead to $|k \cot kl| = \infty$, i.e., $kl = \pi n$, and each level is doubly degenerate, as we anticipated. However, for $a\kappa \gg 1$ we have

$$\coth a\kappa = \frac{1+e^{-2a\kappa}}{1-e^{-2a\kappa}} \sim 1 + 2e^{-2a\kappa}, \quad \tanh a\kappa = \frac{1-e^{-2a\kappa}}{1+e^{-2a\kappa}} \sim 1 - 2e^{-2a\kappa},$$

and the equation for the even solutions becomes

$$\frac{1}{k_e} \tan(k_e l) = -\frac{1}{\kappa}(1 + 2e^{-2a\kappa}),$$

while that for the odd solutions reads

$$\frac{1}{k_o} \tan(k_o l) = -\frac{1}{\kappa}(1 - 2e^{-2a\kappa}).$$

The two levels are now no longer degenerate and we calculate[6] the split between them by setting $k_o = k_e + \epsilon$, obtaining

$$\epsilon = \frac{4k}{\kappa l} e^{-2a\kappa} = \frac{4k\hbar}{l\sqrt{2M(V_0 - E)}} e^{-2a\sqrt{2M(V_0-E)}/\hbar}.$$

Therefore, for a high barrier separating the two wells, the degeneracies are split by

$$\Delta E = \frac{8E\hbar}{l\sqrt{2M(V_0 - E)}} e^{-2a\sqrt{2M(V_0-E)}/\hbar}. \tag{3.26}$$

This result for the energy splitting has direct physical applications. For example, the simple potential shown in Figure 3.6 constitutes a "toy model"

[6] **Do it.**

of the tetrahedral *ammonia molecule* NH_3, whose "inversion spectrum" with a frequency $\nu \sim 0.8\text{cm}^{-1}$ is caused by the effect we have calculated. It corresponds to a classical picture in which the nitrogen atom slowly oscillates from one side of the plane of the three hydrogen atoms to the other (Fig. 3.7). According to (3.26), this oscillation frequency $\nu = \Delta E/2\pi\hbar$ is proportional to the tunneling probability through the potential barrier, as given by (3.15).

3.2.5 The Kronig-Penney potential

The next example we want to investigate is that of a *periodic potential.* Suppose the potential energy is of the form

$$V(x) = \begin{cases} V_0 & \text{if } |x| < a, \\ 0 & \text{if } a < |x| < a + b, \end{cases}$$

and $V(x + l) = V(x)$, with $l = 2a + b$; this is called a *Kronig-Penney* potential (Fig. 3.8). In the n^{th} valley, the solution can be conveniently written in the form

$$\psi_n = A_n \sin k(x + a - nl) + B_n \cos k(x + a - nl),$$

where $k \stackrel{\text{def}}{=} \sqrt{2ME}/\hbar$, while in the bump to the right of the n^{th} valley it may be written in the form

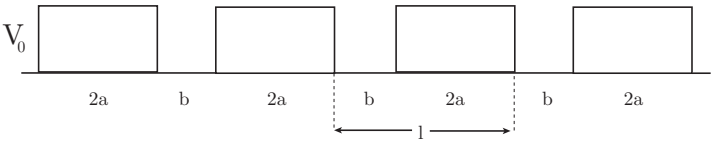

FIGURE 3.8. The Kronig-Penney potential.

$$\psi_n = a_n \sin \kappa(x + a - nl) + b_n \cos \kappa(x + a - nl),$$

where $\kappa \stackrel{\text{def}}{=} \sqrt{k^2 - 2MV_0/\hbar^2}$. Matching ψ_n and its first derivative at $x = nl - a$ gives $B_n = b_n$ and $kA_n = \kappa a_n$; the two matching equations at $x = nl + a$ can be expressed in matrix form by defining

$$\mathcal{A}_n \stackrel{\text{def}}{=} \begin{pmatrix} A_n \\ B_n \end{pmatrix},$$

$$N_1 \stackrel{\text{def}}{=} \begin{pmatrix} -\sin bk & \cos bk \\ \cos bk & \sin bk \end{pmatrix}, \qquad N_2 \stackrel{\text{def}}{=} \begin{pmatrix} (k/\kappa)\sin b\kappa & \cos b\kappa \\ \cos b\kappa & -(\kappa/k)\sin b\kappa \end{pmatrix},$$

so that the equations become $N_1\mathcal{A}_{n+1} = N_2\mathcal{A}_n$, or $\mathcal{A}_{n+1} = \mathcal{M}\mathcal{A}_n$, where $\mathcal{M} \stackrel{\text{def}}{=} N_1^{-1}N_2 = N_1N_2$, since $N_1^2 = \mathbb{1}$. As a result, we have $\mathcal{A}_{n+1} = \mathcal{M}^n\mathcal{A}_1$,

and in order for the solution to remain bounded, both to the left and to the right, $\mathcal{A}_n$ must stay bounded as $n \to \pm\infty$, which means that $\mathcal{M}^n$ must not grow without limit, either as $n \to \infty$ or as $n \to -\infty$. Consider the eigenvalues of $\mathcal{M}$. Since $\det \mathcal{M} = 1$, the equation $\det(\mathcal{M} - m\mathbf{1}) = 0$ for the eigenvalues m of $\mathcal{M}$ reads $m^2 - 2\tau m + 1 = 0$, where $\tau \stackrel{\text{def}}{=} \frac{1}{2}\mathrm{tr}\,\mathcal{M} = \frac{1}{2}(\mathcal{M}_{11} + \mathcal{M}_{22})$. Therefore, $m = \tau \pm \sqrt{\tau^2 - 1}$ and $m_1 m_2 = 1$. In order for $\mathcal{M}^n$ to remain finite as $n \to \pm\infty$, it is necessary and sufficient that $|m_1| = |m_2| = 1$, which requires $\tau^2 \le 1$; if $\tau = \pm 1$ we obtain $m_1 = m_2 = 1$, i.e., $\mathcal{M} = \mathbf{1}$, while $\tau^2 < 1$ leads to $m_1 = m_2^* = \tau + i\sqrt{1 - \tau^2}$. Define $\tau \stackrel{\text{def}}{=} \cos\gamma$, so that $m_1 = e^{i\gamma}$, $m_2 = e^{-i\gamma}$, and the result becomes

$$\cos\gamma = \tau(k) = \cos 2a\kappa \cos bk - \frac{\kappa^2 + k^2}{2k\kappa} \sin 2a\kappa \sin bk. \tag{3.27}$$

[A plot of the function $\tau(k)$ is shown in Figure 3.9.] Those regions of the

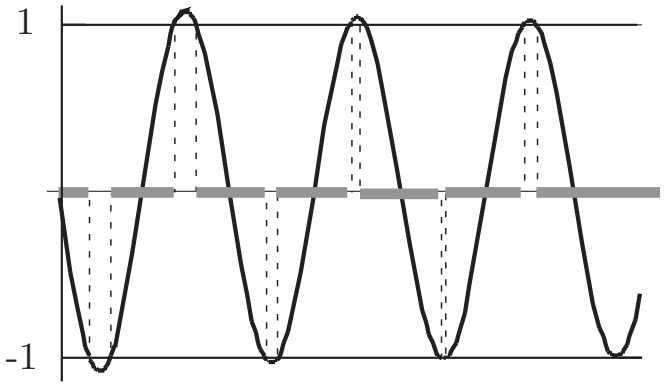

FIGURE 3.9. A plot of the function $F(x) \stackrel{\text{def}}{=} \tau(x/2a)$, where τ is given by (3.27). The heavy gray lines make up the spectrum with its gaps.

real k-axis where $|\tau| \le 1$ produce the allowed values of $E = \hbar^2 k^2/2M$, whereas those stretches where $|\tau| > 1$ are forbidden. In other words, *the spectrum consists of* **bands** *separated by gaps.*

Suppose that the vectors

$$\mathcal{C}^{\pm} \stackrel{\text{def}}{=} \begin{pmatrix} p^{\pm} \\ s^{\pm} \end{pmatrix}$$

are the eigenvectors of $\mathcal{M}$ with the eigenvalues $e^{\pm i\gamma}$, respectively, so that $p^{-*} = p^{+} \stackrel{\text{def}}{=} p$, $s^{-*} = s^{+} \stackrel{\text{def}}{=} s$, and in the n^{th} valley, where $nl - b - a < x < nl - a$,

$$\psi^{\pm}(x) = e^{\pm in\gamma}[p^{\pm} \sin k(x + a - nl) + s^{\pm} \cos k(x + a - nl)]$$

and for $nl - a < x < nl + a$,

$$\psi^{\pm}(x) = e^{\pm in\gamma}[p^{\pm}\frac{k}{\kappa}\sin\kappa(x+a-nl) + s^{\pm}\cos\kappa(x+a-nl)].$$

Define, then, $u^{\pm}(x) \stackrel{\text{def}}{=} e^{\mp ix\gamma/l}\psi^{\pm}(x)$, so that for $-b - a < x - nl < -a$,

$$u^{\pm}(x) = e^{\mp i(x-nl)\gamma/l}[p^{\pm}\sin k(x+a-nl) + s^{\pm}\cos k(x+a-nl)],$$

and for $-a < x - nl < a$,

$$u^{\pm}(x) = e^{\mp i(x-nl)\gamma/l}[p^{\pm}\frac{k}{\kappa}\sin\kappa(x+a-nl) + s^{\pm}\cos\kappa(x+a-nl)],$$

which shows that $u^{\pm}(x+l) = u^{\pm}(x)$. The solution $\psi^{\pm}(x)$ therefore is a product of a function with the same periodicity as the potential and a plane wave with the wave number $K \stackrel{\text{def}}{=} \gamma/l$.[7] The two functions $\psi^{\pm}$ are one another's complex conjugates, $\psi^{-} = \psi^{+*}$, and they are linearly independent, except at the band edges, where $\sin\gamma = 0$, that is, where $Kl = n\pi$, $n = 0, 1, 2\ldots$.

The solutions $\psi^{\pm}(x)$ are called *Bloch functions,* and they play an important role in solid-state physics, since the electrons in a crystalline solid find themselves in the periodic environment provided by the arrangements of the molecules. The one-dimensional Kronig-Penney potential constitutes a very oversimplified model for the motion of electrons along a thin wire. The band structure of their energy spectrum accounts for the existence of electrical conduction, in spite of the fact that if the molecules making up the solid were isolated, the electrons would be bound and localized; it is the regular arrangement of the molecules in a crystal that allows them to penetrate the barriers between neighboring molecules and to roam in the conduction bands as if they were free, albeit with a modified momentum.

3.2.6 The simple harmonic oscillator

The simple one-dimensional harmonic oscillator is probably the most fundamental dynamical system in all of physics. Not only does it arise in many different contexts and guises, but, as stressed already in Chapter 1 and Chapter 2, it lies at the heart of how particles arise from quantum fields in the first place. (See p. 55.)

The potential energy of a simple oscillator of mass M and classical frequency ω, centered at the origin, is given by $V(x) = \frac{1}{2}M\omega^2x^2$, so that its Hamiltonian is $\mathbf{H} = \mathbf{p}^2/2M + \frac{1}{2}M\omega^2\mathbf{q}^2$ and the Schrödinger equation in

[7] According to a theorem by Floquet, wave functions of such a product form exist for any periodic potential, not just for Kronig-Penney potentials.

the configuration representation reads

$$-\frac{\hbar^2}{2M}\frac{d^2\psi}{dx^2}+\frac{1}{2}M\omega^2x^2\psi=E\psi.$$

Defining the operator

$$\mathsf{a}\stackrel{\text{def}}{=}\frac{1}{\sqrt{2}}\left(\xi+\frac{d}{d\xi}\right),\quad\text{so that}\quad\mathsf{a}^\dagger=\frac{1}{\sqrt{2}}\left(\xi-\frac{d}{d\xi}\right),\tag{3.28}$$

where $\xi\stackrel{\text{def}}{=}\sqrt{M\omega/\hbar}\,x$, allows us to transform the Hamiltonian into

$$\mathbf{H}=\hbar\omega(\mathsf{a}^\dagger\mathsf{a}+\frac{1}{2}),$$

while the operators a and $\mathsf{a}^\dagger$ satisfy the commutation relation

$$[\mathsf{a},\mathsf{a}^\dagger]=1.$$

We may now utilize our earlier results for the oscillator in Chapter 1. As we found below Eq.(1.41), the eigenvalues of the operator $\mathsf{N}\stackrel{\text{def}}{=}\mathsf{a}^\dagger\mathsf{a}$ are the non-negative integers, so that the eigenvalues of the simple harmonic oscillator are $E_n=(n+\frac{1}{2})\hbar\omega$, where $n=0,1,2,\ldots$. (You may recall that the "old quantum theory" led to the energies $E_n=n\hbar\omega$, without the *zero-point energy* $\frac{1}{2}\hbar\omega$.)

The ground state is determined by Eq. (1.45) or $\mathsf{a}\phi_0=0$, which now reads

$$\phi_0'+\xi\phi_0=0,$$

the solution of which is $\phi_0(\xi)=ce^{-\frac{1}{2}\xi^2}$; in order to normalize it, choose[8] $c=\pi^{-1/4}$; according to (1.44), the normalized n^{th} eigenfunction is then given by

$$\phi_n(\xi)=\frac{\left[\frac{1}{\sqrt{2}}\left(\xi-\frac{d}{d\xi}\right)\right]^n}{\pi^{1/4}\sqrt{n!}}e^{-\frac{1}{2}\xi^2}.\tag{3.29}$$

The functions defined by

$$H_n(\xi)\stackrel{\text{def}}{=}e^{\frac{1}{2}\xi^2}\left(\xi-\frac{d}{d\xi}\right)^n e^{-\frac{1}{2}\xi^2}\tag{3.30}$$

are called *Hermite polynomials* (see Appendix D.2),[9] and the normalized harmonic oscillator wave functions are expressed in terms of them by

$$\phi_n(\xi)=\frac{1}{2^{n/2}\pi^{1/4}\sqrt{n!}}H_n(\xi)e^{-\frac{1}{2}\xi^2},\tag{3.31}$$

[8] **Show this.**

[9] **Demonstrate that H_n is a polynomial of degree n with n real zeros.**

or explicitly,

$$\psi_n(x) = \frac{(M\omega)^{1/4}}{2^{n/2}(\pi\hbar)^{1/4}\sqrt{n!}} H_n(\sqrt{M\omega/\hbar}\, x) e^{-\frac{M\omega}{2\hbar}x^2}. \tag{3.32}$$

The classical probability density of a particle trapped in a harmonic-oscillator well can be defined by the length of time the particle spends in the interval dx, which is $dx/\dot{x}$. Since classically $x = \sqrt{2E/M\omega^2}\sin\omega t$, the (normalized) classical probability density is given by

$$P_{\rm cl}(x) = \frac{1}{\pi}\sqrt{\frac{M\omega^2}{2E - M\omega^2 x^2}} = \frac{1}{\pi}\sqrt{\frac{M\omega^2}{2E - \omega\hbar\xi^2}}.$$

The quantum mechanical probability density, on the other hand, is given by $P(x) = |\psi(x)|^2$. When comparing the two, it is interesting to note that $H_n(\xi)$ *has no zeros for* $\xi > \sqrt{2n+1}$. [This is proved by writing the Schrödinger equation in the form $\phi''/\phi = \xi^2 - (2n+1)$. When the right-hand side of this equation is positive, $\phi(\xi)$ is convex toward the ξ-axis, and since it vanishes at infinity, it cannot have any zeros.] Therefore all the oscillations

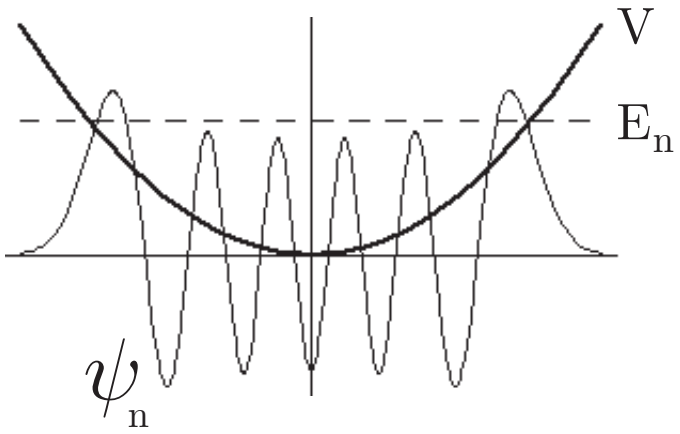

FIGURE 3.10. V is the harmonic-oscillator potential, ψ_n is the corresponding energy eigenfunction for $n = 10$, and E_n is the energy eigenvalue for $n = 10$.

of the wave function are in the region inside the potential, $x^2 < 2E/M\omega^2$, where the particle is classically allowed to be; in the classically forbidden region, $x^2 > 2E/M\omega^2$, the wave functions simply decay (Fig. 3.10).

The expectation value of the potential energy in the n^{th} state can be calculated from the result of #10 of the homework problems and comes

out to be $\langle V\rangle = \frac{1}{2}\hbar\omega(n+\frac{1}{2}) = \frac{1}{2}E_n$, just as it does classically: on average, half the energy is potential energy and half is kinetic.[10]

It is also interesting to see what happens in the course of time to a "minimal wave packet" subject to a harmonic oscillator potential. Suppose that at the time $t = 0$ the wave function is of the form

$$\psi(\xi, 0) = e^{-\frac{1}{2}(\xi-\xi_0)^2}.$$

By means of the generating function (D.43), this can be expanded in the form

$$\psi(\xi) = e^{-\frac{1}{4}\xi_0^2}\sum_{n=0}^{\infty}\frac{(\xi_0/2)^n}{n!}e^{-\frac{1}{2}\xi^2}H_n(\xi).$$

Since $e^{-\frac{1}{2}\xi^2}H_n(\xi) = \phi_n(\xi)$ for the oscillator, the time development of $\psi(\xi, t)$ can therefore be written down directly, using $E_n = \hbar\omega(n+\frac{1}{2})$,

$$\psi(\xi,t) = e^{-\frac{1}{4}\xi_0^2-\frac{1}{2}\xi^2-i\omega t/2}\sum_{n=0}^{\infty}\frac{(\xi_0 e^{-i\omega t}/2)^n}{n!}e^{-\frac{1}{2}\xi^2}H_n(\xi),$$

which can be resummed,

$$\psi(\xi,t) = \exp[-\frac{1}{2}(\xi-\xi_0\cos\omega t)^2]\exp[\frac{i}{2}(\frac{1}{2}\xi_0^2\sin 2\omega t - 2\xi\xi_0\sin\omega t - \omega t)],$$

so that

$$|\psi(\xi,t)|^2 = e^{-(\xi-\xi_0\cos\omega t)^2}.$$

This result is noteworthy for two reasons. First of all, *the packet does not change shape in the course of time*—it does not spread. Second, *the center of the packet simply oscillates with frequency* ω, *just like the classical particle.* Such "minimal wave packets" are special cases of states obtained as follows.

3.2.7 Coherent states

The "lowering operator" a defined by (3.28) is, of course, not Hermitian and does not correspond to an observable; it is nevertheless of interest to find its eigenvalues, if it has any, which satisfy

$$\mathsf{a}\Psi_\alpha = \alpha\Psi_\alpha.$$

Expanding Ψ_α on the basis of the harmonic-oscillator eigenstates $|n\rangle$,

$$\Psi_\alpha = \sum_0^\infty c_n|n\rangle,$$

[10] **Show that this is in agreement with the virial theorem (2.37).**

and using (1.43), which implies that

$$\mathsf{a}|n\rangle = \sqrt{n}|n-1\rangle,$$

we obtain

$$\sum_0^\infty c_n \mathsf{a}|n\rangle = \sum_0^\infty c_{n+1}\sqrt{n+1}|n\rangle = \sum_0^\infty c_n \alpha |n\rangle,$$

from which we conclude $c_{n+1} = \alpha c_n/\sqrt{n+1}$, and therefore

$$c_n = c_0 \frac{\alpha^n}{\sqrt{n!}}.$$

Now let us check under what conditions on α the alleged state Ψ_α is normalizable and therefore acceptable as an eigenstate:

$$(\Psi_\alpha, \Psi_\alpha) = \sum_0^\infty |c_n|^2 = |c_0|^2 \sum_0^\infty \frac{|\alpha|^{2n}}{n!} = |c_0|^2 e^{|\alpha|^2},$$

which means that Ψ_α is normalizable *for all complex values of* α and every complex number α is an eigenvalue of a. (So the "discrete" point spectrum of the non-Hermitian operator a forms a continuum; this cannot happen for Hermitian operators.) In order to normalize Ψ_α we choose

$$c_0 = e^{-\frac{1}{2}|\alpha|^2},$$

and we have explicitly, with $\mathsf{a}\Psi_0 = 0$,

$$\Psi_\alpha = e^{-\frac{1}{2}|\alpha|^2} \sum_0^\infty \frac{(\alpha \mathsf{a}^\dagger)^n}{n!}|0\rangle = e^{-\frac{1}{2}|\alpha|^2 + \alpha \mathsf{a}^\dagger}\Psi_0. \tag{3.33}$$

Since for any eigenvector Ψ_α of a we have $\| \mathsf{a}\Psi_\alpha \| = |\alpha| \, \| \Psi_\alpha \|$ and there is no limit to the size of $|\alpha|$, it follows that the lowering operator a is *unbounded.* Its domain of definition is the set of all vectors $\Psi = \sum b_n|n\rangle$ for which not only $\sum |b_n|^2 < \infty$ but also $\sum n|b_n|^2 < \infty$; the domains of a^m shrink with increasing m.[11]

The eigenstates of a are called *coherent states.* Let us see what they are like in the configuration representation. In terms of[12] $\xi = \sqrt{M\omega/\hbar}\, q$ we have the wave function

$$\psi_0 = \pi^{-1/4} e^{-\frac{1}{2}\xi^2};$$

from (3.28), (3.30), (3.33), and (D.43) we therefore obtain

$$\psi_\alpha \quad = \quad \pi^{-1/4} e^{-\frac{1}{2}|\alpha|^2} \sum_n \frac{1}{n!} \left[\frac{\alpha}{\sqrt{2}}\left(\xi - \frac{d}{d\xi}\right)\right]^n e^{-\frac{1}{2}\xi^2}$$

[11] **Why?**

[12] To avoid confusion, we are going back to using q for the particle's coordinate.

$$\begin{aligned} &= \pi^{-1/4} e^{-\frac{1}{2}|\alpha|^2} \sum_n \frac{1}{n!} \left(\frac{\alpha}{\sqrt{2}}\right)^n e^{-\frac{1}{2}\xi^2} H_n(\xi) \\ &= \pi^{-1/4} e^{-\frac{1}{2}|\alpha|^2} e^{-\frac{1}{2}\xi^2} e^{-\frac{1}{2}\alpha^2 + \sqrt{2}\xi\alpha} \\ &= \pi^{-1/4} \exp\left[\frac{1}{2}\alpha^2 - \frac{1}{2}|\alpha|^2 - \frac{1}{2}(\xi - \sqrt{2}\alpha)^2\right], \end{aligned}$$

and as a result,

$$\psi_\alpha(q) = \pi^{-1/4} \exp\left[\frac{1}{2}\alpha^2 - \frac{1}{2}|\alpha|^2 - \left(\sqrt{M\omega/2\hbar}\, q - \alpha\right)^2\right]. \tag{3.34}$$

So the coherent states are *Gaussian wave packets.* Defining $Z = X + iY \stackrel{\text{def}}{=} \sqrt{2\hbar/M\omega}\,\alpha$ and labeling the coherent states by Z, so that

$$\left(\mathbf{q} + \frac{i}{M\omega}\mathbf{p}\right)\psi_Z(q) = \left(q + \frac{\hbar}{M\omega}\frac{d}{dq}\right)\psi_Z(q) = Z\psi_Z(q),$$

we have explicitly

$$\psi_Z(q) = \pi^{-1/4} \exp\left\{\frac{M\omega}{2\hbar}[-(q-X)^2 + iY(2q+X)]\right\},$$

from which it becomes apparent that $X = \langle \mathbf{q} \rangle$, because that's where the Gaussian packet is centered. Furthermore, since

$$\int dq\, \psi_Z^* \left(-i\hbar \frac{\partial}{\partial q}\right) \psi_Z = -iM\omega Y,$$

it follows that $Y = i\frac{\langle \mathbf{p} \rangle}{M\omega}$, so that

$$\left(\mathbf{q} + \frac{i}{M\omega}\mathbf{p}\right)\psi_Z = \left(\langle \mathbf{q} \rangle + \frac{i}{M\omega}\langle \mathbf{p} \rangle\right)\psi_Z. \tag{3.35}$$

(Note, however, that $\mathbf{q}\psi_Z \neq \langle \mathbf{q} \rangle \psi_Z$ and $\mathbf{p}\psi_Z \neq \langle \mathbf{p} \rangle \psi_Z$.)

We can easily calculate the dispersions of $\mathbf{p}$ and $\mathbf{q}$ in these states: from $\mathbf{q} = \sqrt{\hbar/2M\omega}(\mathsf{a} + \mathsf{a}^\dagger)$ we find that in an eigenstate of a with the eigenvalue α,

$$\begin{aligned} \langle \mathbf{q}^2 \rangle &= \frac{\hbar}{2M\omega} \langle \mathsf{a}^2 + \mathsf{a}^{\dagger 2} + 2\mathsf{a}^\dagger \mathsf{a} + 1 \rangle \\ &= \frac{\hbar}{2M\omega}(\alpha^2 + \alpha^{*2} + 2|\alpha|^2 + 1) \\ &= \frac{\hbar}{2M\omega} + X^2 = \frac{\hbar}{2M\omega} + \langle \mathbf{q} \rangle^2; \end{aligned}$$

therefore,

$$\Delta q^2 = \langle \mathbf{q}^2 \rangle - \langle \mathbf{q} \rangle^2 = \frac{\hbar}{2M\omega},$$

and similarly from $\mathbf{p} = -i\frac{M\omega\hbar}{2}(\mathsf{a} - \mathsf{a}^\dagger)$,

$$\Delta p^2 = \langle \mathbf{p}^2 \rangle - \langle \mathbf{p} \rangle^2 = \frac{M\omega\hbar}{2},$$

so that

$$\Delta p \Delta q = \frac{\hbar}{2}.$$

Thus the coherent states *saturate the uncertainty product.* (That's because they are Gaussian.)

When the oscillator is in a coherent state, the number of quanta, and thus its energy, is not fixed. The time-development of these states is easily calculated, using the Hamiltonian $\mathbf{H} = \hbar\omega(\mathsf{a}^\dagger\mathsf{a} + \frac{1}{2})$. If at the time $t = 0$ the system is in the state

$$|Z\rangle \stackrel{\text{def}}{=} \Psi_Z = e^{-\frac{1}{2}|\alpha|^2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \qquad \alpha = \sqrt{\frac{M\omega}{2\hbar}} Z, \tag{3.36}$$

where $|n\rangle$ is an eigenstate of $\mathsf{N} = \mathsf{a}^\dagger\mathsf{a}$ with the eigenvalue n, then at the time t it is in the state

$$e^{-i\mathbf{H}t/\hbar}|Z\rangle = e^{-\frac{1}{2}|\alpha|^2} e^{-\frac{i}{2}\omega t} \sum_n \frac{\left(\alpha e^{-i\omega t}\right)^n}{\sqrt{n!}} |n\rangle = e^{-\frac{i}{2}\omega t} |Z e^{-i\omega t}\rangle.$$

Since $Ze^{-i\omega t} = X\cos\omega t + Y\sin\omega t + i(Y\cos\omega t - X\sin\omega t)$, we can conclude from (3.35) that

$$\langle \mathbf{q}(t) \rangle = \langle \mathbf{q}_0 \rangle \cos\omega t + \langle \mathbf{p}_0 \rangle \sin\omega t,$$

$$\langle \mathbf{p}(t) \rangle = \langle \mathbf{p}_0 \rangle \cos\omega t - \langle \mathbf{q}_0 \rangle \sin\omega t,$$

which means that these averages behave just like the positions and momenta of the classical oscillator. According to (3.36), the probability w_n of finding n quanta when the system is in a coherent state is given by

$$w_n = \frac{|\alpha|^{2n}}{n!} e^{-|\alpha|^2},$$

and it does not vary with time. For large n, Stirling's formula implies that

$$\log w_n = n - n\log n + 2n\log|\alpha| - |\alpha|^2,$$

which has a maximum at $n_{\max} = |\alpha|^2$. Thus, when the system is in the coherent state Ψ_Z, the number of quanta most likely to be found in it is $n_{\text{ml}} = \frac{M\omega}{2\hbar}|Z|^2$; since each of these quanta has the energy $E = \hbar\omega$, the most likely energy of the oscillator is $E_{\text{ml}} = \frac{1}{2}M\omega^2|Z|^2$, which is exactly the classical energy of a harmonic oscillator of amplitude $|Z|$.

The coherent states of oscillator systems have important physical applications in quantum optics; for further details, see [Klauder].

3.3 Problems and Exercises

1. Prove that

$$\| \left(\frac{d^2}{dx^2} + k^2 \right) e^{ikx - \epsilon x^2} \|^2 \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} dx \left| \left(\frac{d^2}{dx^2} + k^2 \right) e^{ikx - \epsilon x^2} \right|^2 \to 0$$

   as $\epsilon \to 0$.

2. Derive the relativistic dispersion relation in vacuum, i.e., find the relation between the wavelength and the frequency for the de Broglie wave of a free relativistic particle. Find the relation between the phase and group velocities and the particle velocity in the relativistic case. Connect them with those for photons.

3. Calculate the transmission and reflection amplitudes for incidence of electrons of energies 1eV, 1.5eV, 2eV, 2.5eV, 3eV, and 5eV from the left on a rectangular potential barrier of height 2eV and width 3×10^{-8}cm.

4. Use (3.15) to estimate the α-decay lifetime of a uranium nucleus, making reasonable assumptions about the radius of the nucleus ($\sim 10^{-12}$cm), the energy of the α particle ($V_0 - E \sim 12$MeV) and its mean velocity ($\sim 10^9$cm/sec).

5. Calculate the reflection and transmission coefficients for the potential $V(x) = \alpha\delta(x - a)$. (In order to find the effect of the delta-function potential, integrate the Schrödinger equation over a small interval containing the point a.) Find the bound-state energy if there is one. (What is the criterion for the existence of a bound state?)

6. Calculate the reflection and transmission coefficients for the potential $V(x) = \alpha[\delta(x - a) + \delta(x + a)]$.

7. Calculate the reflection and transmission coefficients for the potential $V(x) = 0$ for $x < 0$ and $V(x) = V_0$ for $x > 0$, paying attention to the fact that these coefficients are defined as *flux ratios.*

8. Calculate the reflection and transmission coefficients for both directions of incidence on a potential that vanishes for $x < 0$, has the negative value $-V_0$ for $0 < x < a$, and the positive value V_1 for $x > a$. Also find a transcendental equation for the bound-state eigenvalues.

9. Calculate the normalized bound-state eigenfunction of a particle in the ground state of a square-well potential of width a and depth V_0. What is the probability of finding the particle outside the well? What is the expectation value of its distance from the center of the well?

10. Assuming that a one-particle system is in the n^{th} state of a simple harmonic oscillator, calculate $\langle \mathbf{p}^2 \rangle$, $\langle \mathbf{q}^4 \rangle$, and the Fourier transform of $\psi_n(q)$.

11. Let $|n\rangle, \quad n = 0, 1, \ldots$, be the complete set of normalized eigenstates of a simple harmonic oscillator, and define the operator $\mathcal{E}$ by $\mathcal{E}|n\rangle = |n-1\rangle$. What is $\mathcal{E}^\dagger$? Is $\mathcal{E}$ unitary?

12. Using (1.43), show that if a is the lowering operator for a simple harmonic oscillator, then
$$\mathsf{a}|n\rangle = \sqrt{n}|n-1\rangle.$$

13. Prove that the raising operator $\mathsf{a}^\dagger$ is unbounded and find its domain of definition. Also show that the domains of $\mathsf{a}^{\dagger m}$ shrink with increasing m.

14. Find the eigenvalues of the raising operator $\mathsf{a}^\dagger$.

15. Prove that the *Wronskian* $W(f,g) = fg' - f'g$ of two solutions of the same second-order ordinary differential equation is a constant.

16. Show that the free-particle Schrödinger equation is invariant under Galilean transformations. Do this by showing that when the transformation $x' = x - vt$ is applied (where v is the velocity of the second reference frame with respect to the first), the transformed wave function $\psi'(x',t) \stackrel{\text{def}}{=} f(x,t)\psi(x,t)$ satisfies the same equation as a function x' as does ψ as a function of x, and f involves only $x, t, \hbar, M$, and v. Find the form of f and show that the traveling-wave solution $\psi(x,t) = Ae^{i(kx-\omega t)}$ transforms as expected.

17. Let $\langle \mathbf{q} \rangle$ and $\langle \mathbf{p} \rangle$ be the mean values of $\mathbf{q}$ and its conjugate momentum $\mathbf{p}$ for a system in the state described by the configuration wave function $\psi(q)$. What are the mean values of $\mathbf{q}$ and $\mathbf{p}$ in the state with the wave function
$$e^{-i\langle \mathbf{p} \rangle q/\hbar}\psi(q + \langle \mathbf{q} \rangle)?$$

18. Consider the Schrödinger equation in one dimension with the potential V which consists of the periodic repetition, in both directions, of a given potential $U(x)$ defined for $0 < x < a$. Define two linearly independent solutions on the interval $0 < x < a$ by the boundary conditions $\psi_1(0) = 1$, $\psi_1'(0) = 0$, and $\psi_2(0) = 0$, $\psi_2'(0) = 1$, and derive an inequality that is a criterion for a given energy to be in the spectrum (i.e., to be in an allowed energy band).

19. Consider the Schrödinger equation in one dimension with the potential $V(x) = K^2x^4$. Use the Heisenberg uncertainty relation to estimate the ground-state energy.

4

One- and Two-Particle Systems in Three Dimensions

4.1 Free Particles

In three dimensions, the kinetic energy is of course given by $|\vec{\mathbf{p}}|^2/2M = (\mathbf{p}_x^2+\mathbf{p}_y^2+\mathbf{p}_z^2)/2M$, and the three components of the momentum operator $\vec{\mathbf{p}}$ commute. Since the spectrum of each component is the entire real axis, this is also the common spectrum of all three components, and the spectrum of the kinetic energy operator is the positive real line. However, the quasi-eigenfunctions of $\mathbf{H} = \vec{\mathbf{p}}^2/2M$ in the momentum representation are now given by

$$\widehat{\psi}(\vec{p}\,', \vec{p}) = \delta^3(\vec{p} - \vec{p}\,'), \quad |\vec{p}\,'| = \sqrt{2ME}, \tag{4.1}$$

and each quasi-eigenvalue is *infinitely degenerate,* because for each fixed value of $|\vec{p}\,'|^2$ there is a continuity of directions of the vector $\vec{p}\,'$. The normalization of the free momentum-space wave functions (4.1) is such that

$$\int d^3p\, \widehat{\psi}^*(\vec{p}\,', \vec{p})\widehat{\psi}(\vec{p}\,'', \vec{p}) = \delta^3(\vec{p}\,' - \vec{p}\,'').$$

The free wave functions in the configuration representation are found from (1.70) and (4.1) to be

$$\psi(\vec{k}, \vec{q}) = (2\pi\hbar)^{-3/2} e^{i\vec{k}\cdot\vec{q}} \qquad \vec{k} \stackrel{\text{def}}{=} \vec{p}/\hbar, \tag{4.2}$$

normalized so that

$$\int d^3q\, \psi^*(\vec{k}, \vec{q})\psi(\vec{k}\,', \vec{q}) = \delta^3(\vec{k} - \vec{k}\,')/\hbar^3 = \delta^3(\vec{p} - \vec{p}\,'), \tag{4.3}$$

and they form a complete set in $L^2(\mathbb{R}^3)$,

$$\int d^3p\, \psi^*(\vec{p}/\hbar, \vec{q})\psi(\vec{p}/\hbar, \vec{q}') = \delta^3(\vec{q} - \vec{q}'). \tag{4.4}$$

On the other hand, it is sometimes useful to normalize the wave functions so that

$$\int d^3q\, \psi^*(E, \widehat{k}; \vec{q})\psi(E', \widehat{k}'; \vec{q}) = \delta^2(\widehat{k}, \widehat{k}')\delta(E - E') \tag{4.5}$$

and

$$\int d\Omega_{\vec{k}} \int dE\ \psi^*(E, \widehat{k}; \vec{q})\psi(E, \widehat{k}; \vec{q}') = \delta^3(\vec{q} - \vec{q}'), \tag{4.6}$$

where $\widehat{k} \stackrel{\text{def}}{=} \vec{k}/|\vec{k}|$. In that case the constant $(2\pi\hbar)^{-3/2}$ in (4.2) has to be replaced by $(Mk)^{1/2}/[\hbar(2\pi)^{3/2}]$. These functions, of course, satisfy the free time-independent Schrödinger equation (3.1) in the configuration representation,[1]

$$\left[\nabla^2 + \frac{2ME}{\hbar^2}\right]\psi = 0. \tag{4.7}$$

Since in Cartesian coordinates $\nabla_q^2 = \frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2} + \frac{\partial^2}{\partial q_3^2}$, this partial differential equation is separable and the solution is a product of the form $e^{ik_1q_1}e^{ik_2q_2}e^{ik_3q_3} = e^{i\vec{k}\cdot\vec{q}}$, with $|\vec{k}|^2 = k_1^2 + k_2^2 + k_3^2 = \frac{2ME}{\hbar^2}$ as in (4.2).

Equation (4.7), however, is also separable in another extremely useful coordinate system, namely, spherical polar coordinates (Fig.4.1), in which

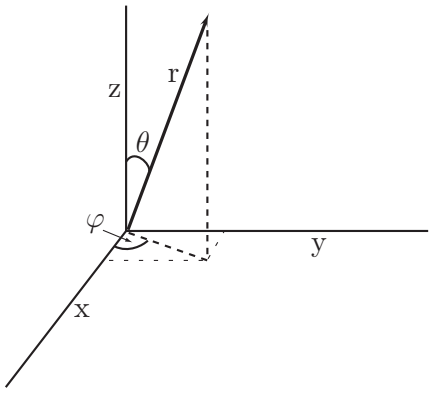

FIGURE 4.1. Spherical polar coordinates.

the Laplace operator takes the form

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} \tag{4.8}$$

[1]With $k^2 = \frac{2ME}{\hbar^2}$ this equation is also known as the Helmholtz equation.

if we define $r \stackrel{\text{def}}{=} |q|$ and θ and φ are the spherical polar angles of $\vec{q}$. Use of this Laplacian in (4.7) and the assumption that ψ can be written as a product $\psi(\vec{q}) = R(r)Y(\theta,\varphi)$ lead to

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + k^2r^2 = -\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}\right],$$

which allows us to conclude that both sides of this equation must be equal to a constant $-\lambda$, since the left-hand side is independent of θ and φ, and the right-hand side is independent of r:

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2} = \lambda Y, \tag{4.9}$$

and

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{\lambda R}{r^2} = -k^2R. \tag{4.10}$$

The solutions of (4.9) are the spherical harmonics[2] $Y_l^m(\widehat{n}) \stackrel{\text{def}}{=} Y_l^m(\theta,\varphi)$, which are discussed in more detail in Appendix D.1. The eigenvalues λ in (4.9) are $-l(l+1)$, where l is a *non-negative integer,* $l = 0, 1, 2, \ldots$, and the l^{th} eigenvalue is $(2l+1)$-fold degenerate, with m taking on $2l+1$ integer values, $-l \le m \le l$.

We are then left with the radial equation (4.10), which can be simplified by setting $R(r) \stackrel{\text{def}}{=} u(r)/r$, so that it becomes

$$-\frac{\hbar^2}{2M}\frac{d^2u}{dr^2} + \frac{l(l+1)\hbar^2}{2Mr^2}u = Eu. \tag{4.11}$$

Before discussing the solution of this equation, let us look at the physical meaning of what has been done.

For a particle of momentum $\vec{p}$ and position $\vec{q}$, the operator representing its *orbital angular momentum* with respect to the origin is given by

$$\vec{\mathbf{L}} = \vec{\mathbf{q}} \times \vec{\mathbf{p}}; \tag{4.12}$$

in the configuration representation this becomes $\vec{\mathbf{L}} = -i\hbar\vec{q} \times \nabla_q$, the z-component of which in spherical polar coordinates is

$$\vec{\mathbf{L}}_z = -i\hbar\frac{\partial}{\partial\varphi}, \tag{4.13}$$

and whose square is given by

$$\begin{aligned}\vec{\mathbf{L}}^2 &= \vec{\mathbf{L}}_x^2 + \vec{\mathbf{L}}_y^2 + \vec{\mathbf{L}}_z^2 = \left(-i\hbar\vec{q}\times\nabla_q\right)^2 \\ &= -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right],\end{aligned} \tag{4.14}$$

[2]The understanding here is that the direction of the unit vector $\widehat{n}$ is given by the polar angles θ and φ.

which commutes with $\vec{\mathbf{L}}_z$. Therefore, the spherical harmonic Y_l^m, which satisfies the equations

$$\mathbf{L}_z Y_l^m = -i\hbar \frac{\partial Y_l^m}{\partial \varphi} = m\hbar Y_l^m, \tag{4.15}$$

$$\vec{\mathbf{L}}^2 Y_l^m = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right] Y_l^m = l(l+1)\hbar^2 Y_l^m, \tag{4.16}$$

is a simultaneous eigenfunction of $\mathbf{L}_z$ and $\vec{\mathbf{L}}^2$, and m and l are the quantum numbers of the z-component and of the magnitude of the angular momentum (or, more precisely, $l(l+1)$ is the quantum number of the square of its magnitude). These eigenfunctions are normalized,

$$\int d\Omega\, Y_l^{m*}(\widehat{n}) Y_{l'}^{m'}(\widehat{n}) = \delta_{mm'}\delta_{ll'}, \tag{4.17}$$

where $d\Omega$ denotes the solid-angle element $d\cos\theta d\varphi$, and they form a complete set in the space $L^2(\widehat{n})$ of square-integrable functions on the unit sphere,

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_l^m(\widehat{n}) Y_l^{m*}(\widehat{n}') = \delta^2(\widehat{n}, \widehat{n}') \tag{4.18}$$

if we write $\delta^2(\widehat{n}, \widehat{n}')$ for the solid-angle delta-function

$$\delta^2(\widehat{n}, \widehat{n}') = \frac{\delta(\theta - \theta')}{\sin\theta} \delta(\varphi - \varphi').$$

In view of the physical meaning attached to the quantum number l, it is clear that the term $l(l+1)\hbar^2/2Mr^2$ in the radial Schrödinger equation (4.11) represents the *centrifugal potential energy*, classically given by $L^2/2Mr^2$, corresponding to the centripetal force exerted on the particle by its motion.

Let us then return to the radial Schrödinger equation (4.11) and write it in the simpler form[3]

$$-\frac{d^2u}{dr^2} + \frac{l(l+1)}{r^2} u = k^2 u, \tag{4.19}$$

[3]The requirement for l to be a non-negative integer, i.e., the quantization of the angular momentum, arises from (4.9) and the behavior of the sperical harmonics as functions of the angles (see Appendix D.1); as far as (4.11) or (4.19) are concerned, the parameter l in them could have any arbitrary value. Physically, however, these equations make sense only for integer l.

which u is to satisfy for $0 < r < \infty$ for a given $k^2 \geq 0$. In the simplest case, for $l = 0$, called the *s-wave*,[4] (4.19) becomes

$$-\frac{d^2u}{dr^2} = k^2 u, \tag{4.20}$$

an equation that has two linearly independent solutions, namely, $u = \sin kr$ and $u = \cos kr$, both of which lead to solutions of the Schrödinger equation (4.7) that are square-integrable at the origin, since $R = u/r$ and the volume element is $r^2 dr \sin\theta d\theta d\varphi$. However, because the choice of $\cos kr$ leads to $\psi \propto r^{-1}$ near $r = 0$, and $\nabla^2 r^{-1} = -4\pi\delta^3(\vec{r})$ (as you should recognize from the Poisson equation in electrostatics, where $\delta^3(\vec{r})$ is the charge density representing a unit point charge at the origin), the use of $\cos kr$ would lead to a delta-function on the right-hand side of (4.7) instead of 0. (Or, put differently, such a function would not be in the domain of the kinetic energy operator.) This solution of (4.19) for $l = 0$ is therefore not acceptable, and there is only *one* acceptable solution of the second-order differential equation (4.20).

For $l > 0$, (4.19) has a *regular singular point* at $r = 0$, and one of its solutions behaves like r^{l+1} at the origin, while the other goes like r^{-l}. Since the latter does not lead to a locally square-integrable solution of (4.7), it cannot be used, and we have, again, only one "regular solution," the *Riccati-Bessel function* $u_l(kr) \stackrel{\text{def}}{=} krj_l(kr)$, where j_l is a *spherical Bessel function* (discussed in more detail in Appendix D.1.5). The first three of these functions are given by

$$\begin{aligned} u_0(x) &= \sin x \\ u_1(x) &= -\cos x + x^{-1}\sin x \\ u_2(x) &= -3x^{-1}\cos x + (3x^{-2} - 1)\sin x. \end{aligned} \tag{4.21}$$

The result is that the general acceptable solution of (4.7) is any linear combination of functions of the form $j_l(kr)Y_l^m(\theta,\varphi)$. That this does not contradict our previous plane-wave solution (4.2) is demonstrated by the general expansion

$$\begin{aligned} e^{i\vec{k}\cdot\vec{r}} &= \frac{4\pi}{kr}\sum_{lm} i^l u_l(kr) Y_l^m(\widehat{r}) Y_l^{m*}(\widehat{k}) \\ &= 4\pi\sum_{lm} i^l j_l(kr) Y_l^m(\widehat{r}) Y_l^{m*}(\widehat{k}), \end{aligned} \tag{4.22}$$

where we have written $\widehat{r}$ and $\widehat{k}$ for unit vectors in the direction of $\vec{r}$ and $\vec{k}$, respectively, and $k = |\vec{k}|$, $r = |\vec{r}|$. The infinite degeneracy of the eigenvalue

[4]The terminology, s-wave for $l = 0$, p-wave for $l = 1$, d-wave for $l = 2$, is of spectroscopic origin, s standing for the *sharp* series of lines, p for *principal,* and d for *diffuse.*

k^2 of (4.7) now manifests itself in the fact that the integers l and m can take on infinitely many values. Note that the only solution of the rotationally invariant equation (4.7) that is spherically symmetric is the s-wave, $\psi = \sin kr/r$. For all other solutions, the degeneracy allows the rotational symmetry to be spontaneously broken.

4.1.1 The propagator

Before leaving the Schrödinger equation for a free particle, let us look at the time-dependence of the solutions of the full Schrödinger equation corresponding to solutions such as (4.2), namely, wave packets of the form

$$\psi(\vec{q},t) = (2\pi)^{-3/2} \int d^3k\, f(\vec{k}) e^{i\vec{k}\cdot\vec{q} - i\hbar k^2 t/2M}, \tag{4.23}$$

where $f(\vec{k})$ is an arbitrary square-integrable function. In order to solve the initial-value problem, in which $\psi(\vec{q},0)$ is given, we would determine f by inverting the Fourier transform (4.23) for $t = 0$,

$$f(\vec{k}) = (2\pi)^{-3/2} \int d^3q\, \psi(\vec{q},0) e^{-i\vec{k}\cdot\vec{q}}. \tag{4.24}$$

We can obtain a direct expression for the solution of the time-dependent Schrödinger equation for $t > 0$ in terms of its initial values by inserting (4.24) in (4.23):

$$\psi(\vec{q},t) = \int d^3q'\, \mathcal{G}_0^+(\vec{q},\vec{q}',t)\psi(\vec{q}',0), \qquad t > 0, \tag{4.25}$$

where for $t > 0$ the time-dependent Green's function or *propagator* $\mathcal{G}_0^+$ is given by

$$\mathcal{G}_0^+(\vec{q},\vec{q}',t) = \frac{1}{(2\pi)^3} \int d^3k\, \exp\left[i\vec{k}\cdot(\vec{q}-\vec{q}') - it\frac{\hbar\vec{k}^2}{2M} \right], \tag{4.26}$$

while for $t < 0$ we define $\mathcal{G}_0^+$ to be equal to zero. The integral in (4.26) can be explicitly evaluated[5] as

$$\mathcal{G}_0^+(\vec{q},\vec{q}',t) = \left(\frac{M}{2\pi\hbar t}\right)^{3/2} \exp\left(\frac{iM|\vec{q}-\vec{q}'|^2}{2\hbar t} - \frac{3\pi i}{4} \right), \qquad t > 0. \tag{4.27}$$

Note that this function becomes highly singular at $t \to 0$, and it vanishes uniformly like $t^{-3/2}$ as $|t| \to \infty$, so that there can be no finite spatial region

[5] **Do it.**

Ω in which the particle can remain localized[6] forever. (This is, of course, quite in agreement with the uncertainty principle, according to which the particle's localization in a region of diameter Δq implies that its momentum is uncertain by at least $\hbar/\Delta q$; therefore, it cannot be expected to remain localized there for longer than the time T it takes to traverse the region of diameter Δq with a velocity equal to $\Delta p/M$, which is $T = M\Delta q/\Delta p \leq M(\Delta q)^2/\hbar$.) Equations (4.25) and (4.27) tell us specifically that, assuming $\psi(\vec{q}, 0)$ to be integrable, as $t \to \infty$,

$$\int_\Omega d^3q\, |\psi(\vec{q}, t)|^2 = O(t^{-3}). \tag{4.28}$$

For $t \neq 0$, the propagator evidently satisfies the Schrödinger equation,

$$\left(\frac{\hbar^2}{2M}\nabla^2 + i\hbar\frac{\partial}{\partial t}\right)\mathcal{G}_0^+(\vec{q}, \vec{q}', t) = 0,$$

and for $t \to 0+$ (4.26) shows that[7]

$$\mathcal{G}_0^+(\vec{q}, \vec{q}', 0+) = \delta^3(\vec{q} - \vec{q}').$$

Since $\mathcal{G}_0^+(\vec{q}, \vec{q}', 0-) = 0$, so that $\mathcal{G}_0^+$ has a step-discontinuity as a function of t at $t = 0$, it follows that the equation satisfied by $\mathcal{G}_0^+$ is[8]

$$\left(\frac{\hbar^2}{2M}\nabla^2 + i\hbar\frac{\partial}{\partial t}\right)\mathcal{G}_0^+(\vec{q}, \vec{q}', t) = i\hbar\delta(t)\delta^3(\vec{q} - \vec{q}'). \tag{4.29}$$

In a similar fashion we can define the back-propagator $\mathcal{G}_0^-(\vec{q}, \vec{q}', t)$, which satisfies the same equation (4.29), but which vanishes for $t > 0$, while for negative times it is given by

$$\mathcal{G}_0^-(\vec{q}, \vec{q}', t) = -\left(\frac{M}{2\pi\hbar|t|}\right)^{3/2} \exp\left(\frac{iM|\vec{q} - \vec{q}'|^2}{2\hbar t} + \frac{3\pi i}{4}\right), \qquad t < 0. \tag{4.30}$$

It allows us to express $\psi_0(\vec{q}, t)$ for negative times in terms of its value at $t = 0$,

$$\psi(\vec{q}, t) = -\int d^3q'\, \mathcal{G}_0^-(\vec{q}, \vec{q}', t)\psi(\vec{q}', 0), \qquad t < 0. \tag{4.31}$$

---

[6] *Localized* here simply means that there is a nonzero probability for the particle to be found in a given finite region.

[7] Remember that the Dirac delta function has the representation $\delta(\vec{r} - \vec{r}') = \frac{1}{(2\pi)^3}\int d^3k e^{i\vec{k}\cdot(\vec{r}-\vec{r}')}$, the three-dimensional analogue of (B.20), which simply expresses the completeness of the exponentials in the sense of Fourier integrals. Also note that $t \to 0+$ means that t tends to zero from above.

[8] Integrate this equation over t from $-\epsilon$ to $+\epsilon$ to see that it gives the correct result.

4.1.2 Two particles

The kinetic-energy operator for two particles of masses M_1 and M_2 is given by

$$\mathbf{H}_0 = \frac{\vec{\mathbf{p}}_1^2}{2M_1} + \frac{\vec{\mathbf{p}}_2^2}{M_2} = -\frac{\hbar^2}{2M_1}\nabla_1^2 - \frac{\hbar^2}{2M_2}\nabla_2^2.$$

Introducing the center-of-mass coordinates

$$\vec{R} \stackrel{\text{def}}{=} \frac{\vec{q}_1 M_1 + \vec{q}_2 M_2}{M_1 + M_2}$$

and the relative coordinates

$$\vec{r} \stackrel{\text{def}}{=} \vec{q}_2 - \vec{q}_1,$$

we find that[9]

$$\mathbf{H}_0 = \frac{\vec{\mathbf{P}}^2}{2M} + \frac{\vec{\mathbf{p}}^2}{2\mu} = -\frac{\hbar^2}{2M}\nabla_R^2 - \frac{\hbar^2}{2\mu}\nabla_r^2,$$

where $\vec{\mathbf{P}} \stackrel{\text{def}}{=} \vec{\mathbf{p}}_1 + \vec{\mathbf{p}}_2$ is the operator of the total momentum, $\vec{\mathbf{p}} \stackrel{\text{def}}{=} \vec{\mathbf{p}}_2 - \vec{\mathbf{p}}_1$ that of the relative momentum, $M \stackrel{\text{def}}{=} M_1 + M_2$, and $\mu \stackrel{\text{def}}{=} M_1 M_2/(M_1 + M_2)$. The Schrödinger equation for two free particles

$$\left[-\frac{\hbar^2}{2M_1}\nabla_1^2 - \frac{\hbar^2}{2M_2}\nabla_2^2\right]\psi(\vec{q}_1, \vec{q}_2) = E\psi(\vec{q}_1, \vec{q}_2)$$

can then be separated by writing $\psi(\vec{q}_1, \vec{q}_2) = \psi_1(\vec{R})\psi_2(\vec{r})$, and the two functions ψ_1 and ψ_2 have to satisfy the equations

$$-\frac{\hbar^2}{2M}\nabla_R^2\psi_1 = E_1\psi_1, \qquad -\frac{\hbar^2}{2\mu}\nabla_r^2\psi_2 = E_2\psi_2, \qquad E_1 + E_2 = E. \quad (4.32)$$

Thus the two-particle equation is reduced to two one-particle equations and the Hilbert space $\mathcal{H} = \mathcal{H}^{\mathrm{I}} \bigotimes \mathcal{H}^{\mathrm{II}}$ is decomposed into the tensor product of one for the center-of-mass motion and another for the relative motion: $\mathcal{H} = \mathcal{H}^{\mathrm{CM}} \bigotimes \mathcal{H}^{\mathrm{rel}}$.

If the only forces on two particles are those exerted by them upon one another, the potential energy is a function of the distance between them, $V(\vec{q}_1, \vec{q}_2) = V(\vec{q}_2 - \vec{q}_1) = V(\vec{r})$, and the two-particle Schrödinger equation can be separated as in the free case. The center-of-mass wave function then satisfies the free Schrödinger equation as in (4.32), and the relative motion is described by the equation

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(\vec{r})\right]\psi = E\psi, \quad (4.33)$$

[9] **Prove this.**

which differs from the one-particle equation only by containing the reduced mass μ in place of the mass M. Two-particle systems are therefore described by the same tools and the same equations as one-particle systems, and our further discussions of the latter are equally applicable to the former.

4.2 Potentials with Spherical Symmerty

Consider now the Schrödinger equation for a particle in three dimensions in the configuration representation,

$$-\frac{\hbar^2}{2M}\nabla^2\psi + V\psi = E\psi \tag{4.34}$$

and suppose that the potential V is a function of $r = |\vec{q}|$ only, which makes it invariant under rotations. In that case, (4.34) is separable like (4.7) in spherical polar coordinates and the Laplacian can be expressed as in (4.8). With the factorization

$$\psi = R(r)Y(\theta, \varphi),$$

the angular equation is the same as (4.9), which leads to $\lambda = -l(l+1)$, with $l = 0, 1, 2, \ldots$, and spherical harmonics Y_l^m as solutions, and the radial equation is given by

$$-\frac{\hbar^2}{2M}\left[\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr}R_l - \frac{l(l+1)}{r^2}R_l\right] + V(r)R_l = ER_l. \tag{4.35}$$

Note that the quantum number m does not appear, so that $R_l(r)$ is independent of m. The definition $\psi_l(r) \stackrel{\text{def}}{=} rR_l(r)$ casts (4.35) into the simpler form

$$-\frac{\hbar^2}{2M}\psi_l'' + \left[\frac{l(l+1)\hbar^2}{2Mr^2} + V(r)\right]\psi_l = E\psi_l,$$

or, multiplying through by $2M/\hbar^2$

$$-\psi_l'' + \left[\frac{l(l+1)}{r^2} + \frac{2M}{\hbar^2}V(r)\right]\psi_l = k^2\psi_l. \tag{4.36}$$

This is the equation that has to be solved for each specific radial potential $V(r)$.

Just as in the case of a particle with one degree of freedom, if $V \to 0$ as $r \to \infty$ we will be looking for negative-energy solutions, i.e., with $k^2 \leq 0$, that are square-integrable — the bound states of the system, if there are any — and positive-energy solutions with $k^2 > 0$ that are bounded but not square-integrable: all positive values of E are in the continuous spectrum of the Hamiltonian. The physical reason is that for $E > 0$ the particle has sufficient energy to escape to infinity, which it lacks when

$E < 0$. The bound-state solutions $\psi_l(r)$ of (4.36) can always be chosen to be real, and, in addition, the angle functions $Y(\theta, \varphi)$ may also be taken to be the real *zonal harmonics*, defined by (D.34) in Appendix D. The resulting bound-state eigenfunctions of the three-dimensional Schrödinger equation, $\psi_l(r)r^{-1}Z_l^{cm}(\theta, \varphi)$ and $\psi_l(r)r^{-1}Z_l^{sm}(\theta, \varphi)$, are then real, and their nodal surfaces are concentric spheres about the origin, cones about the z-axis with apex at the origin, and planes through the z-axis. The choice of the direction of the z-axis, of course, is arbitrary, but every such choice yields a complete set of linearly independent eigenfunctions with those nodal surfaces.

In the absence of a potential, we know that the positive-energy solutions are the Riccati-Bessel functions, $u_l(kr) \stackrel{\text{def}}{=} krj_l(kr)$, whose asymptotic behavior for $kr \gg 1$ is given by (D.41),

$$u_l(kr) \sim \sin(kr - \frac{1}{2}\pi l) = \frac{i}{2}e^{i\frac{1}{2}\pi l}[e^{-ikr} - (-1)^l e^{ikr}],$$

which shows that the ratio of the outgoing amplitude to the incoming amplitude is simply $(-1)^{l+1}$. This ratio will change when there is a potential present, and the change has observable physical effects, as we shall discuss.

4.2.1 The three-dimensional square well

Assume that the potential $V(r)$ has the constant negative value $-V_0$ for $r < b$ and $V = 0$ for $r > b$. Let us first look for the bound-state solutions, for which $E < 0$ and hence $k^2 = -|k|^2 = 2ME/\hbar^2$. In the region $r < b$ we have to pick the regular solution $u_l(\kappa r)$ of (4.19), where $\kappa = \sqrt{2MV_0/\hbar^2 - |k|^2} = \sqrt{2M(E + V_0)}/\hbar$, while for $r > b$ the solution we want is $w_l^{(+)}(i|k|r)$, which is defined in Appendix D.1.5 and which asymptotically decreases as $e^{ikr} = e^{-|k|r}$ [see (D.41)]. Matching the logarithmic derivatives of the two solutions at $r = b$ leads to the equation

$$k\frac{w_l^{(+)\prime}(kb)}{w_l^{(+)}(kb)} = \kappa\frac{u_l'(\kappa b)}{u_l(\kappa b)}\,, \tag{4.37}$$

which for $l = 0$ becomes

$$-|k| = -\sqrt{2MV_0/\hbar^2 - \kappa^2} = \kappa \cot \kappa b,$$

the same equation as (3.16) for the odd solution in the one-dimensional case.[10] (Note that b here corresponds to $a/2$.) In order for this equation to have a solution it is necessary that $\cot \kappa b \leq 0$, which means we need $\sqrt{2MV_0}\, b \geq \pi\hbar/2$; the first bound state will appear, with zero binding energy, when $\sqrt{2MV_0}\, b = \pi\hbar/2$, the second s-wave bound state when

[10] **Why do we get the same equations in these cases?**

$\sqrt{2MV_0}\,b = 3\pi\hbar/2$, etc. In contrast to the one-dimensional case, a non-positive potential does not necessarily lead to a bound state in three dimensions; in order to bind it has to have a certain minimal strength.

The ground state of a spherically symmetric potential is *always an s-sate;* in other words, there can be no state of $l > 0$ with a lower energy than the state of lowest energy with $l = 0$. This is a general theorem, the obvious physical reason for which is the centrifugal repulsion in states of higher angular momentum.

The theorem can be proved by a simple application of what is known as the **Hellmann-Feynman formula:**

Let the Hermitian operator $\mathbf{H}$ *depend on a parameter* α *and let* E *be an eigenvalue of* $\mathbf{H}$*; then*

$$\frac{dE}{d\alpha} = \langle \frac{d\mathbf{H}}{d\alpha} \rangle, \tag{4.38}$$

where the expectation value is taken in the eigenstate of $\mathbf{H}$ *with the eigenvalue* E.

The proof of (4.38) is a simple exercise,[11] using the fact that

$$E = \langle \mathbf{H} \rangle = (\Psi_E, \mathbf{H}\Psi_E)/(\Psi_E, \Psi_E)$$

and differentiating it with respect to α. (All the terms coming from the differentiation of Ψ_E cancel.)

In the case at hand, the operator $\mathbf{H}$ is the radial Hamiltonian

$$\mathbf{H}_l \stackrel{\text{def}}{=} -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + \frac{2M}{\hbar^2}V(r),$$

and the place of α is taken by l, which, in the radial Schrödinger equation, may be regarded as a free parameter, forced to take on only integral values by the angular equation. Now, $d\mathbf{H}_l/dl = (2l+1)/r^2$, so that $\langle d\mathbf{H}_l/dl \rangle > 0$, which therefore implies that the eigenvalue E_l of $\mathbf{H}_l$ is a *monotonely increasing function of* l*:*

$$\frac{dE_l}{dl} > 0.$$

So if there is a bound state for $l > 0$, then changing the l-value continuously down to $l = 0$ necessarily leads to an s-wave bound state of lower energy. □

For positive energies there are no bound states.[12] Just as for a free particle, the positive energies constitute the continuous spectrum of the Hamiltonian, and for reasonable potentials there are no bound states embedded in the continuum. (Recall the physical reason for this: a particle with a positive energy is free to escape to infinity and will thus not remain forever confined to a finite region, except under very special circumstances.) However, whereas the wave function inside the well has to be a multiple of the regular Riccati-Bessel function u_l, the wave function outside the well

---

[11] **Do it.**

[12] If we are dealing with the relative motion of a two-particle system, the total energy, including the kinetic energy of the center-of-mass motion, may of course be positive and its spectrum will, in fact, form a continuum. Remember that here we are fixing our attention on the CM system.

can be a linear combination of the Riccati-Bessel function and the Riccati-Neumann function defined in Appendix D.1.5,

$$\psi_l(r) = A_l u_l(kr) + B_l v_l(kr), \qquad r > b.$$

At the surface of the potential well the logarithmic derivative of this solution has to be matched to β_l, the logarithmic derivative of the inside solution,

$$k\frac{A_l u_l'(bk) + B_l v_l'(bk)}{A_l u_l(bk) + B_l v_l(bk)} = \beta_l, \tag{4.39}$$

which fixes the ratio of the constants A_l and B_l. Inside the well, for $r < b$, the solution of the Schrödinger equation can only be a multiple of the regular $u_l(\kappa r)$, with $\kappa = \sqrt{2MV_0/\hbar^2 + k^2}$, so that we have

$$\beta_l = \kappa\frac{u_l'(b\kappa)}{u_l(b\kappa)}. \tag{4.40}$$

Just as in the one-dimensional case, in which the observed quantitites are the reflection and transmission coefficients, the experimental observations are going to take place at large distances outside the well. When $kr \gg 1$, according to (D.41), the asymptotic form of the wave function is

$$\begin{aligned}\psi_l(r) &\sim A_l \sin(kr - \frac{1}{2}\pi l) - B_l \cos(kr - \frac{1}{2}\pi l)\\ &= \frac{A_l}{\cos\delta_l}\sin(kr - \frac{1}{2}\pi l + \delta_l)\\ &= \frac{iA_l}{2\cos\delta_l}e^{\frac{1}{2}i\pi l - i\delta_l}[e^{-ikr} - (-1)^l e^{ikr}e^{2i\delta_l}],\end{aligned} \tag{4.41}$$

where the *phase shift* δ_l is defined, *modulo* π, by

$$\tan\delta_l = -\frac{B_l}{A_l}.$$

Consequently, the principal long-distance effect of the potential on the wave function for an individual angular momentum is a shift in its phase, and therefore a change in the ratio of the outgoing to the incoming spherical wave amplitudes by a factor of $e^{2i\delta_l}$. Using (4.39), we can calculate this phase shift, obtaining[13]

$$\tan\delta_l = \frac{\beta_l u_l(bk) - k u_l'(bk)}{\beta_l v_l(bk) - k v_l'(bk)}, \tag{4.42}$$

which for the s-wave case becomes

$$\tan\delta_0 = \tan bk\,\frac{bk\cot bk - b\beta_0(k)}{bk\tan bk + b\beta_0(k)}, \tag{4.43}$$

[13] **Do it.**

and for the square well $b\beta_0(k) = b\kappa \cot b\kappa$. The values (D.39) and (D.40) of the Riccati-Bessel and Riccati-Neumann function, used in (4.42), lead to[14] the generic low-energy behavior of the phase shifts, when $k \ll 1/b$

$$\delta_l \sim \text{const.} \times k^{2l+1}, \tag{4.44}$$

which implies the important fact, to which we shall return, that when k is small enough, the s-wave phase shift dominates the others; as the energy rises, the p-wave phase shift begins to contribute, etc.

4.2.2 The scattering amplitude

The principal observable effect of a potential on a positive-energy particle approaching a region with potential energy is a deflection from its initial direction. The requirement that the particle's momentum in the remote past (when it is assumed to have been under our control by means of an accelerator or some other experimental device) was $\vec{p} = \hbar\vec{k}$ is translated into a boundary condition on the physically desired solution $\psi^{(+)}$ of the time-independent Schrödinger equation by subjecting the deviation of $\psi^{(+)}$ from the plane wave $ce^{i\vec{k}\cdot\vec{r}}$,

$$\phi_{\text{scatt}}(\vec{k},\vec{r}) \stackrel{\text{def}}{=} \psi^{(+)}(\vec{k},\vec{r}) - ce^{i\vec{k}\cdot\vec{r}}$$

to the radiation condition,

$$\phi_{\text{scatt}} - ik\frac{\partial}{\partial r}\phi_{\text{scatt}} \to 0 \qquad \text{as} \qquad r \stackrel{\text{def}}{=} |\vec{r}| \to \infty.$$

This implies there exists a function $A(\vec{k}',\vec{k})$ that depends on the initial momentum $\hbar\vec{k}$ and the final momentum $\hbar\vec{k}' \stackrel{\text{def}}{=} \hbar|\vec{k}|\hat{r} = \hbar|\vec{k}|\vec{r}/r$ pointing from the center to the observation point $\vec{r}$, such that as $r \to \infty$,[15]

$$\psi^{(+)}(\vec{k},\vec{r}) = c\left[e^{i\vec{k}\cdot\vec{r}} + A(\vec{k}',\vec{k})\frac{1}{r}e^{ikr}\right] + o(r^{-1}). \tag{4.45}$$

The constant c depends on the desired normalization. Equation (4.2) tells us that if $\psi^{(+)}$ is to be normalized as in (4.3) and (4.4), we must choose $c = (2\pi\hbar)^{-3/2}$; on the other hand, if we want the normalization (4.5) and (4.6), we must take $c = (Mk)^{1/2}/[\hbar(2\pi)^{3/2}]$.

By the same stationary-phase argument used earlier to connect the Fourier-transformed functions to the more physical time-dependent wave functions it follows that in the distant past and at large distances it will be only the plane-wave term that contributes, while in the far future there will also be

[14] **Show it.**

[15] $o(r^{-1})$ means that the remainder tends to zero faster than r^{-1}.

a contribution from the outgoing spherical wave, with the amplitude A, which is called the *scattering amplitude.*[16] The ratio of this outgoing flux in the direction $\vec{k}'$ to the incoming flux in the direction $\vec{k}$ is given by[17]

$$\frac{d\sigma}{d\Omega'} = |A(\vec{k}', \vec{k})|^2, \tag{4.46}$$

which is called the *differential scattering cross section.* (The letter Ω' here denotes the solid angle subtended by $\vec{k}'$.) Thus $(d\sigma/d\Omega')d\Omega' = |A(\vec{k}', \vec{k})|^2 d\Omega'$ is the probability of detecting a particle (whose initial momentum was $\hbar\vec{k}$) scattered and emerging at large distance $r \gg 1/k$ in the element of solid angle $d\Omega'$ around $\vec{k}'$.

In order to see in more detail the magnitude of the incoming spherical wave that the plane wave contributes asymptotically, we have to use (4.22) and the asymptotic values (D.41) of the spherical Bessel functions; as $r \to \infty$

$$e^{i\vec{k}\cdot\vec{r}} = \frac{2\pi i}{kr} \sum_{lm} (-1)^l [e^{-ikr} - (-1)^l e^{ikr}] Y_l^m(\hat{r}) Y_l^{m*}(\hat{k}) + o((kr)^{-1}). \tag{4.47}$$

[Using (D.30) and (D.28), Eq. (4.47) can be expressed as

$$e^{i\vec{k}\cdot\vec{r}} \sim (2\pi/ikr)[e^{ikr}\delta(\hat{k}, \hat{r}) - e^{-ikr}\delta(\hat{k}, -\hat{r})].]$$

Therefore, if $\psi^{(+)}$ is to be of the form (4.45), and[18]

$$\psi^{(+)}(\vec{k}, \vec{r}) = \frac{4\pi}{kr} \sum_{lm} i^l \psi_l(k, r) Y_l^m(\hat{k}) Y_l^{m*}(\hat{r}), \tag{4.48}$$

we have to choose the normalization of ψ_l so that its asymptotically incoming-wave part has the same amplitude as that of the plane wave; this yields the factor $2\pi ic(-1)^l/kr$. Its outgoing-wave part, according to (4.41), then has the amplitude $-2\pi ice^{i\delta_l}/kr$. Adding and subtracting the outgoing wave asymptotically needed to make up the plane wave part, we then obtain

$$\begin{aligned} \psi^{(+)} \quad \sim \quad & \frac{2c\pi i}{kr} \sum_{lm} (-1)^l [e^{-ikr} - (-1)^l e^{ikr}] Y_l^m(\hat{r}) Y_l^{m*}(\hat{k}) \\ & + e^{ikr} \frac{2c\pi i}{kr} \sum_{lm} (1 - e^{2i\delta_l}) Y_l^m(\hat{r}) Y_l^{m*}(\hat{k}), \end{aligned}$$

---

[16] **As an exercise, flesh this out, making the stationary-phase argument explicit.**

[17] Since the kinetic energy at $t \to \infty$ is equal to its initial value at $t \to -\infty$, the magnitudes of the initial and final velocities are equal, $\hbar|\vec{k}'|/M = \hbar|\vec{k}|/M$, and the flux ratio equals the ratio of the squares of the amplitudes. In a more general setting, when inelastic scattering is possible, this is not so, and there appears a factor of $|\vec{k}'|/|\vec{k}|$ on the right-hand side.

[18] Because the spherical symmetry of $V(r)$ implies that $\psi^{(+)}(\vec{k}, \vec{r})$ must be rotationally invariant, the l and m-values of the spherical harmonics in the sum must be equal and the coefficients ψ_l are independent of m; see Appendix D.

from which we conclude that the scattering amplitude has the partial-wave expansion

$$\begin{aligned} A(\vec{k}',\vec{k}) &= \frac{2\pi i}{k}\sum_{lm}(1-e^{2i\delta_l})Y_l^m(\widehat{k}')Y_l^{m*}(\widehat{k}) \qquad (4.49)\\ &= \frac{1}{2ik}\sum_{l}(2l+1)(e^{2i\delta_l}-1)P_l(\cos\theta)\\ &= \frac{1}{k}\sum_{l}(2l+1)e^{i\delta_l}\sin\delta_l P_l(\cos\theta) \stackrel{\text{def}}{=} f(k,\theta), \qquad (4.50)\end{aligned}$$

where $\cos\theta \stackrel{\text{def}}{=} \widehat{k}\cdot\widehat{k}'$ and Eq. (D.32) has been used. (That the scattering amplitude does not depend on the direction of the incident particle but only on the angle θ between the initial and final directions is a consequence of the rotational invariance of the potential.)

Thus what for each individual partial wave appears as a simple shift in phase manifests itself, via interference effects, in the total asymptotic flux as an observable angle-dependent variation: particles sent, for example, by an accelerator, toward the center with the momentum $\hbar\vec{k}$ are detected by counters, positioned at various angles, with an angular distribution given by the differential cross section (4.46).

The total number of particles scattered, per unit time and per unit incident flux, is called the *total scattering cross section.* It is easily calculated from the partial-wave expansion (4.49) by means of Eqs. (D.29) and (D.32),

$$\sigma_{\text{total}} \stackrel{\text{def}}{=} \int d\Omega' \frac{d\sigma(\vec{k}',\vec{k})}{d\Omega'} = \int d\Omega' |A(\vec{k}',\vec{k})|^2 == \frac{4\pi}{k^2}\sum_l (2l+l)\sin^2\delta_l. \qquad (4.51)$$

More generally, we obtain in the same manner[19]

$$\begin{aligned} &\int d\Omega''\, A(\vec{k}'',\vec{k}')A^*(\vec{k}'',\vec{k}) = \left(\frac{2\pi}{k^2}\right)\\ &\sum_{ll'mm'}(e^{2i\delta_l}-1)(e^{-2i\delta_{l'}}-1)\int d\Omega''\, Y_l^m(\widehat{k}'')Y_l^{m*}(\widehat{k}')Y_{l'}^{m'*}(\widehat{k}'')Y_{l'}^{m'}(\widehat{k})\\ = \;& \frac{8\pi^2}{k^2}\sum_{lm}(1-\cos 2\delta_l)Y_l^m(\widehat{k})Y_l^{m*}(\widehat{k}')\\ = \;& \left(\frac{4\pi}{k}\right)^2\sum_{lm}\sin^2\delta_l Y_l^m(\widehat{k})Y_l^{m*}(\widehat{k}') = \frac{4\pi}{k^2}\sum_l(2l+1)\sin^2\delta_l\, P_l(\widehat{k}\cdot\widehat{k}')\\ = \;& \frac{4\pi}{k}\Im A(\vec{k}',\vec{k}). \qquad (4.52)\end{aligned}$$

[19] $\Im A$ means the imaginary part of the complex number A.

The special case in which $\vec{k} = \vec{k}'$ is of particular interest:

$$\sigma_{\text{total}}(k) = \frac{4\pi}{k}\Im A(\vec{k}, \vec{k}) = \frac{4\pi}{k}\Im f(k, 0), \tag{4.53}$$

which is known as the *optical theorem.* Its physical basis is the fact that all the flux in the scattered waves originates from interference between the forward scattering and the incident beam.[20] The formula (4.52) is called the *generalized optical theorem.* In terms of the dimensionless quantity

$$F(k, \cos\theta) \stackrel{\text{def}}{=} kA(\vec{k}', \vec{k})$$

it may be written in the simple form

$$\Im F(k, 1) = \frac{1}{2}\int_{-1}^{1} dx\, |F(k, x)|^2, \tag{4.54}$$

where $x \stackrel{\text{def}}{=} \cos\theta$. Note that the optical theorem immediately implies a lower bound on the forward scattering cross section if the total cross section is given:

$$\frac{d\sigma}{d\Omega}(\vec{k}, \vec{k}) \geq \left(\frac{k}{4\pi}\sigma_{\text{total}}\right)^2.$$

The partial-wave expansions of the scattering amplitude and the cross section are particularly useful at low energies. This is because, as (4.44) shows in the case of a square well but which holds more generally, the l^{th} phase shift goes as k^{2l+1} when k is small, and consequently only the first few terms contribute significantly to the series. Generically, the differential cross section becomes isotropic in the low-energy limit, and for a square well of radius b, $\sigma_{\text{total}} \to 4\pi b^2$. Note that this is *four times the geometric cross section.* In general, the tangent of the s-wave phase shift will go like

$$\tan\delta_0 \sim -ck \qquad \text{as} \qquad k \to 0, \tag{4.55}$$

where the constant c is known as the *scattering length,* and the zero-energy limit of the total cross section is $4\pi c^2$.

There are, however, a variety of more interesting phenomena that can occur at low energies, which we can see by studying Eq. (4.43). Suppose that at some low energy, when $bk \ll 1$, the numerator of the right-hand side of (4.43) is equal to zero, which will be the case when $1 - \frac{1}{3}(bk)^3 = b\beta_0$. All the other phase shifts are then still negligible; and since $\sin\delta_0 = 0$, the s-wave contribution vanishes so that the cross section will be anomalously small. This is the explanation of what is known as the *Ramsauer-Townsend effect,* an unusually small observed scattering cross section of low-energy

[20] For a discussion of the circuitous history of the optical theorem see [Newton 76].

(~ 0.7 eV) electrons on rare-gas atoms. It is analogous to the transparency of a one-dimensional potential well at certain energies. In three dimensions, of course, if the numerator of (4.43) vanishes when $bk \ll 1$, there need be no observable effect because the other partial waves obscure the vanishing of the s-wave.

Another possibility is that the denominator of the right-hand side of (4.43) vanishes. In that case the s-wave phase shift rises through $\frac{1}{2}\pi$ and hence $\sin\delta_0 = 1$: the s-wave scattering amplitude is *maximal.* If that happens at an energy $E_0 = \hbar^2 k_0^2/2M$ when $bk_0 \ll 1$, all other phase shifts will be negligible and the cross section will have a peak of $\sigma = 4\pi/k_0^2 = \lambda_0^2/\pi$: there is a *resonance.* An analogous phenomenon occurs when the denominator of (4.42) vanishes for some $l > 0$, the condition for which is that $kv_l'(bk)/v_l(bk) = \beta_l$. If such an l-wave resonance occurs at a low energy E_0, when $bk_0 \ll 1$, it is particularly striking, because then that particular partial wave completely dominates and the differential cross section has the angular shape of the square of the Legendre polynomial of order l:

$$\frac{d\sigma}{d\Omega} \simeq k_0^{-2}(2l+1)^2 P_l^2(\cos\theta). \tag{4.56}$$

In the vicinity of an l-wave resonance, the tangent of the l^{th} phase shift will generally behave like

$$\tan\delta_l \sim \frac{a}{1-(k/k_0)}, \qquad \text{where} \qquad a > 0. \tag{4.57}$$

This implies that the cross section, which is proportional to $\sin^2\delta_l$ if the other partial waves can be neglected, goes like

$$\frac{d\sigma}{d\Omega} \simeq k_0^{-2}(2l+1)^2 P_l^2(\cos\theta)\frac{a^2}{a^2+[1-(k/k_0)]^2}, \tag{4.58}$$

a function that has the bell shape shown in Figure 4.2, generally known as a *Breit-Wigner* resonance curve.[21] It has its maximimum at $k = k_0$ and decreases to half its maximal value when $k - k_0 = \pm ak_0$; therefore the *half-width at half-maximum* of the Breit-Wigner curve as a function of k is ak_0, and as a function of the energy $E = \hbar^2 k^2/2M$ (if narrow, in the sense that $a \ll 1$) it is

$$\frac{1}{2}\Gamma = \frac{\hbar^2 k_0^2 a}{M}. \tag{4.59}$$

Furthermore, it follows from (4.57) that at the resonance, $d\delta_l/dk = 1/(ak_0)$, which, according to (2.61), means that there is a time delay τ_D in the outgoing wave,

$$\tau_D = \frac{M}{\hbar k_0}\frac{d\delta_l}{dk} = \frac{M}{\hbar a k_0^2}, \tag{4.60}$$

[21] In the context of classical electromagnetic radiation, the same shape is called *Lorentzian.*

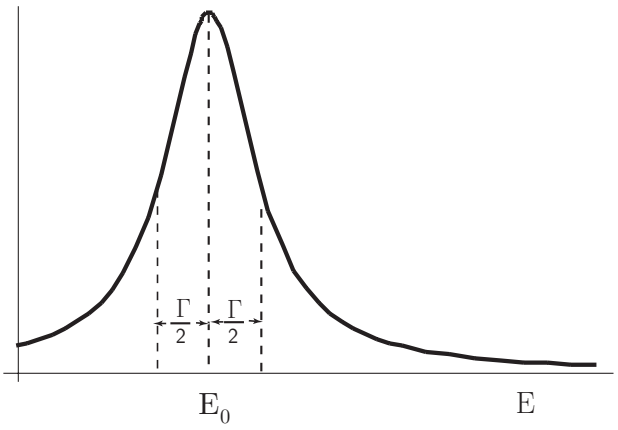

FIGURE 4.2. A Breit-Wigner resonance curve of width Γ.

indicating that the particle is temporarily caught in an unstable kind of "state." This delay is the quantum analogue of the classical phenomenon of "orbiting," in which the particle may circle the attracting center of force many times before finally escaping. Multiplying the two equations (4.60) and (4.59), we find the important relation

$$\Gamma\tau_D = 2\hbar \tag{4.61}$$

between the width Γ of a resonance curve as a function of the energy and the "lifetime" τ_D of the corresponding "state" that is temporarily brought into being. In a vague sort of way (4.61) may be thought of as a consequence of the uncertainty relation between the energy of the unstable "state" and the duration of its existence.

As (4.57) indicates, the phase shift at a sharp resonance steeply rises, increasing by π. This behavior can also be expressed in terms of the *S matrix*,

$$S_l \stackrel{\text{def}}{=} e^{2i\delta_l} = \frac{1 + i\tan\delta_l}{1 - i\tan\delta_l} \simeq \frac{k - k_0 - iak_0}{k - k_0 + iak_0}, \tag{4.62}$$

which shows that near a sharp resonance the S matrix has (or appears to have, because (4.62) may be only an approximation valid on the real axis) a simple pole in the lower half of the k-plane near the real axis. (When the potential strength is increased so as to move the resonance first to the origin and then to produce a bound state, the pole moves to the upper half plane.)

Let's see what happens when the resonance occurs very close to zero energy and the potential strength is increased so as to move the resonance to the origin, in which case the denominator in (4.42) vanishes at $k = 0$. Comparison with (4.37) shows that the critical potential strength needed to accomplish this is exactly the same as that required to introduce, for

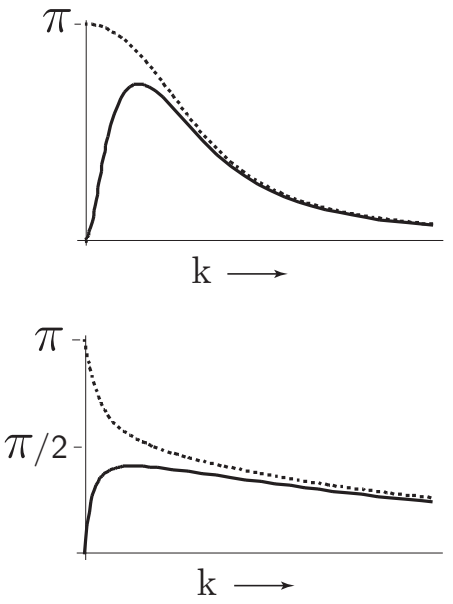

FIGURE 4.3. The upper figure shows a phase shift for $l > 0$ at low energies. The solid curve shows the phase shift if the potential is not quite strong enough to bind; the dashed curve, when there is a shallow bound state. The lower figure shows the behavior of an s-wave phase shift under similar conditions.

the same angular momentum, a new bound state with zero binding energy. According to Eqs. (D.39) and (D.40), used in (4.42), the phase shift near the orgin then goes like

$$\tan \delta_l \sim k^{2l-1},$$

implying that for $l > 0$ the phase shift still vanishes (*modulo* π) at $k = 0$, whereas for $l = 0$ it approaches $\pi/2$ (*modulo* π). (This exceptional situation is called a *zero-energy resonance* or a *half-bound state*.) Before the potential reaches its critical value, the phase shift at $k \simeq 0$ has a positive slope, so that the scattering length is negative; at the critical strength, the scattering length diverges, and it becomes positive when the potential strength is further increased to produce a bound state of negative energy. At the same time, the phase shift, which was an increasing function of k before the introduction of the bound state, *decreases* from π at $k = 0$ after the introduction of the bound state. (The low-energy behavior of the phase shifts for $l > 0$ and for $l = 0$ is shown in Figure 4.3.) Thus, every time the increasing potential strength passes a critical value to introduce a new bound state, the phase shift of the same angular momentum at zero energy discontinuously jumps by π; as a result, the zero-energy phase shift $\delta_l(0)$ (if $\delta_l(k)$ is defined so as to vanish as $E \to \infty$) equals π times the number

n_l of bound states of the same angular momentum:

$$\delta_l(0) = n_l \pi, \tag{4.63}$$

which is known as *Levinson's theorem.*[22]

There is, however, an important difference between the case of $l = 0$ and those of $l > 0$. In the former case, the phase shift never quite rises to $\pi/2$ when the potential is below critical strength, thus not causing a resonance, whereas in the cases of $l > 0$, it rises through $\pi/2$ almost to π, thus causing a resonance at a low value of the energy. The physical reason for this difference is the presence of the centrifugal barrier for $l > 0$, which manages to trap the particle and keep it near the origin for a while.

4.2.3 Transformation to the laboratory system

As we saw, the Schrödinger equation for a two-particle system in its center-of-mass reference frame, in which the two particles interact via a potential V, is identical to that for a one-particle system with V as an external potential, except for the replacement of the mass M of the particle by the reduced mass μ of the two. We therefore always treated two-particle systems in its barycentric frame. Experiments, however, are rarely performed in the center-of-mass frame; more often than not, they are done in a coordinate system, called simply the *laboratory frame*, in which one of the particles is initially at rest. We thus have to find the transformation leading from the scattering angle θ in the barycentric frame to the scattering angle ϑ in the laboratory frame.

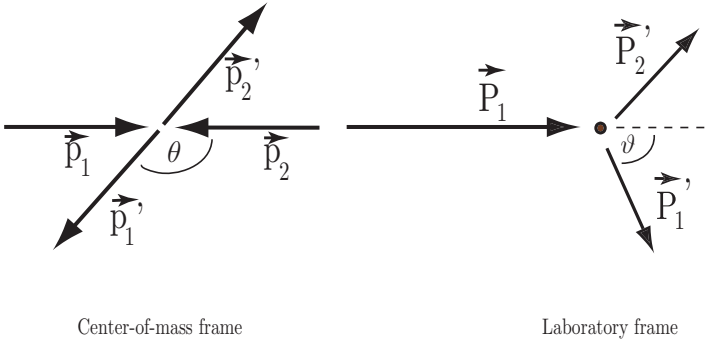

FIGURE 4.4. Initial and final particle momenta in the barycentric and in the laboratory frame.

We will denote the initial and final momenta of the particles in the lab frame by $\vec{P}_1$, $\vec{P}_2$, $\vec{P}_1'$, and $\vec{P}_2'$, respectively, and their counterparts in the

[22] In the exceptional case of $l = 0$ in which there is a zero-energy resonance, (4.63) is replaced by $\delta_0(0) = (n_0 + \frac{1}{2})\pi$.

CM frame $\vec{p}_1$, $\vec{p}_2$, $\vec{p}_1'$, and $\vec{p}_2'$ (Fig. 4.4). In the barycentric frame we have $\vec{p}_1 + \vec{p}_2 = \vec{p}_1' + \vec{p}_2' = 0$, and we define $p \stackrel{\text{def}}{=} p_1 = p_2 = p_1' = p_2'$. In the laboratory, particle #2 will be assumed to be initally at rest: $\vec{P}_2 = 0$. Therefore $0 = \vec{P}_2 = \vec{p}_2 + M_2\vec{v}_{\text{CM}}$ and hence,

$$\vec{v}_{\text{CM}} = \frac{\vec{p}_1}{M_2},$$

and it follows that

$$\vec{P}_1 = \vec{p}_1 + M_1\vec{v}_{\text{CM}} = (1+\gamma)\vec{p}_1, \qquad \gamma \stackrel{\text{def}}{=} \frac{M_1}{M_2}.$$

Similarly,

$$\vec{P}_1' = \vec{p}_1' + M_1\vec{v}_{\text{CM}} = \vec{p}_1' + \gamma\vec{p}_1,$$

which leads to

$$\vec{P}_1 \cdot \vec{P}_1' = P_1 P_1' \cos\vartheta = (1+\gamma)(\gamma + \cos\theta)p^2$$

and

$$P_1'^2 = p^2(1 + \gamma^2 + 2\gamma\cos\theta).$$

As a result we find

$$\cos^2\vartheta = \frac{(\gamma + \cos\theta)^2}{1 + \gamma^2 + 2\gamma\cos\theta}, \tag{4.64}$$

which leads to the simpler formula

$$\tan\vartheta = \frac{\sin\theta}{\gamma + \cos\theta}. \tag{4.65}$$

The cross sections in the CM system and in the lab frame are related by

$$\frac{d\sigma^{\text{LAB}}}{d\Omega} = \frac{d\sigma}{d\cos\vartheta d\varphi} = \frac{d\sigma}{d\cos\theta d\varphi}\frac{d\cos\theta}{d\cos\vartheta} = \frac{d\sigma^{\text{CM}}}{d\Omega}\frac{d\cos\theta}{d\cos\vartheta},$$

and we find from (4.64) that

$$\frac{d\cos\vartheta}{d\cos\theta} = \frac{1 + \gamma\cos\theta}{(1 + \gamma^2 + 2\gamma\cos\theta)^{3/2}},$$

so that

$$\frac{d\sigma^{\text{LAB}}}{d\Omega} = \frac{d\sigma^{\text{CM}}}{d\Omega}\frac{(1 + \gamma^2 + 2\gamma\cos\theta)^{3/2}}{1 + \gamma\cos\theta}. \tag{4.66}$$

Equation (4.65) implies that always $\sin\vartheta < 1/\sqrt{1+\gamma^2}$ and $\vartheta < \theta$. If $M_1 > M_2$, so that a heavier particle is scattered in the laboratory by a lighter one that is initially at rest, then ϑ is always less than 90^o: no particle can be turned back by a lighter one; indeed, when $M_1 \gg M_2$, the scattering in the laboratory is all concentrated near the forward direction; if the two particles have equal masses, then $\vartheta = \frac{1}{2}\theta$, so that backward scattering in the CM system becomes 90^0 scattering in the laboratory. On the other hand, if $M_1 \ll M_2$, the two frames of reference are nearly the same and the laboratory angle is almost equal to the center-of-mass angle.

4.2.4 The Coulomb potential

The hydrogen atom, the simplest atomic system in nature, provided historically the first important test for the quantum theory, initially in the form of the "old quantum theory" of Bohr and Sommerfeld, and subsequently for Schrödinger's, with later refinements by Dirac and by Feynman, Schwinger, and Tomonaga, the first owing to relativity and the second to quantum electrodynamics.

As a first step in the treatment of this two-particle problem we separate out the center-of-mass motion as in Section 4.1.2, after which the wave function of the relative coordinates of the electron with respect to the nucleus has to satisfy a one-particle Schrödinger equation with the reduced mass $\mu = Mm/(m+M)$, if we denote the masses of the electron and the nucleus, respectively, by m and M, and the electric charge of the nucleus by Ze (allowing for the possibility of $Z \neq 1$, say, in the case of a helium ion):

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi - \frac{Ze^2}{r}\psi = E\psi, \tag{4.67}$$

where the reduced mass μ of the electron differs from its actual mass by only 0.05%. The potential being rotationally invariant, we next separate out the angular dependence, writing $\psi = r^{-1}\psi_l(r)Y_l^m(\theta,\varphi)$, so that ψ_l must satisfy the radial equation

$$-\psi_l'' + \left[\frac{l(l+1)}{r^2} - \frac{2\mu Ze^2}{\hbar^2 r}\right]\psi_l = \frac{2\mu E}{\hbar^2}\psi_l.$$

The Coulomb potential has two special characteristics to be recognized immediately: it is singular as r^{-1} at the origin, and it decreases relatively slowly, as r^{-1}, at infinity. The first does not cause any serious difficulties, but the slow decrease at large distance has important consequences, as we shall see, both for the bound states and for the scattering. This physically important potential does not belong to the class (of potentials decreasing faster than r^{-2} at infinity) to which all of our previous mathematical statements are applicable.

4.2.5 The hydrogen atom

Let us first look at the bound states and assume therefore that $E < 0$. It will be convenient to multiply the radial Schrödinger equation by $\hbar^2/(8\mu|E|)$, to define the new independent variable $\rho \stackrel{\text{def}}{=} r\sqrt{8\mu|E|}/\hbar$, and to introduce the constant $\lambda \stackrel{\text{def}}{=} Ze^2\sqrt{\mu}/(\hbar\sqrt{2|E|})$, after which the radial Schrödinger equation for $f_l(\rho) \stackrel{\text{def}}{=} \psi_l(r)$ reads

$$f_l'' + \left(\frac{\lambda}{\rho} - \frac{1}{4} - \frac{l(l+1)}{\rho^2}\right) f_l = 0. \tag{4.68}$$

At this point we factor out the dominant behavior of f_l at infinity and at the origin, defining $g_l(\rho) \stackrel{\text{def}}{=} \rho^{-l-1} e^{\frac{1}{2}\rho} f_l(\rho)$, which leads to the new equation

$$\rho g_l'' + [2(l+1) - \rho] g_l' + [\lambda - (l+1)] g_l = 0. \tag{4.69}$$

Making the Frobenius *ansatz* $g_l = \rho^s \sum_0^\infty a_n \rho^n$ leads to the indicial equation[23]

$$s(s-1) + 2s(l+1) = 0,$$

whose solutions are $s = 0$ and $s = -(2l+1)$, as expected.[24] The coefficients a_n for the regular solution, i.e., the one with $s = 0$, must satisfy the recursion relation[25]

$$a_{n+1} = a_n \frac{n + l + 1 - \lambda}{(n+1)(n+2l+2)},$$

so that for the tail end of the series $a_{n+1}/a_n \sim 1/n$, like the power series for e^ρ. Therefore f_l will grow without bounds for large ρ unless the series breaks up, which will happen if and only if λ is an integer n larger than or equal to $l+1$, $\lambda = n = l + 1 + n'$, $n' = 0, 1, 2, \ldots$. Consequently we find that the bound-state energies of the Coulomb potential are given by

$$E_n = -\frac{Z^2 e^4 \mu}{2\hbar^2 n^2}, \qquad n = l+1,\, l+2,\, l+3, \ldots, \tag{4.70}$$

where n is called the *principal quantum number* and $n' \stackrel{\text{def}}{=} n - l - 1$, which takes on all non-negative integral values, is called the *radial quantum number.*

This result has three significant characteristics, the first two being consequences of the slow decrease of the potential at large distance, and the third special to the Coulomb potential:

1. There are bound states for all angular momenta.

2. For each angular momentum there are *infinitely many bound states.* The energies of these states accumulate at the origin, thus getting denser and denser there. Neither 1 nor 2 is the case for potentials that decrease faster than r^{-2} at infinity.

3. In addition to the normal m-degeneracy due to the rotational symmetry of the potential, there is also an *l-degeneracy:* the ground state, designated $1s$, is necessarily an s-wave and non-degenerate; but the first excited level, with $n = 2$, consists of two states, one with $l = 0$,

[23] **Show this.**
[24] **Why?**
[25] **Show this.**

the $2s$-state, and the other with $l = 1$, the $2p$-state; the second excited level consists of three states, with $l = 0$, $l = 1$, or $l = 2$,, the $3s$, $3p$, and $3d$-states, etc. The total number of states of the n^{th} energy level is $\sum_0^{n-1}(2l+1) = n(n-1) + n = n^2$; so the n^{th} level is n^2-fold degenerate. This is a very special property of the Coulomb potential, to the explanation of which, by means of a symmetry, we shall return in the next Subsection. We shall also see later that not even relativistic corrections serve to remove the l-degeneracy in the hydrogen atom; it takes the effects of quantum electrodynamics to do that.

The bound-state wave functions are obtained by solving (4.69), which is identical to Eq. (D.67) in Appendix D.4. The function g_l, therefore, is an *associated Laguerre polynomial* $L^{2l+1}_{n-l-1}(\rho)$, and according to (D.70), the normalized Coulomb wave functions are given by

$$\begin{aligned}\psi_{nlm}(r,\theta,\varphi) &= 2\frac{[(n-l-1)!]^{1/2}}{n^2[(n+l)!]^{3/2}}\frac{(Ze^2\mu)^{3/2}}{\hbar^3} \\ &\quad\times\left(\frac{R}{n}\right)^l e^{-R/2n}L^{2l+1}_{n-l-1}(\frac{R}{n})Y_l^m(\theta,\varphi),\end{aligned} \tag{4.71}$$

where

$$R \stackrel{\text{def}}{=} \frac{2Ze^2\mu}{\hbar^2}r.$$

Parabolic coordinates

The Schrödinger equation with a Coulomb potential can be separated not only in spherical polar coordinates but also in parabolic coordinates, and it is instructive to to do so. These coordinates are defined by

$$\xi \stackrel{\text{def}}{=} r - z, \qquad \eta \stackrel{\text{def}}{=} r + z, \qquad \varphi \stackrel{\text{def}}{=} \arctan(y/x), \tag{4.72}$$

where r is the distance from the origin, $r \stackrel{\text{def}}{=} \sqrt{x^2+y^2+z^2}$, and φ is the azimuthal angle about the z-axis. The Laplacian in these coordinates is given by[26]

$$\nabla^2 = \frac{4}{\xi+\eta}\left[\frac{\partial}{\partial\xi}\xi\frac{\partial}{\partial\xi} + \frac{\partial}{\partial\eta}\eta\frac{\partial}{\partial\eta}\right] + \frac{1}{\xi\eta}\frac{\partial^2}{\partial\varphi^2}. \tag{4.73}$$

The *ansatz* $\psi = f(\xi)g(\eta)h(\varphi)$ in the Schrödinger equation therefore leads to the equation

$$\frac{2\hbar^2}{\mu}\frac{\xi\eta}{\xi+\eta}\frac{\frac{d}{d\xi}\xi\frac{d}{d\xi}f}{f} + \frac{2\hbar^2}{\mu}\frac{\xi\eta}{\xi+\eta}\frac{\frac{d}{d\eta}\eta\frac{d}{d\eta}g}{g} + \frac{2Ze^2\xi\eta}{\xi+\eta} + E\xi\eta = -\frac{\hbar^2}{2\mu}\frac{\frac{d^2h}{d\varphi^2}}{h},$$

[26] **Prove this.**

from which we may conclude that both sides have to be equal to a constant, which will be called $\hbar^2 m^2/2\mu$, so that

$$h(\varphi) = e^{im\varphi},$$

and the requirement that h be single-valued forces m to be zero or a positive or negative integer. Multiplication by $(\frac{1}{\xi} + \frac{1}{\eta})$ then leads to

$$-\frac{2\hbar^2}{\mu}\frac{\frac{d}{d\xi}\xi\frac{d}{d\xi}f}{f} - Ze^2 - E\xi + \frac{\hbar^2}{2\mu}\frac{m^2}{\xi} = \frac{2\hbar^2}{\mu}\frac{\frac{d}{d\eta}\eta\frac{d}{d\eta}g}{g} + Ze^2 - \frac{\hbar^2}{2\mu}\frac{m^2}{\eta},$$

both sides of which must now be equal to a constant, independent of η and ξ, which we shall call β. As a consequence we obtain the following two equations:

$$\frac{d}{d\xi}\xi\frac{d}{d\xi}f + \left(\frac{Ze^2\mu}{2\hbar^2} - \frac{|E|\mu\xi}{2\hbar^2} - \frac{m^2}{4\xi} + \frac{\beta\mu}{2\hbar^2}\right)f = 0, \tag{4.74}$$

$$\frac{d}{d\eta}\eta\frac{d}{d\eta}g + \left(\frac{Ze^2\mu}{2\hbar^2} - \frac{|E|\mu\eta}{2\hbar^2} - \frac{m^2}{4\eta} - \frac{\beta\mu}{2\hbar^2}\right)g = 0, \tag{4.75}$$

both of which are of the form

$$\frac{1}{\zeta}\frac{d}{d\zeta}\zeta\frac{d}{d\zeta}F + \left(\frac{\lambda}{\zeta} - \frac{m^2}{4\zeta^2} - \frac{1}{4}\right)F = 0, \tag{4.76}$$

where $\zeta \stackrel{\text{def}}{=} (\sqrt{2\mu|E|}/\hbar)\xi$ and $\lambda \stackrel{\text{def}}{=} \frac{Ze^2+\beta}{\hbar}\sqrt{\mu/8|E|}$ for f, while $\zeta \stackrel{\text{def}}{=} (\sqrt{2\mu|E|}/\hbar)\eta$ and $\lambda \stackrel{\text{def}}{=} \frac{Ze^2-\beta}{\hbar}\sqrt{\mu/8|E|}$ for g. The Frobenius *ansatz* leads to the indicial equation $s^2 = m^2/4$,[27] whose solutions are $s = \pm m/2$. Thus the regular solution of (4.76) behaves at the origin like $\zeta^{|m|/2}$; since it is easy to see that at infinity F goes like $e^{-\zeta/2}$, we set

$$F(\zeta) \stackrel{\text{def}}{=} \zeta^{|m|/2}e^{-\zeta/2}G(\zeta),$$

which leads to the equation

$$\zeta G'' + (|m| + 1 - \zeta)G' + [\lambda - \frac{1}{2}(|m| + 1)]G = 0. \tag{4.77}$$

This equation is again of the form of Eq. (D.67), and the solutions f and g are obtained by setting $\lambda_1 = \frac{1}{2}(|m|+1)+n_1$ for f and $\lambda_2 = \frac{1}{2}(|m|+1)+n_2$ for g, where n_1 and n_2 are non-negative integers; hence

$$\lambda_1 + \lambda_2 = \frac{Ze^2}{\hbar}\sqrt{\frac{\mu}{2|E|}} = n_1 + n_2 + |m| + 1 \stackrel{\text{def}}{=} n.$$

[27] **Show this.**

The resulting energy eigenvalues agree with (4.70), and the (non-normalized) eigenfunctions are given by

$$\psi_{mn_1n_2}(\xi,\eta,\varphi) = e^{-\alpha(\xi+\eta)/2n}(\xi\eta)^{|m|/2}L^{|m|}_{n_1+|m|}(\alpha\xi/n)L^{|m|}_{n_2+|m|}(\alpha\eta/n)e^{im\varphi}, \tag{4.78}$$

where

$$\alpha \stackrel{\text{def}}{=} \frac{Ze^2\mu}{\hbar^2}.$$

At this point you may be puzzled by the fact that the eigenfunctions shown in (4.71) vanish when the associated Legendre functions of θ vanish, which means that their zero-surfaces are cones and spheres, whereas the eigenfunctions shown in (4.78) vanish when the associated Laguerre polynomials of $\alpha\eta/n$ and $\alpha\xi/n$ vanish, which implies that their zero-surfaces are paraboloids. The reason why these two sets of eigenfunctions are compatible is the l-degeneracy, owing to which the eigenfunctions are not unique, which shows that there is a connection between the unusual degeneracy of the Coulomb spectrum and the fact that the Schrödinger equation with a Coulomb potential is separable in more than one coordinate system.

More on the l-degeneracy

[This section makes use of some of the results of Chapter 5 and is best read after that chapter.]

For the classical Kepler problem with the Hamiltonian $H = \vec{p}^2/2\mu - a/r$, the vector $\vec{v}\times\vec{L} - a\vec{r}/r$, called the *Laplace-Runge-Lenz vector*, is a constant of the motion.[Goldstein] (Here $\vec{v}$ is the velocity of the particle and $\vec{L} = \vec{r}\times\vec{p}$ is its angular momentum.) Similarly, in quantum mechanics, the operator

$$\vec{\mathbf{M}} \stackrel{\text{def}}{=} \vec{\mathbf{p}}\times\vec{\mathbf{L}} - \vec{\mathbf{L}}\times\vec{\mathbf{p}} - \hbar^2\eta\vec{\mathbf{r}}/\mathbf{r}, \qquad \eta \stackrel{\text{def}}{=} 2\mu Ze^2/\hbar^2 \tag{4.79}$$

commutes with the Hamiltonian $\mathbf{H} = \vec{\mathbf{p}}/2\mu - Ze^2/\mathbf{r}$ of the hydrogen atom:[28]

$$[\vec{\mathbf{M}}, \mathbf{H}] = 0.$$

Furthermore, since $\vec{\mathbf{M}}$ is a vector, its commutation relations with the angular momentum $\vec{\mathbf{L}}$ are (see Section 5.1)

$$[\mathbf{L}_2, \mathbf{M}_3] = i\hbar\mathbf{M}_1$$

and its cyclic permutations, while $[\mathbf{L}_j, \mathbf{M}_j] = 0$. In addition, of course, $\vec{\mathbf{L}}$ is a constant of the motion,

$$[\vec{\mathbf{L}}, \mathbf{H}] = 0$$

[28] **Show it.**

and it also satisfies the commutation relations of a vector,

$$[\mathbf{L}_2, \mathbf{L}_3] = i\hbar \mathbf{L}_1,$$

etc. Obviously $\vec{\mathbf{L}} \cdot \vec{\mathbf{M}} = \vec{\mathbf{M}} \cdot \vec{\mathbf{L}} = 0$, and we find that[29]

$$\vec{\mathbf{M}} \times \vec{\mathbf{M}} = -8i\hbar\mu \vec{\mathbf{L}} \mathbf{H}$$

as well as

$$\vec{\mathbf{M}}^2 - \hbar^4 \eta^2 = 8\mu \mathbf{H}(\vec{\mathbf{L}}^2 + \hbar^2 \mathbb{1}). \tag{4.80}$$

Confining ourselves now to the subspace of a fixed bound-state energy level $-|E|$, and defining $\vec{\mathbf{N}} \stackrel{\text{def}}{=} \vec{\mathbf{M}}/2\sqrt{2\mu|E|}$, we then obtain the commutation relations

$$\frac{1}{2}(\vec{\mathbf{L}} \pm \vec{\mathbf{N}}) \times \frac{1}{2}(\vec{\mathbf{L}} \pm \vec{\mathbf{N}}) = i\hbar \frac{1}{2}(\vec{\mathbf{L}} \pm \vec{\mathbf{N}}),$$

which are identical to the commutation relations of the angular momentum $\vec{\mathbf{L}}$. It therefore follows that the eigenvalues of $[\frac{1}{2}(\vec{\mathbf{L}} \pm \vec{\mathbf{N}})]^2$ must be $\hbar^2 n_\pm(n_\pm + 1)$, where the numbers $n_\pm$ are non-negative integers or half-integers. But since $\vec{\mathbf{L}} \cdot \vec{\mathbf{N}} = \vec{\mathbf{N}} \cdot \vec{\mathbf{L}} = 0$, it follows that $(\vec{\mathbf{L}} + \vec{\mathbf{N}})^2 = (\vec{\mathbf{L}} - \vec{\mathbf{N}})^2$, and consequently the eigenvalues of $\vec{\mathbf{L}}^2 + \vec{\mathbf{N}}^2$ are $\hbar^2(n^2 - 1)$, where the $n \stackrel{\text{def}}{=} 2n_\pm + 1$ are positive integers. Eq.(4.80) therefore implies that $\hbar^2 n^2 = Z^2 e^4 \mu/2|E|$, or

$$E = -\frac{Z^2 e^4 \mu}{2\hbar^2 n^2}$$

in agreement with (4.70).

The interest of this derivation of the energy levels of hydrogenic atoms lies in the fact that the six operators $\vec{\mathbf{L}}$ and $\vec{\mathbf{N}}$, all of which commute with the Hamiltonian $\mathbf{H}$, satisfy exactly the commutation relations of the generators of the group O(4) of rotations in 4-dimensional Euclidean space, so that the l-degeneracy of the discrete spectrum implied by (4.70) is the result of the invariance of the hydrogenic Hamiltonian under that group. The degeneracy is therefore not accidental but normal for this symmetry. For further details about the use of the O(4) symmetry, see [Bander], pp. 330 and 346.

4.2.6 Coulomb scattering

Let us next consider the case of positive energy of a system consisting of two particles of charge $Z_1 e$ and $Z_2 e$.[30] Separating the Schrödinger equation in spherical polar coordinates leads to the radial equation

$$\psi_l'' + k^2 \psi_l = \frac{2n}{r} k\psi_l + \frac{l(l+1)}{r^2} \psi_l. \tag{4.81}$$

[29] **Prove these equations.**

[30] For scattering of an electron by a proton, the system of which hydrogen is a bound state, we would have to take $Z_1 = -1$ and $Z_2 = 1$.

where

$$n \stackrel{\text{def}}{=} \frac{Z_1 Z_2 e^2 \mu}{\hbar^2 k} = \frac{Z_1 Z_2 e^2}{\hbar v}.$$

The parameter n here is the analogue of the principal quantum number for the bound states but is now taking on continuous values, and v is the relative velocity of the two particles when they are far apart.

For large values of r, when the centrifugal term is ignored, the solution of (4.81) can be easily seen to be approximately

$$\psi_l \sim e^{\pm i(kr - n\log r)},$$

rather than the usual $e^{\pm ikr}$. Because of the slow decrease of the Coulomb potential, the usual results for scattering, that the radial wave functions simply suffer a constant phase shift, are thus inapplicable. To solve the radial equation, we again set

$$\psi_l(r) \stackrel{\text{def}}{=} r^{l+1} e^{ikr} f_l(r),$$

obtaining the equation

$$r f_l'' + 2(ikr + l + 1) f_l' + 2[ik(l+1) - nk] f_l = 0. \tag{4.82}$$

Comparison with the confluent hypergeometric equation (D.57) in Appendix D.3 shows that the regular solution of this equation is the confluent hypergeometric function

$$f_l(r) = c_l F(l + 1 + in |\, 2l + 2 \,|\, -2ikr),$$

and therefore,

$$\psi_l(r) = c_l r^{l+1} e^{ikr} F(l + 1 + in |\, 2l + 2 \,|\, -2ikr), \tag{4.83}$$

whose asymptotic behavior for large r, according to (D.59), is

$$\begin{aligned} \psi_l(r) &\sim c_l \frac{\Gamma(2l+2)}{\Gamma(l+1-in)} (2ik)^{-(l+1)} e^{ikr - in\log(2ikr)} \\ &+ c_l \frac{\Gamma(2l+2)}{\Gamma(l+1+in)} (-2ik)^{-(l+1)} e^{-ikr + in\log(-2ikr)} \\ &= \frac{c_l (2l+1)! e^{n\pi/2}}{2^l k^{l+1} |\Gamma(l+1+in)|} \sin[kr - n\log 2kr - \frac{\pi}{2} l + \eta_l], \end{aligned} \tag{4.84}$$

where η_l is defined as the phase of $\Gamma(l+1+in)$:

$$\Gamma(l+1+in) \stackrel{\text{def}}{=} |\Gamma(l+1+in)| e^{i\eta_l}.$$

Therefore the asymptotic form of the radial wave function is of the form

$$\psi_l(r) \propto \sin[kr - n\log 2kr - \frac{\pi}{2} l + \eta_l]. \tag{4.85}$$

As in the case of bound states, it will be instructive to separate the Schrödinger equation in parabolic coordinates, as defined in (4.72), and defining the z-axis as the direction of the incoming particles. We make the *ansatz*

$$\psi \stackrel{\text{def}}{=} e^{ik(\eta-\xi)/2}\Xi(\xi),$$

in the hope of finding a solution $\Xi(\xi)$ that asymptotically differs from unity by a term proportional to $e^{ik\xi} = e^{ik(r-z)}$, which, when multiplied by $e^{ik(\eta-\xi)/2}$, will give us the expected outgoing wave. Insertion of the *ansatz* in the Schrödinger equation with the Laplacian (4.73) yields the equation

$$\xi\Xi'' + (1 - ik\xi)\Xi' - nk\Xi = 0, \tag{4.86}$$

identical with the confluent hypergeometric equation (D.57); therefore, the solution of (4.86) regular at the origin is given by

$$\Xi(\xi) = cF(-in|\,1\,|ik\xi),$$

which, according to (D.59), goes asymptotically as

$$\begin{aligned}\Xi(\xi) \quad \sim \quad & c\left\{\frac{1}{\Gamma(1+in)}e^{in\log(-ik\xi)}\left(1-\frac{n^2}{ik\xi}\right)\right.\\ & \left. +\ \frac{1}{\Gamma(in)}e^{ik\xi}e^{-(1+in)\log(ik\xi)}\left(1-\frac{(1+in)^2}{ik\xi}\right)\right\}.\end{aligned}$$

We therefore find the long-distance behavior

$$\begin{aligned}\psi \quad \sim \quad & \frac{ce^{n\pi/2}}{\Gamma(1+in)}\left\{e^{i[kz+n\log[k(r-z)]]}\left(1-\frac{n^2}{ik(r-z)}+\ldots\right)\right.\\ & \left. +\ f_C(\theta)\frac{1}{r}e^{i[kr-n\log(2kr)]}\left(1-\frac{(1+in)^2}{ik(r-z)}+\ldots\right)\right\},\end{aligned} \tag{4.87}$$

where

$$f_C(\theta) = -\frac{ne^{-in\log\sin^2\theta/2+2i\eta_0}}{2k\sin^2\theta/2}, \tag{4.88}$$

and we have used the fact that

$$\frac{\Gamma(1+in)}{\Gamma(-in)} = \frac{\Gamma(1+in)}{\Gamma(1-in)}\frac{\Gamma(1-in)}{\Gamma(-in)} = -ine^{2i\eta_0}.$$

The leading terms in both parentheses will dominate when $|n^2/k(r - z)| \ll 1$, which requires not only that $kr \gg 1$ but also that the angle θ not be too small, i.e., that we are not looking in the forward direction. Since in all practical cases, kr is extremely large, the excluded forward cone is too small to be of any significance. With this proviso, then, the

Coulomb scattering amplitude is represented by (4.88), and the differential cross section for Coulomb scattering is

$$\frac{d\sigma}{d\Omega} = |f_C(\theta)|^2 = \frac{n^2}{[2k\sin^2(\theta/2)]^2} = \left(\frac{Z_1 Z_2 e^2}{2\mu v^2}\right)^2 \csc^4(\theta/2), \tag{4.89}$$

which by a remarkable coincidence is identical to the *Rutherford cross section* for classical Coulomb scattering. You should note particularly the sharp forward peaking of this scattering cross section and its strong increase at small relative velocities. Another notable feature is the fact that the scattering does not depend on whether the particles have charges of equal or opposite signs. (The phase of the scattering amplitude, however, does depend on it, and especially strongly so near the forward direction.)

If the wave function ψ is to be normalized to unit incident flux, (4.87) shows that we must choose

$$c = \frac{e^{-n\pi/2}\Gamma(1+in)}{\sqrt{v}},$$

and the Coulomb wave function becomes, with $n \stackrel{\text{def}}{=} Z_1 Z_2 e^2 \mu/\hbar^2 k$,

$$\psi(r,k,\theta) = \frac{1}{\sqrt{v}} e^{-n\pi/2}\Gamma(1+in) e^{ikr\cos\theta} F(-in|\,1\,|2ikr\sin^2\theta/2), \tag{4.90}$$

whose absolute magnitude squared at the origin is

$$|\psi(0,k,\theta)|^2 = \frac{1}{v} e^{-n\pi} |\Gamma(1+in)|^2,$$

or, since $\Gamma(1+z)\Gamma(1-z) = \pi z/\sin\pi z$,

$$|\psi(0,k,\theta)|^2 = \frac{1}{v}\frac{2n\pi}{e^{2n\pi}-1}\,. \tag{4.91}$$

This is the probability density of finding the two particles in close proximity to one another, an important fact for nuclear reactions. In the limit as $v \to 0$, this probability density approaches $2|n|\pi/v$ when $n < 0$, i.e., when the force is attractive, and it approaches $2|n|\pi e^{-2n\pi}/v$ when $n > 0$, i.e., when the force is repulsive. The two cases differ by a factor of

$$e^{-2\pi Z_1 Z_2 e^2/\hbar v},$$

called the *Gamow factor,* which expresses the strongly repulsive effect of the Coulomb barrier between two particles whose electrical charges have the same sign. It plays a significant role in nuclear physics, for example, in the explanation of some radioactive decays and in the rate of thermonuclear reactions in the sun.

4.3 The Inverse Problem

The determination of the bound-state energies and the scattering amplitude for a given potential function is called the *direct problem* of the Schrödinger equation. However, from the point of view of explaining experimental data, it is often much more interesting to pose the *inverse problem:* given the scattering amplitude and bound-state energies (call them "the data" for short), can we determine the potential in the Schrödinger equation that produces them? The relation between the data and the potential being nonlinear, the solution to this problem is not straightforward.

A relatively simple procedure sometimes employed is to assume that the potential is a member of a given multi-parameter family of functions for which we know how to solve the Schrödinger equation, either exactly or by numerical computation. We then compare the results calculated for a range of values of the parameters with the data and choose the values leading to the best fit. This has the drawback that there is no guaranty that the potential sought is uniquely determined by the data; moreover, the real underlying potential might not belong to the assumed family at all and may have a quite different shape. It is therefore useful to try to approach the inversion of the map from a large class of potentials to the data as a mathematical problem without any preconceptions.

If we confine ourselves to central potentials and expand the scattering amplitude in a partial-wave series, the inverse problem may be posed in two different ways:

1. Suppose the phase shift of one angular momentum ℓ is given for all energies, can we determine the potential?

2. Suppose all phase shifts are given at one non-zero energy E, can we determine the potential?

The answer to the first question is that if there are no bound states of angular momentum ℓ, then the potential is uniquely determined. However, when there are $n \geq 1$ bound states of angular momentum ℓ, we also need to know all their energies. There still remains one additional free parameter for each bound state, a positive constant which specifies the slope of the normalized bound-state wave function at the origin.

The answer to the second question, from an experimental point of view much more interesting than the first, is that there is always at least a one-parameter family of potentials that produce the same phase shifts at the specified energy E. In fact, there exists at least a one-parameter family of potentials for which all phase shifts vanish at E. Only if we restrict ourselves to potentials that decrease faster than $r^{-3/2}$ at infinity is the solution to the inverse problem at fixed energy unique.

4.3.1 Solution procedure

To illustrate the mathematical method that can actually lead from the data of a fixed angular momentum to the potential, take the case $\ell = 0$ and define the function

$$h(r) \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ikr} [e^{2i\delta(k)} - 1] - \sum_n M_n e^{-\kappa_n r}, \tag{4.92}$$

where δ is the s-wave phase shift, $-\hbar^2\kappa_n^2/2M$ are the energies of the bound states, and the M_n are positive constants. The following linear integral equation,

$$A(r, r') = h(r + r') + \int_r^{\infty} dr''\, h(r' + r'')A(r, r''), \tag{4.93}$$

which is called the *Marchenko equation* after the Russian mathematician V.A. Marchenko, then has to be solved for each given r. [In (4.93), r plays the role of a fixed parameter.] Its solution leads to the potential

$$V(r) = -\frac{\hbar^2}{M} \frac{d}{dr} A(r, r), \tag{4.94}$$

which, when inserted in the s-wave Schrödinger equation yields the phase shift δ and the given bound state energies. A similar procedure allows the construction of the potential from a knowledge of the phase shifts of any other angular momentum and the corresponding bound states.

In order to derive the Marchenko equation, we have to define two auxiliary solutions of the s-wave Schrödinger equation. The first is a regular solution $\varphi(k, r)$ that satisfies the boundary condition at the origin,

$$\varphi(k, 0) = 0, \qquad \varphi'(k, 0) = 1.$$

This function is the solution of the linear integral equation[31]

$$\varphi(k, r) = k^{-1} \sin kr + \frac{2M}{k\hbar^2} \int_0^r dr'\, \sin k(r - r')V(r')\varphi(k, r'), \tag{4.95}$$

which can always be solved uniquely by iteration, even for complex values of k. As a result, $\varphi(k, r)$ is an entire analytic function of k, with no singularities anywhere. Since φ vanishes at $r = 0$, it must be a multiple of the physical wave function ψ.

The other solution we need is defined by the boundary condition at infinity,

$$\lim_{r\to\infty} f(k, r)e^{-ikr} = 1. \tag{4.96}$$

It is the solution of the integral equation

$$f(k, r) = e^{ikr} - \frac{2M}{k\hbar^2} \int_r^{\infty} dr'\, \sin k(r - r')V(r')f(k, r'), \tag{4.97}$$

---

[31] **As an exercise, show that a solution of (4.95) solves the Schrödinger equation.**

which also can always be solved by iteration for all k with $\Im k \geq 0$. This function, called the *Jost solution* of the s-wave Schrödinger equation, is an analytic function of k, regular everywhere in the upper half-plane, but since it generally does not vanish at $r = 0$, it is not physical, except when for $k = i\kappa$, with $\kappa > 0$, the function

$$\mathsf{f}(k) \stackrel{\text{def}}{=} f(k, 0), \tag{4.98}$$

vanishes, $\mathsf{f}(i\kappa) = 0$, in which case $f(i\kappa, r)$ is a solution that vanishes at the origin and also decreases exponentially at infinity; it thus is a bound-state wave function and $-\hbar\kappa^2/2M$ is the energy of a bound state. So the bound-state energies are the zeros of $\mathsf{f}(k)$ in the upper half plane. For large $|k|$ with $\nu \stackrel{\text{def}}{=} \Im k \geq 0$, one finds from (4.97) that

$$f(k, r) \sim e^{ikr}, \qquad \mathsf{f}(k) \sim 1 \tag{4.99}$$

and from (4.95) that

$$\varphi(k, r) = k^{-1} \sin kr + o(|k|^{-1} e^{|\nu|r}). \tag{4.100}$$

Since, for real k, the two function $f(k, r)$ and $f(-k, r)$ are linearly independent, it must be possible to express $\varphi(k, r)$ as a linear combination of these two solution of the Schrödinger equation,

$$\varphi(k, r) = af(k, r) + bf(-k, r),$$

and the two coefficients a and b are calculated by means of the boundary condition of φ as $a = \mathsf{f}(-k)/D(k)$ and $b = -\mathsf{f}(k)/D(k)$, where $D(k) \stackrel{\text{def}}{=} f'(k, 0)f(-k, 0) - f'(-k, 0)f(k, 0)$. The expression $D(k)$ is the Wronskian of the two solutions $f(k, r)$ and $f(-k, r)$ of the Schrödinger equation, a second-order ordinary differential equation, and thus is independent of r, as already remarked in Chapter 3. We can thus evaluate it just as well at $r \to \infty$, in which case we obtain from the boundary condition (4.96) that $D = 2ik$, and as a result

$$\varphi = \frac{1}{2ik}[f(k, r)\mathsf{f}(-k) - f(-k, r)\mathsf{f}(k)]. \tag{4.101}$$

Comparison with (4.41) shows that we can therefore conclude,

$$\frac{\mathsf{f}(-k)}{\mathsf{f}(k)} = e^{2i\delta(k)}. \tag{4.102}$$

Now define the Fourier tranform

$$A(r, r') \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{-ikr'}[f(k, r) - e^{ikr}]. \tag{4.103}$$

When $r' < r$, because of the analyticity of $f(k, r)$ and the asymptotics (4.99), the path of this integral can be closed by a large semicircle in the upper half plane, so that Cauchy's theorem implies that $A(r, r') = 0$ for $r' < r$, and the inverse of (4.103) reads

$$f(k, r) = e^{ikr} + \int_{r}^{\infty} dr'\, A(r, r') e^{ikr'}. \tag{4.104}$$

Next, (4.101) together with (4.102) implies that

$$f(-k, r) = e^{2i\delta} f(k, r) - 2ik\varphi(k, r)/\mathsf{f}(k),$$

which we insert in (4.103) after changing k into $-k$, for $r' > r$

$$\begin{aligned} A(r, r') &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ikr'} \{[e^{2i\delta(k)} - 1][f(k, r) - e^{ikr}] + e^{ikr}[e^{2i\delta} - 1] \\ &+ [f(k, r) - e^{ikr}] - 2ik[\varphi(k, r)/\mathsf{f}(k) - k^{-1} \sin kr]\}. \end{aligned}$$

The first term is the convolution[32] of $A(r, -r')$ with the Fourier transform

$$h(r) \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, [e^{2i\delta(k)} - 1]; \tag{4.105}$$

[32] **Show this.**

the second term is $h(r+r')$; and the third and fourth term vanish by contour integration if there are no bound states. The result is (4.93). If there are bound states, the function h has to be defined as in (4.92) rather than as in (4.105), and again the result is (4.93).

Finally, insert (4.104) in the radial Schrödinger equation and you will find that for $r' > r$, the function $A(r, r')$ has to satisfy the partial differential equation

$$\left(\frac{\partial^2}{\partial r^2} - \frac{\partial^2}{\partial r'^2}\right) A(r, r') = \frac{2M}{\hbar^2} V(r) A(r, r')$$

and the boundary condition (4.94). □

Of course, one also has to prove that the potential obtained by (4.94), when used in the s-wave Schrödinger equation, indeed leads to the data used as input in the definition of the function h. For those and further details, including the precise conditions on the potential under which the procedure works, generalizations to $\ell > 0$, as well as a solution to problem (2), the inverse problem at fixed energy, see [Chadan].

Note that the uniqueness of the inversion procedure from a knowledge of a single phase shift of an angular momentum for which the potential produces no bound states implies that if a scattering amplitude is given for all energies, then the fact that it comes from an underlying central potential[33] not only makes the phase shifts of all angular momenta dependent on one another, but this scattering amplitude also determines the bound states. To see this, all you have to do is use a phase shift of an angular momentum for which there are no bound states—in the specified class there always exists such an ℓ-value—to find the underlying potential V; all the other phase shifts and the bound states, if any, are then determined by V.

An example

As a simple example, suppose that, with $a, b > 0$,

$$e^{2i\delta(k)} = \frac{k - ia}{k + ia} \frac{k + ib}{k - ib},$$

which means that

$$k \cot \delta = \frac{ab + k^2}{b - a}. \tag{4.106}$$

The function $h(r)$ is then easily calculated[34] by evaluating the integral in (4.92) using contour integration (after adding a large semicircle in the upper half plane) and applying the Cauchy residue theorem,

$$h(r) = -2b\beta e^{-br}, \qquad \beta \stackrel{\text{def}}{=} \frac{b - a}{b + a}.$$

---

[33] This potential has to be in a class for which the procedure works, which requires that $rV(r)$ be integrable. Roughly speaking, this means that $V(r)$ is not as singular at the origin as r^{-2} and that it decreases faster than r^{-2} at infinity.

[34] **Do it.**

Thus the Marchenko equation (4.93) becomes

$$A(r,r') = -2b\beta[e^{-b(r+r')} + \int_r^\infty dr''\, e^{-b(r'+r'')}A(r,r'')].$$

The function $B(r,r') \stackrel{\text{def}}{=} A(r,r')e^{b(r+r')}$ therefore must satisfy the equation

$$B(r,r') = -2b\beta\left[1 + \int_r^\infty dr''\, e^{-2br''}B(r,r'')\right],$$

which implies that it is independent of r' and we obtain

$$B(r,r') = -\frac{2b\beta}{1+\beta e^{2br}}$$

or

$$A(r,r') = -2b\beta\frac{e^{-b(r+r')}}{1+\beta e^{-2br}}.$$

Therefore the potential is given by

$$V(r) = \frac{2b\beta\hbar^2}{M}\frac{d}{dr}\frac{1}{\beta+e^{2br}} = -\frac{4b^2\beta\hbar^2}{M}\frac{e^{2br}}{(\beta+e^{2br})^2}. \tag{4.107}$$

4.3.2 The phase of the scattering amplitude

The above inversion procedure starts from the assumption that either one phase shift is known for all energies or all phase shifts are known at one energy. In reality, however, a scattering experiment measures the scattering cross section rather than the scattering amplitude, a knowledge of whose phase is needed to determine the phase shifts. Is there a way of determining the phase of the amplitude if the cross section, i.e., its absolute magnitude, is given?

We have one handle on this problem, and that is the generalized optical theorem (4.52). If the scattering amplitude is written in the dimensionless form

$$A(\vec{k}',\vec{k}) \stackrel{\text{def}}{=} k^{-1}F(x) \stackrel{\text{def}}{=} k^{-1}G(x)e^{i\varphi(x)}$$

with $x = \cos\theta$ and $G(x) \geq 0$,[35] a change of variables in the angle integration[36] changes (4.52) into

$$\Im F(x) = \frac{1}{2\pi}\int\int dy\, dz\, \frac{F(y)F^*(z)}{\sqrt{1-x^2-y^2-z^2+2xyz}},$$

[35] In order not to clutter up the notation, the k-dependence of F and G has been suppressed.

[36] **Carry this out as an exercise.**

where the region of integration is the interior of the ellipse in which the radicand in the integral is positive. This equation may be expressed in terms of the phase φ of the scattering amplitude:

$$\sin\varphi(x) = \int\int dy\,dz\,H(x,y,z)\cos[\varphi(y)-\varphi(z)], \tag{4.108}$$

in which

$$H(x,y,z) \stackrel{\text{def}}{=} \frac{G(y)G(z)}{2\pi G(x)\sqrt{1-x^2-y^2-z^2+2xyz}}.$$

Equation (4.108) may be regarded as a nonlinear integral equation for the phase φ of the scattering amplitude if the cross section $G^2(x)$ is experimentally given. Unfortunately, not much is known about general conditions under which it has a solution and when this solution is unique. One thing is immediately clear: if $\varphi(x)$ is a solution of (4.108), then so is $\pi-\varphi(x)$; thus there can at best be uniqueness of the solution of (4.108) modulo this ambiguity: the generalized optical theorem, together with a knowledge of the differential cross section, can determine the phase shifts, if at all, only to within an overall sign. If the function

$$Q(x) \stackrel{\text{def}}{=} \int\int dy\,dz\,H(x,y,z)$$

exceeds 1 for some value of x, there are particular problems concerning both the existence and the uniqueness of a solution, and explicit examples of non-uniqueness (beyond the one mentioned above) are known [Crichton]. However, it can be proved that if for all $-1 \le x \le 1$, $Q(x) \le M < 1$, then (4.108) has a solution; moreover, if $M < 0.62$, then this solution is unique and can be constructed by a convergent sequence of successive approximations.[37]

The fact that (4.108) always has a solution if only the cross section is small enough (so that M is small) has quite interesting physical implications. Suppose an experimentally given differential cross section is expressible as a linear combination of the first $2L+1$ Legendre polynomials, i.e., from $P_0(\cos\theta)$ to $P_{2L}(\cos\theta)$. One would then be tempted to conclude that the corresponding scattering amplitude contains no partial waves higher than L.[38] But since the phase shifts are real, this would allow for only $L+1$ real coefficients, whereas the expansion of the given differential cross section generally requires $2L+1$ real coefficients. It would therefore follow that the reality of the phase shifts (which is equivalent to the generalized

[37] [Newton 82], pp. 633–636.

[38] **As an exercise, show that if the partial-wave expansion of the amplitude terminates exactly at order L, then the Legendre expansion of its magnitude squared must contain a term of order $2L$.**

optical theorem) imposes strong restrictions on the possible angular shape of a differential cross section. We have, however, found that if the cross section is small enough, this conclusion is incorrect: any shape is compatible with the generalized optical theorem. The only way out of this apparent contradiction is that *almost all differential cross sections that can be expressed in terms of the first* $2L+1$ *Legendre polynomials require infinitely many terms in the partial wave expansions of their corresponding amplitudes.*

4.4 Potentials without Spherical Symmetry

4.4.1 Green's functions

In Section 4.2.2, we discussed the scattering amplitude and its partial-wave expansion for a particle subject to a rotationally invariant potential. We will now take up the case in which such symmetry is lacking. In order to define, or calculate numerically, a solution of the Schrödinger equation that satisfies the boundary condition (4.45), it is most convenient to define a *Green's function* that satisfies the equation

$$(\nabla^2 + k^2) g_0^+(k, \vec{r} - \vec{r}\,') = \delta^3(\vec{r} - \vec{r}\,') \tag{4.109}$$

and the "outgoing-wave boundary condition" for $r = |\vec{r}| \to \infty$,

$$g_0^+(k, \vec{r} - \vec{r}\,') \propto \frac{e^{ikr}}{r}.$$

The physical significance of this function is best appreciated by going back to the time-development of a state.

The retarded propagator $\mathcal{G}^+(t)$, which allows us to express the state vector $\Psi(t)$ in terms of the state vector $\Psi(t')$ at any earlier time $t' < t$ by (2.30), was defined in (2.29). What we now need for solving the time-independent Schrödinger equation is its Fourier transform,

$$\mathbf{G}^+(E) \stackrel{\text{def}}{=} \frac{1}{i\hbar} \int_{-\infty}^{\infty} dt\, e^{iEt/\hbar} \mathcal{G}^+(t), \tag{4.110}$$

which, as it stands, is not well defined, because the integral does not converge. However, since $\mathcal{G}^+(t)$ vanishes for $t < 0$, there will be no problem if E is taken to be complex, with a positive imaginary part $i\varepsilon$. The integral is then easily carried out, with the result

$$\mathbf{G}^+(E) = (E + i\varepsilon - \mathbf{H})^{-1}, \tag{4.111}$$

where the limit $\varepsilon \to 0+$ is understood.

Of course, all of this can be done equally well in terms of the free Hamiltonian, consisting only of the kinetic energy, $\mathbf{H}_0 = \vec{\mathbf{p}}^2/2M$, in which case we form first $\mathcal{G}_0^+(t)$ just as in (2.29), and then take its Fourier transform as in (4.110),

$$\mathbf{G}_0^+(E) \stackrel{\text{def}}{=} \frac{1}{i\hbar}\int_{-\infty}^{\infty} dt\, e^{iEt/\hbar}\mathcal{G}_0^+(t), \tag{4.112}$$

with the result

$$\mathbf{G}_0^+(E) = (E + i\varepsilon - \mathbf{H}_0)^{-1}, \tag{4.113}$$

which is called the free Green's function (in contrast to (4.111), which is the *complete Green's function*). In the momentum representation $\mathbf{G}_0^+(E)$ is diagonal and given explicitly by the function

$$\widehat{G}_0^+(E,\vec{p}) = 1\Big/\left(E + i\varepsilon - \frac{\vec{p}^2}{2M}\right). \tag{4.114}$$

From this we calculate the Green's function in the configuration representation by means of (1.71),

$$G_0^+(E;\vec{r},\vec{r}\,') = (2\pi\hbar)^{-3}\int d^3p\, e^{i(\vec{r}-\vec{r}\,')\cdot\vec{p}/\hbar}\widehat{G}_0^+(E,\vec{p}),$$

with the result[39]

$$G_0^+(E;\vec{r},\vec{r}\,') = -\frac{M}{2\pi\hbar^2}\frac{e^{ik|\vec{r}-\vec{r}\,'|}}{|\vec{r}-\vec{r}\,'|}, \tag{4.115}$$

where $k = +\sqrt{2ME}/\hbar$. It satisfies the differential equation[40]

$$\left(\frac{\hbar^2}{2M}\nabla^2 + E\right)G_0^+(E;\vec{r},\vec{r}\,') = \delta^3(\vec{r}-\vec{r}\,') \tag{4.116}$$

and the *outgoing-wave* boundary condition. This function therefore differs from that defined in (4.109) simply by a factor of $2M/\hbar^2$. In fact, since for $|\vec{r}| \gg |\vec{r}\,'|$ and $k|\vec{r}| \gg (k|\vec{r}\,'|)^2$,

$$k|\vec{r}-\vec{r}\,'| = k[|\vec{r}| - \vec{r}\cdot\vec{r}\,'/|\vec{r}| + O((|\vec{r}\,'|/|\vec{r}|)^2))] = kr - \vec{k}'\cdot\vec{r}\,' + o(1),$$

where $\vec{k}' \stackrel{\text{def}}{=} k\vec{r}/r$, we have in that limit

$$G_0^+(E;\vec{r},\vec{r}\,') \simeq -\frac{M}{2\pi\hbar^2}\frac{e^{ikr}}{r}e^{-i\vec{k}'\cdot\vec{r}\,'}. \tag{4.117}$$

The function $G_0^+(E;\vec{r},\vec{r}\,')$ is nothing but the Fourier transform of the propagator already calculated earlier, with the result (4.27). Similarly as in

[39] **Do this integral as an exercise in contour integration in the complex plane.** You should do the angle integration first.

[40] **Verify this.**

(4.30), there is also a (free) *back-propagator*, which determines the state vector $\Psi_0(t)$ in terms of $\Psi_0(t')$ for $t < t'$. Its Fourier transform, the corresponding Green's function, is the operator[41]

$$\mathbf{G}_0^-(E) = (E - i\varepsilon - \mathbf{H}_0)^{-1}, \tag{4.118}$$

which in the configuration representation takes the form[42]

$$G_0^-(E;\vec{r},\vec{r}\,') = -\frac{M}{2\pi\hbar^2}\frac{e^{-ik|\vec{r}-\vec{r}\,'|}}{|\vec{r}-\vec{r}\,'|} = G_0^{+*}(E;\vec{r},\vec{r}\,'). \tag{4.119}$$

Regarded as an operator, we have

$$\mathbf{G}_0^-(E) = \mathbf{G}_0^{+\dagger}(E). \tag{4.120}$$

Let us now use these Green's functions for solving the Schrödinger equation. In order to solve the equation

$$\left(\frac{\hbar^2}{2M}\nabla^2 + E\right)\psi = V\psi$$

for $E > 0$, with the outgoing-wave boundary condition, apply the inverse of the operator on the left, which according to (4.113) is the Green's function G_0^+, and you obtain the equation

$$\psi^+(\vec{k},\vec{r}) = (2\pi)^{-3/2}e^{i\vec{k}\cdot\vec{r}} + \int d^3r'\, G_0^+(E;\vec{r},\vec{r}\,')V(\vec{r}\,')\psi^+(\vec{k},\vec{r}\,'), \tag{4.121}$$

or explicitly,

$$\psi^+(\vec{k},\vec{r}) = (2\pi)^{-3/2}e^{i\vec{k}\cdot\vec{r}} - \frac{M}{2\pi\hbar^2}\int d^3r'\,\frac{e^{ik|\vec{r}-\vec{r}\,'|}}{|\vec{r}-\vec{r}\,'|}V(\vec{r}\,')\psi^+(\vec{k},\vec{r}\,'), \tag{4.122}$$

which is called the *Lippmann-Schwinger equation.* The inhomogeneous term $\psi_0(\vec{k},\vec{r}) \stackrel{\text{def}}{=} (2\pi)^{-3/2}e^{i\vec{k}\cdot\vec{r}}$ in it arises because the operator $E - \mathbf{H}_0 = E + \frac{\hbar^2}{2M}\nabla^2$ annihilates it, so that the inverse $(E - \mathbf{H}_0)^{-1}$ cannot be well defined, which is also reflected in the need for the $i\varepsilon$ in (4.113).

Any solution of (4.122) solves the Schrödinger equation.[43] Moreover, according to (4.117), this solution has the asymptotic form

$$\psi^+(\vec{k},\vec{r}) = (2\pi)^{-3/2}e^{i\vec{k}\cdot\vec{r}} - \frac{M}{2\pi\hbar^2 r}e^{ikr}\int d^3r'\, e^{-i\vec{k}'\cdot\vec{r}\,'}V(\vec{r}\,')\psi^+(\vec{k},\vec{r}\,') + o(\frac{1}{r}), \tag{4.123}$$

[41] **Show this.**
[42] **Show this.**
[43] **Verify this explicitly.**

where $\vec{k}' \stackrel{\text{def}}{=} k\vec{r}/r$. The meaning of the label $\vec{k}$ on the wave function $\psi^+(\vec{k},\vec{r})$ resides entirely in the use of the inhomogeneity $\psi_0(\vec{k},\vec{r}) = (2\pi)^{-3/2}e^{i\vec{k}\cdot\vec{r}}$ in (4.122), which means, physically, that it reflects the fact that the time-dependent origin of $\psi^+(\vec{k},\vec{r})$ was a wave function describing a particle that *in the remote past* had the (approximate)[44] momentum $\hbar\vec{k}$. (The momentum is, of course, not a constant of the motion!)

The scattering wave functions $\psi^+(\vec{k},\vec{r})$, together with the bound-state functions $\psi_{E_n}(\vec{r})$, form a complete set in the Hilbert space in the sense of the completeness relation (B.16) and the generalization of (B.20) in the appendix, which may be written symbolically

$$\sum_n \psi_{E_n}(\vec{r})\psi^*_{E_n}(\vec{r}') + \int d^3k\, \psi^+(\vec{k},\vec{r})\psi^{+*}(\vec{k},\vec{r}') = \delta^3(\vec{r}-\vec{r}'), \qquad (4.124)$$

meaning that every square-integrable function on $\mathbb{R}^3$ may be expanded in the form

$$f(\vec{r}) = \sum_n c_n\psi_{E_n}(\vec{r}) + \int d^3k\, g(\vec{k})\psi^+(\vec{k},\vec{r}),$$

where

$$c_n = \int d^3r\, f(\vec{r})\psi^*_{E_n}(\vec{r}), \qquad g(\vec{k}) = \int d^3r\, f(\vec{r})\psi^{+*}(\vec{k},\vec{r}).$$

Applying the operator $(E \pm i\varepsilon - \mathbf{H})^{-1}$ to (4.124) leads to the following expression for the complete Green's function:

$$G^{\pm}(E;\vec{r},\vec{r}') = \sum_n \frac{\psi_{E_n}(\vec{r})\psi^*_{E_n}(\vec{r}')}{E-E_n} + \int d^3k\, \frac{\psi^+(\vec{k},\vec{r})\psi^{+*}(\vec{k},\vec{r}')}{E \pm i\varepsilon - \frac{\hbar^2}{2M}k^2}. \qquad (4.125)$$

This formula carries important information about the dependence of the complete Green's function on the variable E: it implies that for fixed $\vec{r}$ and $\vec{r}'$, the functions $G^+(E;\vec{r},\vec{r}')$ and $G^-(E;\vec{r},\vec{r}')$ are the boundary values on the real axis of an analytic function $G(E;\vec{r},\vec{r}')$ of E, which is *meromorphic in the complex plane cut along the continuous spectrum from $E = 0$ to $E = \infty$, with simple poles at the bound-state eigenvalues E_n*. Moreover, the residue of $G(E;\vec{r},\vec{r}')$ at the pole E_n is equal to

$$\text{res}_n = \psi_{E_n}(\vec{r})\psi^*_{E_n}(\vec{r}'), \qquad (4.126)$$

where $\psi_{E_n}(\vec{r})$ is the normalized bound-state eigenfunction. The discontinuity of $G(E)$ across the cut is found by Eq.(A.2) in the appendix to be

[44] Remember that a particle whose state is described by a well-defined time-dependent wave function in a Hilbert space cannot have a precisely defined momentum; if square-integrable, it must be a superposition of quasi-eigenfunctions of $\vec{\mathbf{p}}$ with various values of $\vec{p}$.

given by

$$(E + i\varepsilon - \mathbf{H})^{-1} - (E - i\varepsilon - \mathbf{H})^{-1} = -2\pi i \delta(E - \mathbf{H}), \tag{4.127}$$

where the symbolic expression $\delta(E - \mathbf{H})$ is, according to (4.125), explicitly given by

$$G(E + i\varepsilon; \vec{r}, \vec{r}') - G(E - i\varepsilon; \vec{r}, \vec{r}') = \frac{2\pi M k}{i\hbar^2} \int d\Omega'' \psi^+(\vec{k}'', \vec{r}) \psi^{+*}(\vec{k}'', \vec{r}'), \tag{4.128}$$

or

$$\delta(E - \mathbf{H}) \stackrel{\text{def}}{=} \frac{Mk}{\hbar^2} \int d\Omega'' \psi^+(\vec{k}'', \vec{r}) \psi^{+*}(\vec{k}'', \vec{r}').$$

The complete Green's function can be used to solve (4.121) and express ψ^+ in the form[45]

$$\psi^+(\vec{k}, \vec{r}) = \psi_0(\vec{k}, \vec{r}) + \int d^3r' \, G^+(E; \vec{r}, \vec{r}') V(\vec{r}') \psi_0(\vec{k}, \vec{r}'), \quad E = \frac{\hbar^2 k^2}{2M}, \tag{4.129}$$

which is useful for many purposes, even though G^+ can rarely be explicitly constructed.

4.4.2 The scattering amplitude

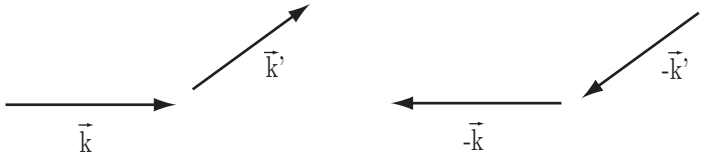

FIGURE 4.5. Initial and final particle momenta in the original scattering and in its reciprocal.

Comparison of (4.123) with (4.45) shows that the scattering amplitude can be expressed in terms of ψ^+ as

$$A(\vec{k}', \vec{k}) = -\frac{M}{2\pi\hbar^2} \int d^3r' \, \psi_0^*(\vec{k}', \vec{r}') V(\vec{r}') \psi^+(\vec{k}, \vec{r}'). \tag{4.130}$$

In contrast to the case of a central potential discussed in Section 4.2.2, the scattering amplitude now depends separately on the direction of incidence and the scattering direction rather than only on the angle between the two.

[45] **You should verify that the function defined by (4.129) solves the Schrödinger equation.**

This dependence, however, is subject to an important symmetry, called *reciprocity*,

$$A(\vec{k}', \vec{k}) = A(-\vec{k}, -\vec{k}'), \tag{4.131}$$

depicted in Figure 4.5. It says that the scattering from the direction $\vec{k}$ to $\vec{k}'$ is the same as that in the reverse direction, from $-\vec{k}'$ to $-\vec{k}$. An crucial ingredient in its derivation is the assumption that the potential V is *real*, which, as we shall see later, is equivalent to *time-reversal invariance.*

To demonstrate (4.131), write (4.130) in the form $-4\pi A(\vec{k}', \vec{k}) = (\psi_0(\vec{k}'), \mathcal{V}\psi^+(\vec{k}))$, $\mathcal{V} \stackrel{\text{def}}{=} 2MV/\hbar^2$. Substituting (4.121) and using the fact that $\mathbf{G}_0^- = \mathbf{G}_0^{+\dagger}$ can be employed as in (4.121) to define a solution $\psi^-(\vec{k}, \vec{r}) = \psi^{+*}(-\vec{k}, \vec{r})$, one obtains

$$\begin{aligned} -4\pi A(\vec{k}', \vec{k}) \quad &= \quad (\psi_0(\vec{k}'), \mathcal{V}\psi_0(\vec{k})) + (\psi_0(\vec{k}'), \mathcal{V}\mathbf{G}^+ V\psi_0(\vec{k})) = (\psi_0(\vec{k}'), \mathcal{V}\psi_0(\vec{k})) \\ &\quad + (\mathbf{G}^- V\psi_0(\vec{k}'), \mathcal{V}\psi_0(\vec{k})) = (\psi^-(\vec{k}'), \mathcal{V}\psi_0(\vec{k})) \\ &\quad = (\psi_0(-\vec{k}), \mathcal{V}\psi^+(-\vec{k}')) = -4\pi A(-\vec{k}, -\vec{k}'). \end{aligned}$$

Furthermore, the scattering amplitude satisfies the *optical theorem,*

$$\Im A(\vec{k}, \vec{k}) = \frac{k}{4\pi}\sigma_{\text{total}}(\vec{k}) = \frac{k}{4\pi}\int d\Omega' \, |A(\vec{k}', \vec{k})|^2, \tag{4.132}$$

which we already encountered in (4.53) for the special case of a central potential.

To prove (4.132) in the physically most meaningful way, call the direction of incidence the z-axis, so that the wave function at large distance $z \gg 1/k$ in the nearly forward direction, in terms of the incident wave ψ_{inc}, has the form

$$\psi^+ \simeq [1 + r^{-1}e^{ik(r-z)}A(0)]\psi_{\text{inc}}e^{ikz}$$

if we write $A(0)$ for the forward scattering amplitude $A(\vec{k}, \vec{k})$. Let this wave (or the corresponding particles) be detected on a screen whose dimensions are small compared to the distance from the scattering center, so that on it not only

$$x^2 + y^2 \ll z^2,$$

because the scattering angle is small, but also

$$zk\left(\frac{x^2 + y^2}{z^2}\right)^2 \ll 1.$$

Expansion then gives

$$k(r - z) = k[\sqrt{z^2 + (x^2 + y^2)} - z] = k(x^2 + y^2)/2z + o(1)$$

and thus

$$r^{-1}e^{ik(r-z)} \simeq z^{-1}e^{ik(x^2+y^2)/2z}.$$

At a point (x, y) on the screen we therefore have

$$\begin{aligned} |\psi^+|^2 &\simeq |\psi_{\rm inc}|^2 |1 + z^{-1} e^{ik(x^2+y^2)/2z} A(0)|^2 \\ &\simeq |\psi_{\rm inc}|^2 \left\{ 1 + 2z^{-1} \Re[e^{ik(x^2+y^2)/2z} A(0)] \right\}. \end{aligned}$$

To calculate the total number of particles detected, integrate this over the area of the screen. Assuming that the screen is *large* in the sense that on its edge

$$x^2 + y^2 \gg \frac{2\pi z}{k},$$

(while still $x^2 + y^2/z^2 \ll 1$) the resulting integral $\int dx\, dy\, e^{ik(x^2+y^2)2z}$ can be extended, with little error, all the way to infinity, and we have[46]

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, e^{ik(x^2+y^2)/2z} = 2\pi i z k^{-1}.$$

If the area of the screen is a, we therefore find that the total number of particles detected on the screen is

$$|\psi_{\rm inc}|^2 [a - 4\pi k^{-1} \Im A(0)],$$

in which the second term represents the amount by which the particle count has been diminished. The condition on the diameter D of the screen for this approximation to be good is that

$$\sqrt{z/k} \ll D \ll z,$$

so that the screen must be *neither too large nor too small;* specifically, it should include many diffraction rings, and not only the central spot. The effect here calculated is generally known as *shadow scattering*; it is the result of an interference between the incident beam and the forward-scattered particles.

Since the number of particles (or the probability of their detection) is conserved, the flux missing in the forward direction must have gone somewhere else: it must be equal to the *total number of particles scattered.* We therefore conclude that

$$\sigma_{\rm total} = 4\pi k^{-1} \Im A(0),$$

which is (4.132).[47]

---

[46] **Do this as an exercise.**

[47] This manner of deriving the optical theorem clearly shows that if there is not only elastic scattering but also absorption or inelastic scattering, $\sigma_{\rm total}$ has to include those cross sections as well.

Eq.(4.132) is a special case of the *generalized optical theorem,*

$$\begin{aligned} A(\vec{k}',\vec{k}) - A^*(\vec{k},\vec{k}') &= \frac{ik}{2\pi}\int d\Omega''\, A^*(\vec{k}'',\vec{k}')A(\vec{k}'',\vec{k}) \\ &= \frac{ik}{2\pi}\int d\Omega''\, A(\vec{k}',\vec{k}'')A^*(\vec{k},\vec{k}''), \end{aligned} \tag{4.133}$$

which we have already encountered for radial potentials in (4.52) and from which (4.132) follows when $\vec{k}' = \vec{k}$. If the *S matrix* is defined by

$$S(\vec{k}',\vec{k}) \stackrel{\text{def}}{=} \delta^2(\Omega',\Omega) - \frac{k}{2\pi i}A(\vec{k}',\vec{k}), \tag{4.134}$$

then (4.133) is the statement that S is *unitary*:[48]

$$\int d\Omega''\, S^*(\vec{k}'',\vec{k}')S(\vec{k}'',\vec{k}) = \int d\Omega''\, S(\vec{k}',\vec{k}'')S^*(\vec{k},\vec{k}'') = \delta(\Omega,\Omega'). \tag{4.135}$$

As the derivation of the optical theorem indicated, the basic origin of the unitarity of the S matrix is the *conservation of probability* or the *conservation of flux.*

For the derivation of (4.133) it is most convenient to write the solution of the Lippmann-Schwinger equation in the form (4.129) in terms of the complete Green's function G^+. Therefore, the scattering amplitude may be written in the form $-4\pi A(\vec{k}',\vec{k}) = (\psi_0(\vec{k}'),\mathcal{V}\psi_0(\vec{k})) + (\psi_0(\vec{k}'),\mathcal{V}\mathbf{G}^+(E)V\psi_0(\vec{k}))$, $\mathcal{V} \stackrel{\text{def}}{=} 2MV/\hbar^2$, and as a result, by (4.128),

$$\begin{aligned} A(\vec{k}',\vec{k}) - A^*(\vec{k},\vec{k}') &= -\frac{M}{2\pi\hbar^2}(\psi_0(\vec{k}'),V[\mathbf{G}^+(E) - \mathbf{G}^-(E)]V\psi_0(\vec{k})) \\ &= \frac{ik}{(2\pi)(4\pi)^2}\int d\Omega''\,(\psi_0(\vec{k}'),\mathcal{V}\psi^+(\vec{k}''))(\psi^+(\vec{k}''),\mathcal{V}\psi_0(\vec{k})) \\ &= \frac{ik}{2\pi}\int d\Omega''\, A(\vec{k}',\vec{k}'')A^*(\vec{k},\vec{k}''), \end{aligned}$$

which proves the second version of (4.133); the first version then follows from (4.131).□

4.4.3 Solving the Lippmann-Schwinger equation

The most straightforward method of solving the integral equation (4.121) is to iterate it, that is, to re-insert it in itself repeatedly, thereby generating the series

$$\begin{aligned} \psi^+(\vec{k},\vec{r}) &= e^{i\vec{k}\cdot\vec{r}} + \int d^3r'\, G_0^+(E;\vec{r},\vec{r}')V(\vec{r}')e^{i\vec{k}\cdot\vec{r}'} \\ &+ \int d^3r'd\vec{r}''\, G_0^+(E;\vec{r},\vec{r}')V(\vec{r}')G_0^+(E;\vec{r}',\vec{r}'')V(\vec{r}'')e^{i\vec{k}\cdot\vec{r}''} \\ &+ \ldots, \end{aligned} \tag{4.136}$$

[48] **Show this.**

called the *Born series.*[49] If the potential V is multiplied by a "coupling constant" λ, (4.136) is a power series in λ. For sufficiently weak potentials (or small enough λ), the first term in (4.136) may suffice, which results in the *Born approximation* to the scattering amplitude

$$A_{\text{Born}}(\vec{k}', \vec{k}) \stackrel{\text{def}}{=} -\frac{M}{2\pi\hbar^2} \int d^3r\, V(\vec{r}) e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}. \tag{4.137}$$

As you will notice, it has the special property of being a *function of the momentum transfer* $\hbar(\vec{k} - \vec{k}')$ *only.* On purely dimensional grounds, the criterion for its validity may be expected to be roughly that

$$\frac{\overline{V}R}{\hbar v} \ll 1 \tag{4.138}$$

if $\overline{V}$ is the "average" strength of the potential, R is its range, and v is the velocity of the incoming particle. This means that the time R/v that the particle spends inside the potential should be short compared to the characteristic time $\hbar/\overline{V}$ required by the interaction to influence it. As a consequence, the Born approximation can be expected to be good at high energies, and indeed one can show that in general the scattering amplitude approaches it as $k \to \infty$. The criterion (4.138), however, is not foolproof and does not always work. As a notable case in point, the Born approximation to the Rutherford cross section is the exact answer at all energies![50]

The great drawback of the Born series is that it generally fails to converge. While it converges for "sufficiently large" energies and for "sufficiently weak" potentials, the precise meaning of "sufficiently" here is hard to pin down. (See Appendix C for more on the radius of convergence of the Born series.)

There is, however, another procedure of solving the integral equation (4.121), namely, the *Fredholm method.* If the equation is written in the formal shorthand

$$[\mathbb{1} - \mathbf{G}_0^+(E)V]\psi^+(\vec{k}) = \psi_0(\vec{k}),$$

so that its solution is

$$\psi^+(\vec{k}) = [\mathbb{1} - \mathbf{G}_0^+(E)V]^{-1}\psi_0(\vec{k}),$$

and the operator $[\mathbb{1} - \mathbf{G}_0^+(E)V]$ is regarded as a generalized matrix, then we might be tempted to proceed with the construction of $[\mathbb{1} - \mathbf{G}_0^+(E)V]^{-1}$ as with that of the inverse of a matrix, beginning with the calculation of the determinant of $[\mathbb{1} - \mathbf{G}_0^+(E)V]$. As is shown in Appendix C, this idea indeed works, and the determinant is known as the *Fredholm determinant.* Replacing V by λV leads to a series-expansion of $\det_2[\mathbb{1} - \lambda\mathbf{G}_0^+(E)V]$ in

[49] Mathematicians call it a Neumann series.

[50] Not the phase of the amplitude, though.

powers of λ, which (provided V is in a suitable, large class of functions, given by (C.22)) *always converges,* no matter how large λ. (The reason for the replacement of the ordinary Fredholm determinant by the modified one denoted by $\det_2$ is explained in Appendix C.) Thus we have

$$[\mathbb{1} - \lambda \mathbf{G}_0^+(E)V]^{-1}(\vec{r}, \vec{r}') = \frac{\mathcal{N}(E; \vec{r}, \vec{r}')}{\det_2[\mathbb{1} - \lambda \mathbf{G}_0^+(E)V]}, \tag{4.139}$$

where both the numerator and the denominator can be expanded in a power series that converges for all (real and complex) λ.

The determinant $\det_2[\mathbb{1} - \lambda \mathbf{G}_0^+(E)V]$ serves another useful purpose. The *Fredholm alternative* states that an inhomogeneous Fredholm integral equation has a unique solution if and only if its homogeneous version has only the trivial solution, $\psi = 0$. Just as for matrices, a nontrivial solution of the homogeneous equation exists if the determinant vanishes. But since the complete Green's function can be written in the form

$$\mathbf{G}^+(E) = [E + i\varepsilon + \frac{\hbar^2}{2M}\nabla^2 - V]^{-1} = [\mathbb{1} - \mathbf{G}_0^+(E)V]^{-1}\mathbf{G}_0^+(E) \tag{4.140}$$

the vanishing of $\det_2[\mathbb{1} - \lambda \mathbf{G}_0^+(E)V]$ leads to a pole of $\mathbf{G}^+(E)$, indicating a discrete eigenvalue. A nontrivial solution of the homogeneous Lippmann-Schwinger equation can therefore occur only for $E < 0$, and *the zeros of* $\det_2[\mathbb{1} - \mathbf{G}_0^+(E)V]$ *are the bound-state energies.*

4.5 Problems and Exercises

1. Derive (4.46) from (4.45), using (2.39).

2. Consider the one-particle Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2M}\nabla^2\psi(\vec{r},t) + \int d^3r'\, V(\vec{r},\vec{r}')\psi(\vec{r}',t).$$

   Derive the conditions for the nonlocal potential $V(\vec{r},\vec{r}')$ that ensures the conservation of probability. Does a current that depends only on ψ at a point exist in this case?

3. Consider the one-particle Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2M}\nabla^2\psi(\vec{r},t) + [V_1(\vec{r}) + iV_2(\vec{r})]\psi(\vec{r},t),$$

   where V_1 and V_2 are real. Show that probability is not conserved unless $V_2 \equiv 0$ and give an expression for the rate at which probability is "lost" or "gained" in a given spatial volume Ω.

4. If ψ_1 and ψ_2 are two solutions of the same Schrödinger equation, show that

$$\int d^3r\, \psi_1^*(\vec{r},t)\psi_2(\vec{r},t)$$

   does not depend on the time.

5. Show that for a three-dimensional wave packet,

$$\frac{d}{dt}\langle \mathbf{x}^2\rangle = \frac{1}{M}\langle \mathbf{x}\mathbf{p}_x + \mathbf{p}_x\mathbf{x}\rangle.$$

6. Suppose that in the Schrödinger equation for a bounded region the potential is changed everywhere by the addition of a constant c. What is the resulting change in the wave functions and the eigenvalues?

7. Consider the Schrödinger equation in two dimensions with a potential that depends only on the radial distance r from the center. Separate the equation; what are the angle functions and what is the radial equation?

8. Calculate (a) the normalized bound-state eigenfunction of a particle in the ground state of a spherical square-well potential of radius b and depth V_0; (b) the probability of finding the particle outside the well; and (c) the expectation value of its distance from the center of the well.

9. Calculate the mean value $\langle \mathbf{r} \rangle$ of the distance between the electron and the proton in a hydrogen atom in a state of angular momentum l and principal quantum number n.

10. Calculate $\langle \mathbf{r}^2 \rangle$ for a hydrogen atom in a state of angular momentum l and principal quantum number n. From this and the result of the last problem, calculate Δr.

11. Calculate the phase shifts for scattering by an impenetrable sphere, i.e., $V(r) = 0$ for $r > R$, and $V(r) = \infty$ for $r < R$. What is the low-energy limit of the l^{th} phase shift for $Rk \ll 1$? What is the low-energy limit of the total cross section?

12. Consider the special cases of (4.106) with $a = 0$ and with $b = 0$. What are the corresponding potentials? Is there anything unusual about either of these?

13. Find the eigenvalues and eigenfunctions of the Schrödinger equation in three dimensions with the potential $V = \frac{1}{2}c(x^2 + y^2 + z^2)$. What are the degeneracies of the three lowest eigenvalues?

14. Consider the radial Schrödinger equation with $l = 0$ and the potential $V(r) = -\lambda e^{-r/a}$, $\lambda > 0$. Change variables from r to $z \stackrel{\text{def}}{=} e^{-r/2a}$ and show that the result is Bessel's equation. What boundary conditions are to be imposed on the solutions as functions of z, and how can these be used to determine the bound-state energy levels? What is the smallest value of λ for which a bound state can exist?

15. The Schrödinger equation for a rigid body of moment of inertia I about a given axis, constrained to rotate about this axis, is

$$i\hbar \frac{\partial \psi(\varphi, t)}{\partial t} = -\frac{\hbar^2}{2I} \frac{\partial^2 \psi(\varphi, t)}{\partial \varphi^2}.$$

What boundary condition must be supplied for the solutions of this equation? Find the general solution.

16. Consider the radial Schrödinger equation for $l = 0$ with the potential $V(r) = \frac{1}{2}K^2 r^2$ for $r > a$ and $V = 0$ for $r < a$. Derive a transcendental equation that determines the energy eigenvalues.

17. Show that the expectation value of the potential energy of an electron in the nth state of a hydrogen atom is given by

$$-\frac{e^4 M}{\hbar^2 n^2}.$$

From this result find the expectation value of its kinetic energy.

18. Calculate the Born approximation to the differential cross section (a) in the center-of-mass system and (b) in the laboratory system for collisions of two particles of equal mass M if the interaction potential is
$$V(r) = \gamma e^{-\alpha r}.$$

19. Do the same for the potential
$$V(r) = \gamma e^{-\alpha^2 r^2}.$$

20. Do the same for the potential
$$V(r) = \gamma e^{-\alpha r} r^{-3/2}.$$

21. Do the same for the potential
$$V(r) = \gamma e^{-\alpha r}/r.$$
Take the limit $\alpha \to 0$ and comment on the result.

22. The definition
$$\psi_l \simeq A_l \sin(kr - \pi l/2 + \delta_l)$$
of the phase shift δ_l refers explicitly to the wave number k and hence to the energy. How do you reconcile this with the fact that, physically, the energy can be shifted by a constant without effect? Is the zero-point of the energy fixed in scattering theory? If so, by what?

23. Using the fact that the function $G^+(\vec{r}, \vec{r}') \stackrel{\text{def}}{=} -e^{ik|\vec{r}-\vec{r}'|}/4\pi|\vec{r} - \vec{r}'|$ satisfies the equation $(\nabla^2 + k^2)G^+(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}')$ and an outgoing-wave boundary condition, show that the coefficients $G_l^+(r, r')$ in the expansion
$$G^+(\vec{r}, \vec{r}') = \sum_{l,m} Y_l^m(\hat{r}) Y_l^{m*}(\hat{r}') G_l^+(r, r') r r'$$
satisfy the differential equation
$$\left[\frac{\partial^2}{\partial r^2} - \frac{l(l+1)}{r^2} + k^2\right] G_l^+(r, r') = \delta(r - r')$$
and the boundary condition that, for fixed r', as $r \to \infty$ G_l^+ contains outgoing spherical waves only. How does G_l^+ behave when r and r' are interchanged?

24. Use the results of the last problem to construct $G_l^+(r, r')$.

25. (a) Calculate the s-wave scattering phase shift for a repulsive potential shell that vanishes for $r < R$ and for $r > 3R/2$, while for $R < r < 3R/2$ it has the constant value $V = V_0$. (b) What is the amplitude of the $l = 0$ wave function in the region outside the potential shell when V_0 is large (with respect to what?)?

26. For the same potential as in the last problem: (a) Show that for special values of the energy that are very close to eigenvalues for the case when $V_0 = \infty$, the amplitude outside is very small. Also calculate the s-wave phase shift for these special values of the energy. (b) Calculate the derivative of the phase shift at these energies and relate it to the time delay. Are these resonances? (c) Calculate the approximate widths of the peaks in the partial-wave cross section from the derivatives of the phase shifts and relate them to the time-delays. Explain what is going on here physically.

27. Two particles of mass 3×10^{-25}g interact with one another with a force of range $\sim 10^{-12}$ cm. A scattering experiment is carried out at 200 keV center-of-mass energy. If the differential cross section is measured with an accuracy of a few percent, what do you expect the general form of the angle dependence to be?

28. Let the potential in the Schrödinger equation be nonlocal of the form

$$V(\vec{r}, \vec{r}') = \sum_1^N \lambda_n f_n(\vec{r}) f_n(\vec{r}').$$

Calculate the scattering amplitude.

29. Find the energy eigenvalues of the Schrödinger equation in two dimensions for the harmonic oscillator potential $V(x, y) = \frac{1}{2}K(x^2 + 4y^2)$. Is there degeneracy? What are the eigenfunctions?

30. (a) Calculate the scattering cross section in the Born approximation for two particles in their center-of-mass system if their mutual interaction is described by the potential

$$V(r) = \lambda r^{-1} e^{-\beta r} \cos(\alpha r).$$

Discuss what happens when α and β tend to zero. (b) What is the center-of-mass cross section if the two particles are indistinguishable?

31. Consider a particle subject to a central potential of the form $V(r) = \infty$ for $r < a$, $V(r) = -V_0$ for $a < r < b$ and $V(r) = 0$ for $r > b$, with $V_0 > 0$. (a) Calculate the s-wave phase shift δ. How does δ behave for large energy? Comment. Are there resonances at large energy? (b) Derive the transcendental equation that has to be satisfied by the

energy of a bound state. (c) For what values of V_0 are new bound states introduced? What is the value of the cross section at $E = 0$ when V_0 has one of those values?

32. By setting $r = x^2$, show that the Schrödinger equation with a Coulomb potential can be reduced to that for a spherically symmetric harmonic oscillator.

This page intentionally left blank

5
Symmetries

5.1 The Angular-Momentum Operator

As we saw in Section 2.2, the angular momentum of a system of particles, and indeed of any physical system, is the generator of rotations for that system. As a result, the properties of the angular-momentum operator are intimately connected to the system's behavior under rotations. In this chapter we shall study this connection in detail.

If the coordinate frame employed to specify the parameters of a given physical system is rotated, all state vectors will be mapped into new ones, $\Psi \mapsto \Psi'$, in such a way that

$$|(\Psi, \Phi)|^2 = |(\Psi', \Phi')|^2,$$

because all probabilities of physical measurements have to remain unchanged. According to a theorem by Eugene Wigner, the state vectors can therefore always be chosen in one of two ways: the vectors Ψ' are related to Ψ either by a unitary transformation, $\Psi' = \mathsf{U}\Psi$, $\mathsf{U}\mathsf{U}^\dagger = \mathsf{U}^\dagger\mathsf{U} = \mathbb{1}$, or else by an *antiunitary* one, $\Psi' = \mathsf{V}\Psi$, $\mathsf{V}\mathsf{V}^\dagger = \mathsf{V}^\dagger\mathsf{V} = \mathbb{1}$, where V is such $\mathsf{V}a\Psi = a^*\mathsf{V}\Psi$, which means that V is not a linear operator but a so-called *antilinear* one. However, since all pure rotations are continuously connected to the unit operator (no rotation at all), which is unitary, the second possibility is ruled out, and we must have $\Psi' - \Psi = (\mathsf{U} - \mathbb{1})\Psi$.

Consider, then, a rotation of the coordinate frame about the axis $\widehat{n}$ by the infinitesimal angle ε in the right-handed screw sense. (According to Euler's theorem, every rotation in a space of odd dimensions leaves at least

one direction invariant; in three dimenions this direction, called the axis of rotation, is unique—unless the rotation is the identity transformation—and the transformation is a rotation *about that axis.*) The form of the unitary transformation must in that case be

$$\mathsf{U} - \mathbb{1} \stackrel{\text{def}}{=} \frac{1}{i\hbar}\varepsilon\widehat{n}\cdot\vec{\mathbf{J}},$$

where $\vec{\mathbf{J}}$ is Hermitian in order for U to be unitary,[1] and the $\hbar$ is inserted for convenience, making $\varepsilon\widehat{n}\cdot\vec{\mathbf{J}}/\hbar$ dimensionless. By definition, and in agreement with our discussion in Section 2.2, the generator $\vec{\mathbf{J}}$ of rotations is the *angular momentum* operator for the system. For a finite rotation by an angle θ about the axis $\widehat{n}$ the operator U can be obtained as a limit of a large number of small rotations,

$$\mathsf{U} = \lim_{m\to\infty}\left(\mathbb{1} - \frac{i\theta}{m\hbar}\widehat{n}\cdot\vec{\mathbf{J}}\right)^m = e^{-\frac{i}{\hbar}\theta\widehat{n}\cdot\vec{\mathbf{J}}}. \tag{5.1}$$

For operators, we have to require that under rotations of the coordinate axes,

$$(\Psi, \mathbf{G}\Phi) = (\Psi', \mathbf{G}'\Phi') = (\Psi, \mathsf{U}^{-1}\mathbf{G}'\mathsf{U}\Phi),$$

and therefore,

$$\mathbf{G}' = \mathsf{U}\mathbf{G}\mathsf{U}^{-1},$$

which implies that for infinitesimal rotations of the coordinate axes,

$$\mathbf{G}' = \left(\mathbb{1} - \frac{i}{\hbar}\varepsilon\widehat{n}\cdot\vec{\mathbf{J}}\right)\mathbf{G}\left(\mathbb{1} + \frac{i}{\hbar}\varepsilon\widehat{n}\cdot\vec{\mathbf{J}}\right) = \mathbf{G} - \frac{i}{\hbar}\varepsilon[\widehat{n}\cdot\vec{\mathbf{J}}, \mathbf{G}];$$

thus when the coordinate frame is infinitesimally rotated, the infinitesimal change in the operator $\mathbf{G}$ is given by

$$\mathbf{G}' - \mathbf{G} \stackrel{\text{def}}{=} \delta_p\mathbf{G} = -\frac{i}{\hbar}\varepsilon[\widehat{n}\cdot\vec{\mathbf{J}}, \mathbf{G}] = \frac{i}{\hbar}[\mathbf{G}, \vec{\mathbf{J}}]\cdot\widehat{n}\varepsilon. \tag{5.2}$$

5.1.1 Active and passive rotations

To understand the meaning of (5.2), consider two kinds of rotations and carefully distinguish between them: on one hand, those of the coordinate system, that is, of the three axes with respect to which a quantity or the components of an operator such as $\mathbf{G}$ are specified, and, on the other hand, those of the physical system itself. The former are called *passive transformations,* and the latter, *active transformations.* If the reference axes are rotated, the coordinates of a point of a physical system undergo a linear transformation A; if the physical system is rotated and the axes are held

[1] **Prove this.**

fixed, the system's coordinates undergo the transformation A^{-1}, the inverse of the first. The behavior of quantities under passive transformations defines their formal, mathematical classification as scalars, vectors, tensors of rank two, etc. But a physical quantity may be a scalar and yet not be invariant under rotations of the system of which it is a property: the energy of a system of charged particles is a scalar, but in the presence of an external electric field, it is not necessarily invariant under rotations of the system. If $\vec{P}$ is the total momentum of a system and $\vec{A}$ is some given fixed vector, $\vec{A}\cdot\vec{P}$ is a scalar, but it is not invariant under active rotations of the system, because under such rotations $\vec{P}$ changes but $\vec{A}$ remains unaltered. On the other hand, if, like $\vec{P}$, the vector $\vec{A}$ is a property of the system rather than being externally given, then the fact that $\vec{A}\cdot\vec{P}$ is a scalar implies that it is also invariant under active rotations. From the point of view of physical consequences, both active and passive transformations are important to consider.

In quantum mechanics, the distinction between being an externally given quantity and being a dynamical property of a physical system is implemented by the distinction between a *number*[2] and an *operator.* All dynamical variables are operators on the Hilbert space of the system. Therefore, if the components of an operator transform under passive rotations like those of a vector, then they do so also under active rotations. But note that the inner product of two vectors, $\vec{A}\cdot\vec{B}$, though a scalar under passive rotations, is invariant under active rotations only if either both $\vec{A}=\vec{\mathbf{A}}$ and $\vec{B}=\vec{\mathbf{B}}$ are operators or neither one is.

Suppose, then, that $\mathbf{G}$ is an observable that is invariant under active rotations of the system about the axis $\widehat{n}$. The left-hand side of (5.2) must therefore vanish, and it follows that

$$[\mathbf{G},\vec{\mathbf{J}}\cdot\widehat{n}]=0;$$

furthermore, if $\mathbf{G}$ is invariant under *all* rotations, we can conclude that

$$[\mathbf{G},\vec{\mathbf{J}}]=0, \tag{5.3}$$

that is, $\mathbf{G}$ commutes with all three components of the angular momentum. One such quantity is $\mathbf{G}=\vec{\mathbf{J}}\cdot\vec{\mathbf{J}}$, so that we have

$$[\vec{\mathbf{J}}\cdot\vec{\mathbf{J}},\vec{\mathbf{J}}]=0; \tag{5.4}$$

all components of the angular momentum commute with the square of its magnitude, $\vec{\mathbf{J}}^2 \stackrel{\text{def}}{=} \vec{\mathbf{J}}\cdot\vec{\mathbf{J}}=\mathbf{J}_x^2+\mathbf{J}_y^2+\mathbf{J}_z^2$.

[2]By a number I mean what Dirac calls a *c-number.* Do not confuse this with being a numerical function in a specific representation of the Hilbert space, like the potential in the Schrödinger equation in the configuration representation.

For another example, if the Hamiltonian $\mathbf{H}$ (which is a scalar) of a system is invariant under active rotations, it must commute with all components of the angular momentum,

$$[\mathbf{H}, \vec{\mathbf{J}}] = 0,$$

and therefore also with $\vec{\mathbf{J}} \cdot \vec{\mathbf{J}}$. This implies that *every component of the angular momentum is a constant of the motion,* and so is the square of its magnitude, a specific instance of Noether's theorem.

Suppose next that $\vec{\mathbf{G}}$ is a vector variable of a system, that is, it has three components that behave like those of a vector under passive rotations, but it is also internal to the system and therefore must behave like a vector under active rotations as well. Under an active infinitesimal rotation $\varepsilon\widehat{n}$ it must thus change as

$$\delta_a \vec{\mathbf{G}} = -\delta_p \vec{\mathbf{G}} = \varepsilon \widehat{n} \times \vec{\mathbf{G}},$$

or by (5.2),

$$\widehat{n} \times \vec{\mathbf{G}} = \frac{1}{i\hbar}[\vec{\mathbf{G}}, \vec{\mathbf{J}}] \cdot \widehat{n}, \tag{5.5}$$

which means that $\frac{1}{i\hbar}[\mathbf{G}_x, \mathbf{J}_y] = \mathbf{G}_z$, $\frac{1}{i\hbar}[\mathbf{G}_x, \mathbf{J}_z] = -\mathbf{G}_y$, and its cyclic permutations of indices, while all other commutators of components of $\vec{\mathbf{J}}$ and $\vec{\mathbf{G}}$ vanish. A special instance is that of $\vec{\mathbf{G}} = \vec{\mathbf{J}}$, in which case we obtain

$$[\mathbf{J}_x, \mathbf{J}_y] = i\hbar \mathbf{J}_z,$$

and cyclic permutations of its indices, which can also be written in the compact form

$$\vec{\mathbf{J}} \times \vec{\mathbf{J}} = i\hbar \vec{\mathbf{J}}. \tag{5.6}$$

Thus, *the three components of the angular momentum do not commute,* and they cannot all be simultaneously diagonalized. On the other hand, since the operator $\vec{\mathbf{J}}^2 = \vec{\mathbf{J}} \cdot \vec{\mathbf{J}}$ commutes with all the components, we can choose an arbitrary z-axis and use the simultaneous eigenvalues of $\vec{\mathbf{J}}^2$ and $\mathbf{J}_z$ to label states. Furthermore, if the Hamiltonian of the system is invariant under rotations, these eigenvalues are *good quantum numbers*, i.e., they are conserved.

In order to find the eigenvalues of $\mathbf{J}_z$ and $\vec{\mathbf{J}}^2$, define

$$\mathsf{j}_\pm \stackrel{\text{def}}{=} [\mathbf{J}_x \pm i\mathbf{J}_y]/\hbar, \;\; \mathsf{j}_z \stackrel{\text{def}}{=} \mathbf{J}_z/\hbar, \tag{5.7}$$

so that $\mathsf{j}_- = \mathsf{j}_+^\dagger$, and

$$\mathsf{j}^2 \stackrel{\text{def}}{=} \vec{\mathbf{J}}^2/\hbar^2 = \mathsf{j}_-\mathsf{j}_+ + \mathsf{j}_z(\mathsf{j}_z+1) = \mathsf{j}_+\mathsf{j}_- + \mathsf{j}_z(\mathsf{j}_z-1) = \frac{1}{2}(\mathsf{j}_-\mathsf{j}_+ + \mathsf{j}_+\mathsf{j}_-) + \mathsf{j}_z^2. \tag{5.8}$$

The commutation relations (5.6) then lead to

$$\mathsf{j}_z\mathsf{j}_+ = \mathsf{j}_+(\mathsf{j}_z + 1), \qquad \mathsf{j}_z\mathsf{j}_- = \mathsf{j}_-(\mathsf{j}_z - 1), \tag{5.9}$$

as well as

$$[\mathsf{j}_+, \mathsf{j}_-] = 2\mathsf{j}_z. \tag{5.10}$$

Now let $|j, m\rangle$ be a simultaneous eigenstate of j_z with the eigenvalue m and of j^2 with the eigenvalue $j(j+1)$,

$$\mathsf{j}_z|j, m\rangle = m|j, m\rangle, \qquad \mathsf{j}^2|j, m\rangle = j(j+1)|j, m\rangle,$$

or

$$\mathsf{J}_z|j, m\rangle = \hbar m|j, m\rangle, \qquad \vec{\mathbf{J}}^2|j, m\rangle = \hbar^2 j(j+1)|j, m\rangle;$$

then *the spectrum of* $\vec{\mathbf{J}}^2$ *is formed by the values of* j *that are integers or half-integers,* $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$, *and the spectrum of* $\mathsf{J}_z/\hbar$ *consists, for each given* j, *of the* $2j+1$ *values* $m = -j, -j+1, \ldots, j-1, j$.

Proof: Without any prejudgment of the values of m and j, it follows from (5.9) that

$$\mathsf{j}_z\mathsf{j}_+|j, m\rangle = \mathsf{j}_+(\mathsf{j}_z + 1)|j, m\rangle = (m+1)\mathsf{j}_+|j, m\rangle,$$

and

$$\mathsf{j}_z\mathsf{j}_-|j, m\rangle = \mathsf{j}_-(\mathsf{j}_z - 1)|j, m\rangle = (m-1)\mathsf{j}_-|j, m\rangle.$$

Therefore, if m is an eigenvalue of j_z, so is $m+1$, unless $\mathsf{j}_+|j, m\rangle = 0$, and so is $m-1$, unless $\mathsf{j}_-|j, m\rangle = 0$. However, we have

$$\begin{aligned} 0 \leq \| \mathsf{j}_-|j, m\rangle \|^2 &= \langle jm|\mathsf{j}_+\mathsf{j}_-|j, m\rangle = \langle jm|\mathsf{j}^2 - \mathsf{j}_z(\mathsf{j}_z - 1)|j, m\rangle \\ &= j(j+1) - m(m-1) = (j+m)(j-m+1), \end{aligned}$$

and

$$\begin{aligned} 0 \leq \| \mathsf{j}_+|j, m\rangle \|^2 &= \langle jm|\mathsf{j}_-\mathsf{j}_+|j, m\rangle = \langle jm|\mathsf{j}^2 - \mathsf{j}_z(\mathsf{j}_z + 1)|j, m\rangle \\ &= j(j+1) - m(m+1) = (j-m)(j+m+1). \end{aligned}$$

Consequently, $\mathsf{j}_-|j, m\rangle = 0$ implies that either $m = -j$ or $m = j+1$; but the latter possibility is ruled out by the fact that $\mathsf{j}^2 = \frac{1}{2}(\mathsf{j}_-\mathsf{j}_+ + \mathsf{j}_+\mathsf{j}_-) + \mathsf{j}_z^2$ implies[3] $m^2 \leq j(j+1)$. Similarly, $\mathsf{j}_+|j, m\rangle = 0$ implies $m = j$. We can therefore conclude that m has to run in integral steps from $-j$ to $+j$, taking on $2j+1$ values, and thus j and m both have to be either integers or half-integers. □

Just as in the case of the simple harmonic oscillator, the spectra of $\mathbf{J}_z$ and $\vec{\mathbf{J}}^2$ follow directly from their commutation relations. Furthermore, we find that[4]

$$\mathsf{j}_+|j, m\rangle = \sqrt{(j-m)(j+m+1)}|j, m+1\rangle, \tag{5.11}$$

and

$$\mathsf{j}_-|j, m\rangle = \sqrt{(j+m)(j-m+1)}|j, m-1\rangle, \tag{5.12}$$

[3] **Why?**

[4] **Prove these two equations.**

if $|j,m\rangle$, $|j,m+1\rangle$, and $|j,m-1\rangle$ are all normalized to unity. These equations imply that

$$\langle j,m+1|\mathrm{j}_+|j,m\rangle = \langle j,m|\mathrm{j}_-|j,m+1\rangle = \sqrt{(j-m)(j+m+1)},$$

and all other matrix elements of $\mathrm{j}_\pm$ vanish. Consequently the only nonvanishing matrix elements of the x and y components of the angular momentum are just above and below the diagonal, given by

$$\langle j,m+1|\mathbf{J}_x|j,m\rangle = \langle j,m|\mathbf{J}_x|j,m+1\rangle = \frac{1}{2}\hbar\sqrt{(j-m)(j+m+1)}, \quad (5.13)$$

$$-\langle j,m+1|\mathbf{J}_y|j,m\rangle = \langle j,m|\mathbf{J}_y|j,m+1\rangle = \frac{i}{2}\hbar\sqrt{(j-m)(j+m+1)}. \quad (5.14)$$

As another application of (5.11), note that if the Hamiltonian of a physical system is invariant under all rotations, we saw that the angular momentum is a "good quantum number," i.e., both m and j are conserved. But we can say more: the three quantities, $\mathbf{H}$, $\vec{\mathbf{J}}^2$, and $\mathbf{J}_z$ can all be simultaneously diagonalized; let $|E,j,m\rangle$ be an eigenfunction of the three operators with the eigenvalues E, $j(j+1)\hbar^2$, and $m\hbar$, respectively (or a quasi-eigenfunction, if E is a quasi-eigenvalue). We then have

$$\begin{aligned}\mathbf{H}|E,j,m+1\rangle &= c\mathbf{H}\mathrm{j}_+|E,j,m\rangle = c\mathrm{j}_+\mathbf{H}|E,j,m\rangle \\ &= Ec\mathrm{j}_+|E,j,m\rangle = E|E,j,m+1\rangle,\end{aligned}$$

and similarly for j_-. Therefore *the eigenvalue (or quasi-eigenvalue) E of* $\mathbf{H}$ *must be* $(2j+1)$*-fold degenerate.* This important m-degeneracy of the energies of any rotationally invariant system is a special instance of Theorem E.9 in the Appendix, which also tells us that the so-labeled states $|E,j,m\rangle$ span the entire Hilbert space.

5.1.2 An oscillator model

The technique of generating the spectrum of the angular momentum operators $\vec{\mathbf{J}}^2$ and $\vec{\mathbf{J}}_z$ is so similar to that of constructing the spectrum of a harmonic oscillator that it suggests the search for an underlying model that relates the two. In particular, we might wonder if there is an operator with the eigenvalues j, rather than just j^2 with the eigenvalues $j(j+1)$. The following model, which accomplishes both these aims, was invented by Schwinger.

Suppose there are two independent harmonic oscillators, each described as in Sections 1.4.3 and 3.2.6. Let their lowering operators be j_+ and j_-, so that

$$[\mathrm{j}_+,\mathrm{j}_-] = [\mathrm{j}_+,\mathrm{j}_-^\dagger] = 0, \qquad [\mathrm{j}_+,\mathrm{j}_+^\dagger] = [\mathrm{j}_-,\mathrm{j}_-^\dagger] = 1,$$

and the number operators are defined by $\mathsf{N}_\pm \stackrel{\text{def}}{=} \mathsf{j}_\pm^\dagger \mathsf{j}_\pm$. We then know that the spectra of the two commuting Hermitian operators $\mathsf{N}_\pm$ are the non-negative integers, and their simultaneous eigenstates $|n_+, n_-\rangle$ can be generated from the "vacuum state" $|0,0\rangle$, defined by

$$\mathsf{j}_+|0,0\rangle = \mathsf{j}_-|0,0\rangle = 0, \tag{5.15}$$

by means of the creation operators as in (1.44),

$$|n_+, n_-\rangle = \frac{\mathsf{j}_+^{\dagger n_+}\mathsf{j}_-^{\dagger n_-}}{\sqrt{n_+!n_-!}}|0\rangle. \tag{5.16}$$

Now set

$$\mathsf{j}_+ \stackrel{\text{def}}{=} \mathsf{j}_+^\dagger \mathsf{j}_-, \qquad \mathsf{j}_- \stackrel{\text{def}}{=} \mathsf{j}_-^\dagger \mathsf{j}_+, \qquad \mathsf{j}_z \stackrel{\text{def}}{=} \frac{1}{2}(\mathsf{j}_+^\dagger \mathsf{j}_+ - \mathsf{j}_-^\dagger \mathsf{j}_-) = \frac{1}{2}(\mathsf{N}_+ - \mathsf{N}_-), \tag{5.17}$$

and you find that the $\mathsf{j}_\pm$ and j_z satisfy the commutation relations (5.9) and (5.10). Furthermore, if we define $\mathcal{J} \stackrel{\text{def}}{=} \frac{1}{2}(\mathsf{j}_+^\dagger \mathsf{j}_+ + \mathsf{j}_-^\dagger \mathsf{j}_-) = \frac{1}{2}(\mathsf{N}_+ + \mathsf{N}_-)$, Eq.(5.8) leads, after a bit of algebra,[5] to

$$\mathsf{j}^2 = \mathcal{J}(\mathcal{J}+1), \tag{5.18}$$

and it follows that

$$\mathsf{j}_z|n_+, n_-\rangle = \frac{1}{2}(n_+ - n_-)|n_+, n_-\rangle, \quad \mathcal{J}|n_+, n_-\rangle = \frac{1}{2}(n_+ + n_-)|n_+, n_-\rangle. \tag{5.19}$$

Therefore, the labels n_+ and n_- on the states $|n_+, n_-\rangle$ might as well be replaced by the eigenvalues m and j of j_z and $\mathcal{J}$, respectively, with $m = \frac{1}{2}(n_+ - n_-)$ and $j = \frac{1}{2}(n_+ + n_-)$, or $n_+ = j + m$ and $n_- = j - m$. Consequently, the state $|n_+, n_-\rangle$ is an eigenstate of $\vec{\mathbf{J}}^2$ and of $\vec{\mathbf{J}}_z$ with the eigenvalues $\hbar^2 j(j+1)$ and $\hbar m$, respectively, and we have by (5.16),

$$|j, m\rangle = \frac{\mathsf{j}_+^{\dagger n_+}\mathsf{j}_-^{\dagger n_-}}{\sqrt{n_+!n_-!}}|0,0\rangle = \frac{\mathsf{j}_+^{\dagger\, j+m}\mathsf{j}_-^{\dagger\, j-m}}{\sqrt{(j+m)!(j-m)!}}|0,0\rangle. \tag{5.20}$$

For $j = 1/2$, this leads to

$$|\frac{1}{2}, \frac{1}{2}\rangle = \mathsf{j}_+^\dagger|0,0\rangle, \qquad |\frac{1}{2}, -\frac{1}{2}\rangle = \mathsf{j}_-^\dagger|0,0\rangle,$$

so that the two operators $\mathsf{j}_+^\dagger$ and $\mathsf{j}_-^\dagger$ create the up and down states of a spin-1/2 particle. Thus (5.20) may be regarded as a decomposition of states of arbitrary angular momentum into those of spin-1/2.

[5] **Do it.**

The technique employed here, which consists essentially of *factorizing* an operator whose spectrum is bounded below, has found fruitful physical applications in other contexts, such as, for example, in what is called "supersymmetry." Intuitively helpful and appealing though it may be, there is, however, no reason to endow the procedure with any deeper physical significance than its mathematical usefulness.

5.1.3 States of spin-1 and spin-1/2

Suppose, then, that χ^1_m, $m = 1, 0, -1$, are the three normalized simultaneous eigenvectors of $\vec{\mathbf{J}}^2$ with the eigenvalue $2\hbar^2$ (i.e., $j = 1$) and of $\mathbf{J}_z$ with the eigenvalues $\hbar, 0, -\hbar$, respectively, and we write them as

$$\chi^1_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \chi^1_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \chi^1_{-1} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{5.21}$$

In a representation using these three vectors as a basis, the operator $\mathbf{J}_z$ then is represented by the matrix

$$\mathcal{D}^1(\mathbf{J}_z) = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{5.22}$$

and from (5.13) and (5.14) we find that

$$\mathcal{D}^1(\mathbf{J}_x) = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{D}^1(\mathbf{J}_y) = \frac{i\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{5.23}$$

Similarly, if

$$\chi^{1/2}_{1/2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \chi^{1/2}_{-1/2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5.24}$$

are the two simultaneous eigenvectors of $\vec{\mathbf{J}}^2$ with the eigenvalue $3\hbar^2/4$ (i.e., $j = 1/2$) and of $\mathbf{J}_z$ with the eigenvalues $\frac{1}{2}\hbar, -\frac{1}{2}\hbar$, respectively, then the three components of the angular momentum are represented by the matrices $\mathcal{D}^{1/2}(\mathbf{J}_x) = \frac{1}{2}\hbar\sigma_x$, $\mathcal{D}^{1/2}(\mathbf{J}_y) = \frac{1}{2}\hbar\sigma_y$, and $\mathcal{D}^{1/2}(\mathbf{J}_z) = \frac{1}{2}\hbar\sigma_z$, where σ_x, σ_y, and σ_z, are given by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.25}$$

These are called the *Pauli spin matrices,* and they satisfy the equations

$$\sigma_j^2 = \mathbf{1}, \quad j = x, y, z, \qquad \sigma_x\sigma_y = i\sigma_z \quad \text{and cyclic permutations}, \tag{5.26}$$

as well as

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 0, \quad \text{if} \quad i \neq j,$$

which means

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}. \tag{5.27}$$

Since they are representations of the components of the vector $\vec{\mathbf{J}}$, they form a vector $\vec{\sigma}$ under rotations, so that for any given numerical vector $\widehat{n}$, $\widehat{n}\cdot\vec{\sigma}$ is a scalar, which moreover has the property

$$(\widehat{n}\cdot\vec{\sigma})^2 = \mathbb{1}. \tag{5.28}$$

In any adopted coordinate frame, we are, nevertheless, free to use the matrices (5.25) for the three components of $\vec{\sigma}$.

For many calculations it is convenient to note that every 2×2 matrix M can be expressed as a linear combination of the unit matrix and the three Pauli matrices,

$$M = a\mathbf{1} + \sum_j b_j\sigma_j,$$

and the coefficients are given by[6] $a = \frac{1}{2}\mathrm{tr}M$, $b_j = \frac{1}{2}\mathrm{tr}M\sigma_j$.

5.2 The Rotation Group

Since the angular momentum is the generator of rotations, its properties are directly related to those of the three-dimensional rotation group, and it is instructive to connect the physical properties of the angular momentum to mathematical results obtained by studying the representations of the group SO(3) of proper rotations in three dimensions.

Let T be a (real, orthogonal) three-dimensional rotation matrix that transforms the old Cartesian coordinates x_1, x_2, x_3 of a given point into those with respect to new, rotated coordinate axes:[7]

$$x_i' = \sum_{j=1}^{3} T_{ji}x_j, \quad i = 1,2,3, \quad \text{with} \sum_i x_i'^2 = \sum_i x_i^2, \quad \text{and} \quad \det T = 1. \tag{5.29}$$

According to Euler's theorem, mentioned earlier, every rotation about the origin in three dimensions is a rotation *about an axis* through the origin. This axis $\vec{a}$ is a vector left invariant by T, so that $T\vec{a} = \vec{a}$. If its length is taken to be the angle ψ of the rotation (in the right-handed screw sense),

---

[6] **Show it.**

[7] **Explain why it is convenient to define the rotation matrix so that the sum on the right-hand side runs over the first index rather than the second.**

$\vec{a}$ completely determines the rotation, and in spherical polar coordinates it is given by

$$\vec{a} = \psi \begin{pmatrix} \cos\varphi \sin\theta \\ \sin\varphi \sin\theta \\ \cos\theta \end{pmatrix}. \tag{5.30}$$

Thus three angles are required to identify the rotation, $0 \leq \varphi \leq 2\pi$, $0 \leq \theta \leq \pi$, and $0 \leq \psi \leq \pi$. (Any rotation about the axis $\vec{a}$ by an angle $\psi > \pi$ is equivalent to one about $-\vec{a}$ by $\psi < \pi$.) The group manifold, consisting of all the possible endpoints of $\vec{a}$, is therefore a ball of radius π. *This group manifold is doubly connected.*

The reason is that, because a rotation by π about $\vec{a}$ is identical to a rotation by π about $-\vec{a}$, every point on the surface of the ball (the group manifold) has to be identified with its polar opposite. As a result, any curve starting at a point $\mathcal{P}$ on the surface and ending at the polar opposite of $\mathcal{P}$ is *closed,* and such curves cannot be continuously contracted to a point.

As discussed in Appendix E, we can therefore expect the rotation group SO(3) to have double-valued quasi-representations.

Using the orthogonality property of the matrix T, the equation for $\vec{a}$ may be written in the form $(T - \widetilde{T})\vec{a} = 0$, which implies that the components of $\vec{a}$ may be determined by the triple ratio

$$a_1 : a_2 : a_3 = (T_{32} - T_{23}) : (T_{13} - T_{31}) : (T_{21} - T_{12}). \tag{5.31}$$

In order to determine the angle of rotation (and thus the length of $\vec{a}$), do the rotation in three steps: 1) rotate the axis to the z-axis, 2) perform a rotation by the angle ψ (in the right-handed screw sense) about the z-axis, and 3) rotate the axis back to its original position. The rotation about the z-axis is of the form

$$x_i' = \sum_j T_{ji}^{(3)} x_j,$$

where the matrix $T^{(3)}$ is given by

$$T^{(3)} = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If the matrix M takes the z-axis into the axis of rotation with the polar angles φ, θ, then the entire transformation has to be of the form $T = MT^{(3)}M^{-1}$.[8] Therefore we have

$$\operatorname{tr} T = \operatorname{tr} MT^{(3)}M^{-1} = \operatorname{tr} T^{(3)} = 1 + 2\cos\psi, \tag{5.32}$$

---

[8]This implies that, group-theoretically, *all rotations by the same angle ψ are in the same class.*

which makes it easy to calculate ψ from T.

The Hurwitz invariant integrals, to be used for all integrations over functions of the three-dimensional rotation group, according to (E.67) in the Appendix, have a Haar measure $d\mu$ with the weight function $2(1-\cos\psi)\sin\theta$; they are thus of the form

$$\int d\mu \ldots = 2\int_0^{\pi} d\psi(1-\cos\psi)\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta \ldots . \tag{5.33}$$

If the integral is over a class function, i.e., a function of ψ only, it becomes

$$2\int_0^{\pi} d\psi(1-\cos\psi)\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta\, f(\psi) = 8\pi \int_0^{\pi} d\psi\,(1-\cos\psi) f(\psi),$$

and the "volume" of the group is $h = \int d\mu = 8\pi^2$.

Another convenient parametrization of rotations employs *Euler angles*, defined by the following sequence of rotations of the axes:

1. rotate the coordinate frame by α about the z-axis;
2. rotate the new frame by β about the new y'-axis;
3. rotate the new frame by γ about the new z''-axis,

where $0 \le \alpha \le 2\pi$, $0 \le \beta \le \pi$, and $0 \le \gamma \le 2\pi$. Therefore, the rotation matrix T is given by the following product of three matrices:

$$T = \tag{5.34}$$
$$\begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

from which we can calculate the total rotation angle ψ by (5.32), with the result[9]

$$\cos\left(\frac{1}{2}\psi\right) = \left|\cos\left(\frac{1}{2}\beta\right)\cos\left[\frac{1}{2}(\alpha+\gamma)\right]\right|. \tag{5.35}$$

The Haar measure in terms of Euler angles turns out to be[10]

$$d\mu = d\alpha\, d\beta\, d\gamma\, \sin\beta. \tag{5.36}$$

A third useful parametrization of rotations uses the *Cayley-Klein parameters,* defined in terms of the Euler angles by

$$a = e^{-i\frac{1}{2}(\alpha+\gamma)}\cos\left(\frac{1}{2}\beta\right), \qquad b = e^{i\frac{1}{2}(\alpha-\gamma)}\sin\left(\frac{1}{2}\beta\right). \tag{5.37}$$

[9] **You should do this calculation as an exercise.**
[10] **Prove this as an exercise.**

The geometric connection between rotations in three dimensions and the Cayley-Klein parameters is based on a stereographic projection of a sphere of unit diameter, sitting with its south pole on the origin of the complex plane: every point P on the sphere is projected to a point C in the complex plane by drawing a ray from the north pole through P to C in the plane (Fig. 5.1). A rotation of the plane then corresponds to the homographic transformation

$$w = \frac{az + b}{-b^* z + a^*}.$$

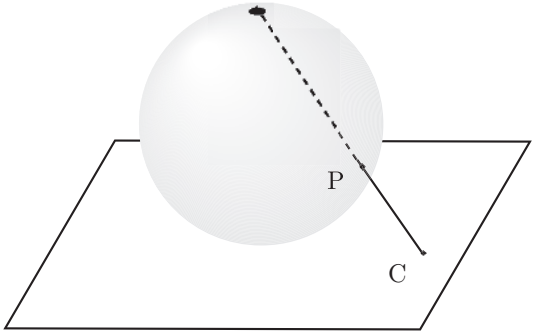

FIGURE 5.1. The stereographic projection C in the complex plane of the point P on the sphere.

The matrix

$$S = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \quad |a|^2 + |b|^2 = 1, \tag{5.38}$$

with a and b defined by (5.37), thus assigns to every rotation in three dimensions a unitary, unimodular 2×2-matrix; these matrices form the group SU(2). Thus the matrices S appear to form a two-dimensional representation of SO(3), except for one fact: a rotation by 2π (take $\alpha = \gamma = 0$, and $\beta = 2\pi$) corresponds to $S = -\mathbb{1}$, so that to every rotation in three dimensions there correspond *two* matrices S: this "representation" is double-valued. (We noted earlier that, because the group manifold of SO(3) is doubly connected, such double-valued "representations" could be expected to exist.) Since the group manifold of SU(2) is simply connected, it has no multiple-valued representations: *SU(2) is the universal covering group of SO(3).* Conversely, the 3×3 matrices T form an unfaithful representation of the group SU(2).

There are infinitely many irreducible representations of SU(2) (see Appendix E), one each of dimensionality $2j + 1$, where $j = \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$. Those with integral j are called *even representations* and they are the irreducible representations of the rotation group SO(3); those with half-integral j are called *odd representations* and they form double-valued quasi-representations of SO(3). From the physical perspective, both the integral

and the half-integral values of j have the significance that $\hbar^2 j(j+1)$ is the eigenvalue of the square of the angular momentum, $\vec{\mathbf{J}}^2 = \vec{\mathbf{J}} \cdot \vec{\mathbf{J}}$. Furthermore, the angular momentum operators $\mathbf{J}_x, \mathbf{J}_y, \mathbf{J}_z$ differ only by a common constant factor from the infinitesimal generators of the rotation group, from which its Lie algebra (see Appendix E) is formed, while the operator $\vec{\mathbf{J}}^2$ differs from the Casimir operator of that algebra again only by a constant factor.

The fact that the even-dimensional "representations" of SO(3) (i.e., those with half-integral values of j) are not single-valued explains why the numbers $l(l+1)\hbar^2$ with $l = \frac{1}{2}, \frac{3}{2}, \ldots$ do not appear as eigenvalues of the square of the orbital angular momentum $\vec{\mathbf{q}} \times \vec{\mathbf{p}}$, a variable that has a classical analogue. The eigenvectors of the angular momentum with half-integral quantum numbers change sign under rotations by 2π, but since all observable quantities contain squares of matrix elements, they are unaffected. Thus these half-integral angular-momentum quantum numbers and all their consequences are typical quantum effects without classical analogues.

5.2.1 Angular momentum and the carrier space

If, for a fixed value of j, the matrices $\{D^j_{\mu\nu}(\alpha, \beta, \gamma)\}$ form the $(2j+1)$-dimensional irreducible representation Γ^j of SU(2), where α, β, γ are the Euler angles of the corresponding rotation in three dimensions, then the carrier space of this representation is spanned by a set of orthonormal vectors Ψ^j_μ, each of which belongs to one row of $D^j_{\mu\nu}$ in the sense that if the rotation operator $\mathcal{O}(\alpha, \beta, \gamma)$ takes Ψ^j_μ into Ψ'^j_μ, then

$$\Psi'^j_\mu = \mathcal{O}(\alpha, \beta, \gamma)\Psi^j_\mu = \sum_\nu D^j_{\nu\mu} \Psi^j_\nu. \tag{5.39}$$

For $j = 1$, Cartesian coordinates lead to the usual form of vectors with three components, V_μ, which transform among each other like those of the Cartesian coordinates of a point in three-dimensional Euclidean space, that is, as in (5.29). This means, if we define

$$\vec{X} \stackrel{\text{def}}{=} \frac{1}{r} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

then under a rotation of the coordinate axes, $\vec{X}' = \widetilde{T}\vec{X}$, where T is the matrix given by (5.29), and if the three components of the *Cartesian vector* $\vec{V}$ are similarly arranged in the form of a column matrix, then $\vec{V}' = \widetilde{T}\vec{V}$. Alternatively, if the three eigenvectors of $\vec{\mathbf{J}}^2$ with the eigenvalue $2\hbar^2$ (i.e., $j = 1$) and the eigenvalues $\hbar, 0, -\hbar$ of $\mathbf{J}_z$ are used (which means that spherical polar coordinates are employed), then the relevant comparison is with the transformation properties of spherical harmonics of order 1: the

three components of a *spherical vector* transform under rotations like the three spherical harmonics Y_1^{-1}, Y_1^0, Y_1^1, and we form

$$\begin{aligned}\mathcal{Y} &\stackrel{\text{def}}{=} \begin{pmatrix} Y_1^1(\theta,\varphi) \\ Y_1^0(\theta,\varphi) \\ Y_1^{-1}(\theta,\varphi) \end{pmatrix} = i\sqrt{\frac{3}{8\pi}} \begin{pmatrix} -e^{i\varphi}\sin\theta \\ \sqrt{2}\cos\theta \\ e^{-i\varphi}\sin\theta \end{pmatrix} \\ &= \frac{i}{r}\sqrt{\frac{3}{4\pi}} \begin{pmatrix} -(x+iy)/\sqrt{2} \\ z \\ (x-iy)/\sqrt{2} \end{pmatrix}.\end{aligned}$$

One then readily finds that[11]

$$\mathcal{Y} = -i\sqrt{\frac{3}{4\pi}}\mathcal{M}\vec{X}, \tag{5.40}$$

where $\mathcal{M}$ is the unitary matrix

$$\mathcal{M} \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i & 0 \\ 0 & 0 & -\sqrt{2} \\ -1 & i & 0 \end{pmatrix}. \tag{5.41}$$

Therefore, under a coordinate rotation $\mathcal{Y}$ transforms into $\mathcal{Y}' = \mathcal{D}^1\mathcal{Y}$, and the transformation matrix $\mathcal{D}^1$ is related to T by[12]

$$\mathcal{D}^1(\alpha,\beta,\gamma) = \mathcal{M}\widetilde{T}(\alpha,\beta,\gamma)\mathcal{M}^\dagger. \tag{5.42}$$

Every Cartesian vector $\vec{V}$ can be re-expressed by $\mathcal{M}$ as a spherical vector $\mathcal{V}$ by

$$\mathcal{V} = \begin{pmatrix} \mathcal{V}_1 \\ \mathcal{V}_0 \\ \mathcal{V}_{-1} \end{pmatrix} = \mathcal{M}\vec{V}, \tag{5.43}$$

so that its three components belong to the three eigenvalues of $\mathbf{J}_z$, as well as belonging to the angular momentum $j = 1$. For example, when dealing with a vector field ψ_μ, it is physically more meaningful to express its three components as those of a spherical vector, using the matrix $\mathcal{M}$ given by (5.41), so that the particles that emerge from its quantization are clearly labeled as having the intrinsic angular momentum or *spin* 1. *This is the connection between the fact that the electromagnetic field is a vector field and the consequence that its quantum, the photon, has spin 1.*

---

[11] **Check this.**

[12] Note two things: 1) whereas the normal enumeration of columns increases from left to right and that of rows from top to bottom, here it is customary to do the opposite; 2) as $\mathcal{Y}$ is a column vector, the square matrix $\mathcal{D}^1$ has to stand on the left, which is the opposite of what (5.39) indicates, so that $\mathcal{D}^1 = \widetilde{D}^1$. As a result, the order of two successive rotations is reversed in these matrices: $\mathcal{D}^1(R_1R_2) = \mathcal{D}^1(R_2)\mathcal{D}^1(R_1)$.

For $j = 1/2$, the two-component objects that transform under rotations using the matrices $\{D^{1/2}_{mm'}\}$ given in the Appendix, are called *spinors.* In other words,

$$\chi^{1/2'}_m = \mathcal{O}(\alpha,\beta,\gamma)\chi^{1/2}_m = \sum_{m'=-1/2}^{1/2} D^{1/2}_{m'm}(\alpha,\beta,\gamma)\chi^{1/2}_{m'}, \quad m = -\frac{1}{2}, \frac{1}{2}. \tag{5.44}$$

This means that if the spinor is written as a column vector,

$$\chi^{1/2} = \begin{pmatrix} a \\ b \end{pmatrix} = a\chi^{1/2}_{1/2} + b\chi^{1/2}_{-1/2}, \tag{5.45}$$

where $\chi^{1/2}_{1/2}$ and $\chi^{1/2}_{-1/2}$ are spin-up and spin-down spinors as given in (5.24), and the transformation is written in the form $\chi^{1/2\prime} = \mathcal{D}^{1/2}\chi^{1/2}$, then the matrix $\mathcal{D}^{1/2}$ for a rotation by the Euler angles α, β, γ is given by

$$\mathcal{D}^{1/2}(\alpha,\beta,\gamma) = \begin{pmatrix} e^{\frac{i}{2}\alpha} & 0 \\ 0 & e^{-\frac{i}{2}\alpha} \end{pmatrix} \begin{pmatrix} \cos\frac{\beta}{2} & \sin\frac{\beta}{2} \\ -\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} e^{\frac{i}{2}\gamma} & 0 \\ 0 & e^{-\frac{i}{2}\gamma} \end{pmatrix}. \tag{5.46}$$

Thus for spinors, a rotation of the coordinate system by the angle α about the z-axis leads to the new spinor

$$\chi^{1/2'} = \begin{pmatrix} e^{\frac{i}{2}\alpha} & 0 \\ 0 & e^{-\frac{i}{2}\alpha} \end{pmatrix} \chi^{1/2}, \tag{5.47}$$

while a rotation of the coordinate axies about the y-axis by the angle β transforms it into

$$\chi^{1/2'} = \begin{pmatrix} \cos\frac{\beta}{2} & \sin\frac{\beta}{2} \\ -\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} \chi^{1/2}. \tag{5.48}$$

These transformations can be written in terms of the Pauli matrices (5.25) in the following way. Since by (5.26) each of the three Pauli matrices σ_n, $n = 1, 2, 3$, is such that $\sigma_n^2 = \mathbb{1}$, the power series expansion shows that[13]

$$e^{i\omega\sigma_n} = \mathbb{1}\cos\omega + i\sigma_n \sin\omega. \tag{5.49}$$

Therefore it is easily seen that

$$e^{i\frac{\alpha}{2}\sigma_z} = \begin{pmatrix} e^{\frac{i}{2}\alpha} & 0 \\ 0 & e^{-\frac{i}{2}\alpha} \end{pmatrix},$$

and

$$e^{i\frac{\beta}{2}\sigma_y} = \begin{pmatrix} \cos\frac{\beta}{2} & \sin\frac{\beta}{2} \\ -\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix},$$

[13] **Verify this.**

so that for a rotation of the coordinate system by the angle α about the z-axis,

$$\chi^{1/2\prime} = e^{i\frac{\alpha}{2}\sigma_z}\chi^{1/2},$$

and for a rotation by β about the y-axis,

$$\chi^{1/2\prime} = e^{i\frac{\beta}{2}\sigma_y}\chi^{1/2}.$$

These equations are nothing but restatements of (5.1), together with the fact that for $j = 1/2$, the components of the angular momentum operator $\vec{\mathbf{J}}$ are represented by $\mathcal{D}^{1/2}(\mathbf{J}_n) = \frac{1}{2}\hbar\sigma_n$, $n = 1, 2, 3$. For an arbitrary rotation by ω about the axis $\widehat{n}$, we have

$$\chi^{1/2\prime} = e^{i\frac{\omega}{2}\widehat{n}\cdot\vec{\sigma}}\chi^{1/2} \tag{5.50}$$

in terms of the vector $\vec{\sigma}$ defined after (5.27). Because of (5.28), the matrix $e^{-i\frac{\omega}{2}\widehat{n}\cdot\vec{\sigma}}$ can be written as

$$e^{i\frac{\omega}{2}\widehat{n}\cdot\vec{\sigma}} = \mathbb{1}\cos\frac{\omega}{2} + i\widehat{n}\cdot\vec{\sigma}\sin\frac{\omega}{2}, \tag{5.51}$$

and use of (5.25) leads to

$$\widehat{n}\cdot\vec{\sigma} = \begin{pmatrix} \cos\theta & e^{-i\varphi}\sin\theta \\ e^{i\varphi}\sin\theta & -\cos\theta \end{pmatrix} \tag{5.52}$$

if the vector $\widehat{n}$ has the components $(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, so that θ and φ are the polar angles of the direction of $\widehat{n}$. Eq.(5.50) shows, when the rotation is by 2π, the rotation matrix is $e^{i\pi\widehat{n}\cdot\vec{\sigma}} = -\mathbb{1}$, so that spinors change sign under a rotation by 360^0; this is the double valuedness of the quasi-representations for half-integral j. Since the three components σ_j, $j = 1, 2, 3$, form a vector, i.e., they transform among one another under rotations of the coordinate system like the Cartesian coordinates (so that $\widehat{n}\cdot\vec{\sigma}$ is a scalar), they must satisfy the equation

$$e^{-i\frac{\omega}{2}\widehat{n}\cdot\vec{\sigma}}\sigma_j e^{i\frac{\omega}{2}\widehat{n}\cdot\vec{\sigma}} = \sum_k T_{kj}\sigma_k, \tag{5.53}$$

where T is the rotation matrix given by (5.34) in terms of Euler angles.

For any given normalized spinor $\chi^{1/2}$, written in the form (5.45), define the angles ξ, η, ζ by writing $a = e^{i\eta}\cos\xi$, $b = e^{i\zeta}\sin\xi$. It then follows that, if the system is in the state $\chi^{1/2}$, the probability of finding the spin to be *up*, i.e., for the system to be in the state $\chi^{1/2}_{+1/2}$, is $\cos^2\xi$. We calculate the expectation value of the spin by using (5.25), finding that[14]

$$\frac{2}{\hbar}\langle\vec{\mathbf{J}}\rangle = \langle\vec{\sigma}\rangle = (\sin 2\xi\cos(\zeta-\eta), \sin 2\xi\sin(\zeta-\eta), \cos 2\xi),$$

[14] **Check this.**

which means that $\langle\vec{\sigma}\rangle$ is a unit vector with the polar angles $\theta = 2\xi, \varphi = \zeta - \eta$. We may now ask if there is a direction $\widehat{n}$ such that

$$\widehat{n} \cdot \vec{\sigma}\chi^{1/2} = \chi^{1/2}.$$

If such a spinor $\chi^{1/2} \neq 0$ exists, then in the rotated coordinate system in which the z-axis points in the direction $\widehat{n}$, it has "spin up" (i.e., $\chi^{1/2}$ is an eigenvector of σ_z with the eigenvalue 1). A simple calculation[15], using (5.45) and (5.52), leads to the conclusion that indeed such a $\chi^{1/2}$ exists, and the polar angles of $\widehat{n}$ are given by the equation $a/b = e^{-i\varphi}\cot\frac{\theta}{2}$, so that

$$\cot\frac{\theta}{2} = \frac{|a|}{|b|} = \cot\xi, \qquad \varphi = \zeta - \eta. \tag{5.54}$$

This result has the physical interpretation that for the spinor (5.45), the angles $\theta = 2\xi$ and $\varphi = \zeta - \eta$ are the polar angles of the "direction of the spin."[16] However, the "spin direction," as such, of a particle is not a directly experimentally observable vector.

Behavior under reflections

Whereas under coordinate rotations the components of $\vec{\mathbf{J}}$ have to behave like those of a vector, they, like the components of the rotation axis $\widehat{n}$, must not change sign under inversion (again, so that $\widehat{n} \cdot \vec{\mathbf{J}}$ remains invariant): $\vec{\mathbf{J}}$ must be an *axial* vector. This implies that its three components should really be regarded as those of an anti-symmetric tensor of rank 2, and they should be written in the form

$$\mathbf{J}_1 \stackrel{\text{def}}{=} \mathbf{J}_{23} = -\mathbf{J}_{32} \tag{5.55}$$

and its cyclic permutations. From this perspective, $\mathbf{J}_{kl}$ is the generator of rotations in the (kl)-plane, and the commutation relations (5.6) become

$$[\mathbf{J}_{kl}, \mathbf{J}_{rs}] = i\hbar[\delta_{kr}\mathbf{J}_{ls} - \delta_{ks}\mathbf{J}_{lr} + \delta_{ls}\mathbf{J}_{kr} - \delta_{lr}\mathbf{J}_{ks}], \tag{5.56}$$

while the anti-commutation relations (5.27) appropriate for spin-1/2 take the form

$$\frac{1}{2}\{\sigma_{kl}, \sigma_{rs}\} = \delta_{kr}\delta_{ls} - \delta_{ks}\delta_{lr}. \tag{5.57}$$

[15] **Do it.**

[16] Be cautioned that this works for spin-1/2 only! For higher spins there does not always exist a direction of the z-axis such that the given spin function is an eigenvector of $\vec{\mathbf{J}}_z$ with the maximal eigenvalue $\hbar j$.

5.2.2 Polarization and the spin density matrix

Suppose, now, that $|+\rangle \stackrel{\text{def}}{=} \chi_+^{1/2}$ is of the form (5.45), so that it denotes an eigenstate of *spin up* in the direction $\widehat{n}$ specified by $a/b = e^{-i\varphi}\cot\frac{\theta}{2}$, and

$$|-\rangle \stackrel{\text{def}}{=} \chi_-^{1/2} = \begin{pmatrix} b^* \\ -a^* \end{pmatrix} = b^*\chi_{1/2}^{1/2} - a^*\chi_{-1/2}^{1/2}$$

denotes the corresponding *spin down* state. The density operator

$$\boldsymbol{\rho} = p_+|+\rangle\langle+| + p_-|-\rangle\langle-|. \tag{5.58}$$

then defines a mixed state in which p_+ is the probability of finding *spin up* in the direction $\widehat{n}$, and p_- is the probability of finding *spin down* in that direction. A little calculation[17] shows that in this state

$$\vec{P} \stackrel{\text{def}}{=} \langle\vec{\sigma}\rangle = (p_+ - p_-)\widehat{n}. \tag{5.59}$$

The vector $\vec{P}$ is called the *polarization vector* of the state; its direction has the polar angles θ, φ, while its magnitude is called the *degree of polarization,*

$$P \stackrel{\text{def}}{=} |\vec{P}| = |p_+ - p_-|.$$

It can be directly calculated from the density operator by the formula[18]

$$\mathrm{tr}\boldsymbol{\rho}^2 = \frac{1}{2}(1 + P^2), \tag{5.60}$$

and the density matrix (the representation of the density operator on the basis of spin-up and spin-down states with respect to the z-axis) can be expressed in the form

$$\boldsymbol{\rho} = \frac{1}{2}(\mathbb{1} + \vec{P}\cdot\vec{\sigma}). \tag{5.61}$$

If $P = 0$, the spin state is unpolarized, with equal probability for spin up and down; if this is the case for one direction $\widehat{n}$, it is equally the case for any other direction.[19]

For spin-1 (e.g., photons), the spin operator acting on a Cartesian vector $\mathbf{f}$ may be defined, as we already did in (2.107), by

$$\mathsf{S}_j\vec{f} \stackrel{\text{def}}{=} i\hbar\widehat{\mathrm{e}}^{(j)} \times \vec{f}, \qquad j = 1, 2, 3, \tag{5.62}$$

where $\widehat{\mathrm{e}}^{(j)}$ is the unit vector in the direction of the j-axis. The three eigenstates of the z-projection of the spin are then given by[20]

$$|\pm 1\rangle = \frac{1}{\sqrt{2}}(\widehat{\mathrm{e}}^{(2)} \mp i\widehat{\mathrm{e}}^{(1)}), \qquad |0\rangle = i\widehat{\mathrm{e}}^{(3)}.$$

---

[17] **Do it.**
[18] **Show this.**
[19] **Show this.**
[20] **Show that** $\mathsf{S}_3|\pm 1\rangle = \pm\hbar|\pm 1\rangle$ **and** $\mathsf{S}_3|0\rangle = 0$.

Now define

$$|\varphi\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}(e^{i\varphi}|+1\rangle + e^{-i\varphi}|-1\rangle) = \hat{\mathrm{e}}^{(1)}\sin\varphi + \hat{\mathrm{e}}^{(2)}\cos\varphi$$

and

$$|\bar{\varphi}\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}(e^{i\varphi}|+1\rangle - e^{-i\varphi}|-1\rangle) = -i(\hat{\mathrm{e}}^{(1)}\cos\varphi - \hat{\mathrm{e}}^{(2)}\sin\varphi),$$

so that $\langle\bar{\varphi}|\varphi\rangle = 0$. These states correspond to *plane polarization* for a beam in the z direction, φ being the angle of polarization with respect to the y axis for $|\varphi\rangle$, and orthogonal to that for $|\bar{\varphi}\rangle$. For the two probabilities $p_\varphi \geq 0$ and $p_{\bar{\varphi}} \geq 0$, with $p_\varphi + p_{\bar{\varphi}} = 1$, the density operator for a partially polarized beam can be expressed in the form

$$\boldsymbol{\rho}_\varphi = p_\varphi|\varphi\rangle\langle\varphi| + p_{\bar{\varphi}}|\bar{\varphi}\rangle\langle\bar{\varphi}|,$$

and the degree of polarization is given by

$$P = p_\varphi - p_{\bar{\varphi}} = 2p_\varphi - 1.$$

A bit of algebra shows that[21] $\mathrm{tr}\boldsymbol{\rho}_\varphi^2 = \frac{1}{2} + \frac{1}{2}P^2$, or

$$P^2 = 2\mathrm{tr}\boldsymbol{\rho}_\varphi^2 - 1.$$

Furthermore, we find[22]

$$\begin{aligned}
\hat{\mathrm{e}}^{(1)} \cdot \boldsymbol{\rho}_\varphi \cdot \hat{\mathrm{e}}^{(1)} &= \frac{1}{2}(1 - P\cos 2\varphi), \\
\hat{\mathrm{e}}^{(2)} \cdot \boldsymbol{\rho}_\varphi \cdot \hat{\mathrm{e}}^{(2)} &= \frac{1}{2}(1 + P\cos 2\varphi), \\
\hat{\mathrm{e}}^{(2)} \cdot \boldsymbol{\rho}_\varphi \cdot \hat{\mathrm{e}}^{(1)} &= \frac{1}{2}P\sin 2\varphi,
\end{aligned}$$

so that the polarization angle may be expressed as

$$\tan 2\varphi = \frac{2\hat{\mathrm{e}}^{(2)} \cdot \boldsymbol{\rho}_\varphi \cdot \hat{\mathrm{e}}^{(1)}}{\hat{\mathrm{e}}^{(2)} \cdot \boldsymbol{\rho}_\varphi \cdot \hat{\mathrm{e}}^{(2)} - \hat{\mathrm{e}}^{(1)} \cdot \boldsymbol{\rho}_\varphi \cdot \hat{\mathrm{e}}^{(1)}}.$$

5.2.3 The magnetic moment

The most direct handle on the experimental determination of the projection of the angular momentum of a particle in a given direction is the fact that if an electrically charged particle has an angular momentum, it has a *magnetic*

[21] **Do it.**
[22] **Check these equations.**

moment. Immersed in a magnetic field, it is therefore subject to torques and other forces whose effects are observable, the best known of which is its deflection in a Stern-Gerlach experiment. As we shall discuss later in more detail, the energy of a point particle of charge e, mass M, and orbital angular momentum $\vec{\mathbf{L}} = \hbar\vec{\mathcal{L}}$ in a weak uniform magnetic field $\vec{B}$ is given by $-(e/2Mc)\vec{B}\cdot\vec{\mathbf{L}}$, which means that it has an effective magnetic moment $\vec{\mu} = (e/2Mc)\vec{\mathbf{L}} = (e\hbar/2Mc)\vec{\mathcal{L}}$. If $-e$ and M are the charge and mass, respectively, of the electron, the factor $\mu_0 = (e\hbar/2Mc)$ is called the *Bohr magneton.*[23] If the particle, in addition, has the intrinsic spin angular momentum $\vec{\mathsf{S}}$, this adds the energy $-(e/Mc)\mathsf{S}$, which means that a spin-1/2 particle has an intrinsic magnetic moment $\vec{\mu} = (e/Mc)\vec{\mathsf{S}} = (e\hbar/2Mc)\vec{\sigma}$. *Note the factor of 2, called the g-factor, by which the gyromagnetic ratio (the ratio of the magnitude of the magnetic moment to that of the angular momentum) of the spin contribution differs from the orbital.* We shall return to it again later.

As a result, the magnetic moment of a system of charged particles, such as the electrons in an atom, is given by

$$\vec{\mu} = -\frac{e}{2Mc}(\vec{\mathbf{L}} + 2\vec{\mathsf{S}}) = -\mu_0(\vec{\mathcal{L}} + \sum_i \vec{\sigma}^{(i)}), \tag{5.63}$$

where $\hbar\vec{\mathcal{L}}$ is the total orbital angular momentum, and $\frac{1}{2}\hbar\sum_i \vec{\sigma}^{(i)}$ is the sum of all the spin angular momenta of the electrons. It is important to remember that $\vec{\mu}$ is *not* proportional to the total angular momentum.

The experimentally measured magnetic moment of a charged particle with spin always differs from the Bohr magneton (i.e., its g-factor differs from 2) by a certain amount, which is explained by its interaction with the quantum field. In the case of the electron, the deviation of the g-factor from 2 is extremely small but very accurately predicted by quantum electrodynamics. For the proton—in this instance, the Bohr magneton has to be replaced by the *nuclear magneton*, which is about 1846 times smaller because of the mass ratio—the deviation of the g-factor from 2 is quite large ($g \simeq 2.79$) and roughly predicted by the field theory of the strong interactions, where precise calculations are difficult. The same field-theoretic mechanism that leads to such anomalous magnetic moments of charged particles also induces magnetic moments in some neutral ones, so that, as a result of its strong interaction with the chromodynamic field producing quarks, the neutron has a magnetic moment of about -1.9 nuclear magnetons.

[23] Because of its measurement by means of a magnetic field, the eigenvalue m of the z-component of the angular momentum is usually called the *magnetic* quantum number.

5.2.4 Addition of angular momenta

Suppose two independent physical systems, one in a state $\Psi^{(1)}(j_1, m_1)$ with angular momentum quantum numbers j_1, m_1, and the other in the state $\Psi^{(2)}(j_2, m_2)$, are combined, so that the system as a whole is in the state $\Psi(j_1, j_2; m_1, m_2) = \Psi^{(1)}(j_1, m_1) \otimes \Psi^{(2)}(j_2, m_2)$. This state may alternatively be written as a superposition of states of various total angular momenta

$$\Psi(j_1, j_2; m_1, m_2) = \sum_{j,m} \langle j, m | j_1, m_1; j_2, m_2 \rangle \Psi(j, m), \tag{5.64}$$

in which the numbers

$$\langle j, m | j_1, m_1; j_2, m_2 \rangle = (\Psi(j, m), \Psi(j_1, j_2; m_1, m_2))$$

are called *Clebsch-Gordan coefficients.*[24] Since all the components of $\vec{\mathbf{J}}^{(1)}$ commute with all the components of $\vec{\mathbf{J}}^{(2)}$ (the systems are assumed independent), and $\mathbf{J}_z = \mathbf{J}_z^{(1)} + \mathbf{J}_z^{(2)}$, we must have $m = m_1 + m_2$, which implies, first of all, that if both j_1 and j_2 are either integral or half-integral, then j must be integral; otherwise j is half-integral; second, $|j_1 - j_2| \leq j \leq j_1 + j_2$, because $|m| \leq j$, $|m_1| \leq j_1$, and $|m_2| \leq j_2$. If we think of the angular momentum as a classical vector (as in the "old quantum theory"), then the inequalities $|j_1 - j_2| \leq j \leq j_1 + j_2$ simply restate the triangle inequalities

$$||\vec{\mathbf{J}}^{(1)}| - |\vec{\mathbf{J}}^{(2)}|| \leq |\vec{\mathbf{J}}^{(1)} - \vec{\mathbf{J}}^{(2)}| \leq |\vec{\mathbf{J}}^{(1)}| + |\vec{\mathbf{J}}^{(2)}|$$

of the old vector-addition model. Consequently, we have

$$\Psi(j_1, j_2; m_1, m_2) = \sum_{j=|j_1-j_2|}^{j_1+j_2} \langle j, m_1 + m_2 | j_1, m_1; j_2, m_2 \rangle \Psi(j, m_1 + m_2). \tag{5.65}$$

If the basis functions are normalized, the matrix of the Clebsch-Gordan coefficients has to be unitary and $\langle j_1, m_1; j_2, m_2 | j, m \rangle = \langle j, m | j_1, m_1; j_2, m_2 \rangle^*$ so that we also have

$$\Psi(j, m) = \sum_{m_1+m_2=m} \langle j, m | j_1, m_1; j_2, m_2 \rangle^* \Psi(j_1, j_2; m_1, m_2). \tag{5.66}$$

[Note that j_1 and j_2 are not summed over in (5.66).] Expressions for the Clebsch-Gordan coefficients are given in Appendix E, but here we shall construct some of them.

---

[24] Other names used for the Clebsch-Gordan coefficients in the physics literature are *Gaunt coefficients*, *Slater coefficients*, and *Wigner coefficients.*

Two spin-1/2 systems

A system consisting of two spin-1/2 particles can be in four states, with each of the individual particles in states $\chi^{1/2}_{1/2}$ or $\chi^{1/2}_{-1/2}$, making up the two-particle states $\chi^{1/2(1)}_{\pm 1/2} \otimes \chi^{1/2(2)}_{\pm 1/2}$. Using these states, we can form two-particle states with total angular momentum $j = 1$ or $j = 0$ $(2 \cdot 2 = 3 + 1)$; for obvious reasons, the first is called a *triplet,* the second, a *singlet.* In order to construct these states out of the individual spinors (assumed normalized), we start by noting that if, in the triplet, $m = -1$, it is necessary for both m_1 and m_2 to have the values $-1/2$, which means that $\chi^1_{-1} = \chi^{1/2(1)}_{-1/2} \otimes \chi^{1/2(2)}_{-1/2}$; similarly we must have $\chi^1_1 = \chi^{1/2(1)}_{1/2} \otimes \chi^{1/2(2)}_{1/2}$. In order to construct χ^1_0, we form the operators j_+ as in (5.7), so that $\mathsf{j}_+ = \mathsf{j}^{(1)}_+ + \mathsf{j}^{(2)}_+$. Equation (5.11) then tells us that

$$\begin{aligned} \sqrt{2}\chi^1_0 = \mathsf{j}_+\chi^1_{-1} &= (\mathsf{j}^{(1)}_+ + \mathsf{j}^{(2)}_+)\chi^{1/2(1)}_{-1/2} \otimes \chi^{1/2(2)}_{-1/2} \\ &= \chi^{1/2(1)}_{1/2} \otimes \chi^{1/2(2)}_{-1/2} + \chi^{1/2(1)}_{-1/2} \otimes \chi^{1/2(2)}_{1/2}. \end{aligned}$$

Thus the three triplet states are given by

$$\begin{aligned} \chi^1_{-1} &= \chi^{1/2(1)}_{-1/2} \otimes \chi^{1/2(2)}_{-1/2}, \\ \chi^1_0 &= \frac{1}{\sqrt{2}}(\chi^{1/2(1)}_{1/2} \otimes \chi^{1/2(2)}_{-1/2} + \chi^{1/2(1)}_{-1/2} \otimes \chi^{1/2(2)}_{1/2}), \\ \chi^1_1 &= \chi^{1/2(1)}_{1/2} \otimes \chi^{1/2(2)}_{1/2}. \end{aligned} \tag{5.67}$$

The singlet state has to be orthogonal to all three of the triplet, which allows us to conclude that it must be

$$\chi^0 = \frac{1}{\sqrt{2}}(\chi^{1/2(1)}_{1/2} \otimes \chi^{1/2(2)}_{-1/2} - \chi^{1/2(1)}_{-1/2} \otimes \chi^{1/2(2)}_{1/2}). \tag{5.68}$$

Notice the important fact that *the singlet state is anti-symmetric* under an interchange of the two spins, while *all three components of the triplet state are symmetric.*

If a system consisting of two-particles of spin-1/2 is in the singlet state (5.68), each of the one-particle subsystems alone, that is, with the other one ignored, is in a mixed state whose spin density operator is given by[25]

$$\boldsymbol{\rho} = \frac{1}{2}(\mathsf{P}_+ + \mathsf{P}_-) = \frac{1}{2}\mathbb{1},$$

where $\mathsf{P}_\pm$ are the projections on the spin-up and spin-down states. It is as "incoherent" as is possible for a spin-1/2 state to be, with $\mathrm{tr}\boldsymbol{\rho}^2 = \frac{1}{2}$.[26]

[25] **Prove this.**

[26] **As an exercise, prove that for any spin-1/2 system,** $\mathrm{tr}\boldsymbol{\rho}^2 \geq \frac{1}{2}$.

On the other hand, if the other particle is not ignored, the *entanglement* of the states $\chi^{1/2(1)}_{-1/2} \otimes \chi^{1/2(2)}_{1/2}$ and $\chi^{1/2(1)}_{1/2} \otimes \chi^{1/2(2)}_{-1/2}$ in the expression (5.68) has important physical consequences. For example, if the two-particle system is in the state χ^0, what is the probability that a measurement of the variables A on particle #1 and B on particle #2 will yield the results A and B, respectively? Denoting the two eigenstates by $\chi^{(1)}_A$ and $\chi^{(2)}_B$, this probability is

$$\begin{aligned} P(A,B) &= |(\chi^0, \chi^{(1)}_A \otimes \chi^{(2)}_B)|^2 \\ &= \frac{1}{2}|(\chi^{1/2}_{1/2}, \chi_A)(\chi^{1/2}_{-1/2}, \chi_B) - (\chi^{1/2}_{-1/2}, \chi_A)(\chi^{1/2}_{1/2}, \chi_B)|^2, \end{aligned}$$

which differs from the product of the two probabilities of finding A for particle #1 and B for #2. The result of the measurement of A on particle #1 thus depends on the result of the measurement of B on particle #2, even though the two particles may be very far apart. Specifically, if two spin-1/2 particles are the decay products of a spin-0 parent system, they fly off in opposite directions in a singlet state like (5.68), and their entanglement leads to the kind of correlation described by Bohm in his version of the EPR *Gedanken* experiment discussed in Chapter 1; it allows us to infer the spin projection of one of them by measuring that of the other.

The singlet and triplet states of two spin-1/2 particles could also have been constructed by exploiting their opposite exchange properties, in the following way. Suppose a system of two spin-1/2 particles is in the state $\chi^{1/2(1)}_a \otimes \chi^{1/2(2)}_b$. Then the operator

$$\mathsf{P}_{\text{exch}} \stackrel{\text{def}}{=} \frac{1}{2}[\mathbf{1} + \vec{\sigma}^{(1)} \cdot \vec{\sigma}^{(2)}], \tag{5.69}$$

called the *spin exchange operator* switches the two spin states, in the sense that[27]

$$\mathsf{P}_{\text{exch}} \chi^{1/2(1)}_a \otimes \chi^{1/2(2)}_b = \chi^{1/2(1)}_b \otimes \chi^{1/2(2)}_a.$$

Since two such exchanges lead back to the original state, it follows that $\mathsf{P}^2_{\text{exch}} = \mathbf{1}$, which implies that its two eigenvalues are ± 1. The eigenstate with the eigenvalue 1 is the triplet state of the two spin-1/2 particles, and the state with the eigenvalue -1 is the singlet.[28]

Spin-1/2 plus orbital angular momentum

The wave function of a spin-0 particle subject to a rotationally invariant Hamiltonian, we have seen, can be written as a superposition of products of a radial function and a spherical harmonic, fixing the magnitude and the z-component of its orbital angular momentum, $R_l(r)Y^{m_l}_l(\theta, \varphi)$. If the particle

[27] **Prove this.**

[28] **As an exercise, construct the four states in this manner.**

has spin-1/2, each such product wave function has to be multiplied by a spinor $\chi^{1/2}_{m_s}$, so that, apart from the energy, the wave function is labeled by (l, m_l, m_s) (the additional, fixed label 1/2 for the magnitude of the spin being understood[29]). Rotational invariance of the Hamiltonian now, however, does not necessarily entail conservation of the orbital angular momentum, so that l and m_l need not be good quantum numbers; instead, the good quantum numbers are the total angular momentum j and its z-component $m = m_l + m_s$. It is therefore useful to construct angle-dependent spinors that are eigenfunctions of $\vec{\mathbf{J}}^2 = (\vec{\mathbf{L}} + \vec{\mathbf{S}})^2$ and of $\mathbf{J}_z = \mathbf{L}_z + \mathbf{S}_z$, if by $\vec{\mathbf{S}}$ we mean the spin angular momentum operator.[30]

For convenience, define $\vec{\mathcal{J}} \stackrel{\text{def}}{=} \vec{\mathbf{J}}/\hbar = \vec{\mathcal{L}} + \frac{1}{2}\vec{\sigma}$, so that

$$\vec{\mathcal{J}}^2 = \vec{\mathcal{L}}^2 + \vec{\mathcal{L}} \cdot \vec{\sigma} + \frac{1}{4}\vec{\sigma}^2 = \vec{\mathcal{L}}^2 + \vec{\mathcal{L}} \cdot \vec{\sigma} + \frac{3}{4}. \tag{5.70}$$

Both $\vec{\mathcal{L}}^2$ and $\vec{\mathcal{L}} \cdot \vec{\sigma}$ are rotationally invariant and therefore commute with $\vec{\mathbf{J}}^2$ and $\mathbf{J}_z$; they therefore have simultaneous eigenvectors with $\vec{\mathbf{J}}^2$ and $\mathbf{J}_z$. It will be one of your homework assignments to show that for the states with $j = l + \frac{1}{2}$, $\vec{\mathcal{L}} \cdot \vec{\sigma}$ has the eigenvalue l, and for the states with $j = l - \frac{1}{2}$, $\vec{\mathcal{L}} \cdot \vec{\sigma}$ has the eigenvalue $-l - 1$. We may thus define two projection operators

$$\mathsf{P}_+ \stackrel{\text{def}}{=} [(l+1)\mathbb{1} + \vec{\mathcal{L}} \cdot \vec{\sigma}]/(2l+1), \quad \mathsf{P}_- \stackrel{\text{def}}{=} [l\mathbb{1} - \vec{\mathcal{L}} \cdot \vec{\sigma}]/(2l+1), \quad \mathsf{P}_+ + \mathsf{P}_- = \mathbb{1} \tag{5.71}$$

so that $\mathsf{P}_+ = 1$ when acting on a state with $j = l + \frac{1}{2}$ and $\mathsf{P}_+ = 0$ when acting on a state with $j = l - \frac{1}{2}$, while $\mathsf{P}_- = 1$ when acting on a state with $j = l - \frac{1}{2}$ but $\mathsf{P}_- = 0$ when acting on a state with $j = l + \frac{1}{2}$.

These two projections may be used to construct spin functions $\mathfrak{Z}^m_{l,j}$ with $j = l + \frac{1}{2}$ and with $j = l - \frac{1}{2}$,

$$P_+ Y_l^{m-1/2}(\theta, \varphi)\chi^{1/2}_{1/2} = A\mathfrak{Z}^m_{l,l+\frac{1}{2}}(\theta, \varphi),$$

$$P_- Y_l^{m-1/2}(\theta, \varphi)\chi^{1/2}_{1/2} = B\mathfrak{Z}^m_{l,l-\frac{1}{2}}(\theta, \varphi).$$

In order for these $\mathfrak{Z}^m_{l,j}$ to be normalized, A and B have to be chosen (apart from an arbitrary phase factor),[31]

$$A = \sqrt{\frac{1}{2} + \frac{m}{2l+1}}, \quad B = \sqrt{\frac{1}{2} - \frac{m}{2l+1}}. \tag{5.72}$$

---

[29] If we are talking about the Schrödinger equation of a system consisting of two spin-1/2 particles, on the other hand, the total spin may be 0 or 1, and these two states may be coupled.

[30] This is analogous to the need in electromagnetic theory for constructing *vector spherical harmonics*, which are eigenfunctions of the total angular momentum made up of the "orbital" angular momentum and spin-1 functions, in accordance with the fact that the photon has spin-1 (i.e., the electromagnetic field is a vector field).

[31] **Prove this as an exercise.**

For these calculations it is useful to introduce $\sigma_\pm \stackrel{\text{def}}{=} \frac{1}{2}(\sigma_x \pm i\sigma_y)$ and $\mathcal{L}_\pm \stackrel{\text{def}}{=} \mathcal{L}_x \pm i\mathcal{L}_y$, so that $\vec{\mathcal{L}} \cdot \vec{\sigma} = \mathcal{L}_z\sigma_z + \mathcal{L}_+\sigma_- + \mathcal{L}_-\sigma_+$ and $\sigma_+\chi^{1/2}_{-1/2} = \chi^{1/2}_{1/2}$, while $\sigma_-\chi^{1/2}_{1/2} = \chi^{1/2}_{-1/2}$.

As a result of evaluating the actions of $\mathsf{P}_\pm$ on $Y_l^{m-1/2}(\theta,\varphi)\chi^{1/2}_{1/2}$ we then find that[32]

$$\begin{aligned}
\mathfrak{Z}^m_{l,l+\frac{1}{2}} &= \sqrt{\frac{1}{2} + \frac{m}{2l+1}}\, Y_l^{m-\frac{1}{2}}\chi^{1/2}_{1/2} \\
&\quad + \sqrt{\frac{1}{2} - \frac{m}{2l+1}}\, Y_l^{m+\frac{1}{2}}\chi^{1/2}_{-1/2}, \qquad (5.73)\\
\mathfrak{Z}^m_{l,l-\frac{1}{2}} &= \sqrt{\frac{1}{2} - \frac{m}{2l+1}}\, Y_l^{m-\frac{1}{2}}\chi^{1/2}_{1/2} \\
&\quad - \sqrt{\frac{1}{2} + \frac{m}{2l+1}}\, Y_l^{m+\frac{1}{2}}\chi^{1/2}_{-1/2}. \qquad (5.74)
\end{aligned}$$

5.2.5 Spherical tensors and selection rules

Consider an operator $\mathcal{T}^0_0$ that is invariant under active rotations, so that it commutes with all components of the angular momentum. If $\mathcal{O}_R$ is a rotation operator, we then have (assuming the irreducible representations D^j and $D^{j'}$ are unitary),

$$\begin{aligned}
\langle \alpha' j' m' | \mathcal{T}^0_0 | \alpha j m\rangle &= \langle \alpha' j' m' | \mathcal{O}_R \mathcal{T}^0_0 \mathcal{O}_R^{-1} | \alpha j m\rangle \\
&= \sum_{nn'} D^{j'}_{m'n'}(R) \langle \alpha' j' n' | \mathcal{T}^0_0 | \alpha j n\rangle D^{j*}_{mn}(R),
\end{aligned}$$

which implies that

$$\sum_n \langle \alpha' j' m' | \mathcal{T}^0_0 | \alpha j n\rangle D^j_{nm}(R) = \sum_{n'} D^{j'}_{m'n'}(R) \langle \alpha' j' n' | \mathcal{T}^0_0 | \alpha j m\rangle.$$

Lemmas (E.2) and (E.3) (Schur's lemma) in Appendix E therefore allow us to draw the conclusion that $\langle \alpha' j' m' | \mathcal{T}^0_0 | \alpha j m\rangle = 0$ unless $j = j'$ and $m = m'$. One says that for rotationally invariant operators, *the selection rule is* $\Delta j = 0$ *and* $\Delta m = 0$*;* moreover, *the nonvanishing matrix elements are independent of* m (so that $\langle \alpha' j m' | \mathcal{T}^0_0 | \alpha j m\rangle$ is not just diagonal, but a multiple of the unit matrix in spin space).

[32] **Show this.**

Next take a Cartesian vector operator $\vec{\mathbf{G}}$ and its commutation relation (5.5) with the angular momentum. Applying (5.5) twice yields[33]

$$[\vec{\mathbf{G}}, \vec{\mathbf{J}}^2] = i\hbar(\vec{\mathbf{J}} \times \vec{\mathbf{G}} - \vec{\mathbf{G}} \times \vec{\mathbf{J}}).$$

Now using the fact that the left-hand side of this equation is a vector operator, substitute $[\vec{\mathbf{G}}, \vec{\mathbf{J}}^2]$ for $\vec{\mathbf{G}}$, and you find[34]

$$[[\vec{\mathbf{G}}, \vec{\mathbf{J}}^2], \vec{\mathbf{J}}^2] = 2\hbar^2(\vec{\mathbf{J}}^2\vec{\mathbf{G}} + \vec{\mathbf{G}}\vec{\mathbf{J}}^2) - 4\hbar^2\vec{\mathbf{J}}(\vec{\mathbf{G}} \cdot \vec{\mathbf{J}}). \tag{5.75}$$

Taking matrix elements of this equation for $j \neq j'$, and using other needed quantum numbers α that are eigenvalues of a rotationally invariant operator, we obtain

$$\langle \alpha' j' m' | [[\vec{\mathbf{G}}, \vec{\mathbf{J}}^2], \vec{\mathbf{J}}^2] - 2\hbar^2(\vec{\mathbf{J}}^2\vec{\mathbf{G}} + \vec{\mathbf{G}}\vec{\mathbf{J}}^2) | \alpha j m \rangle = 0,$$

which, after a little algebra[35], leads to

$$[(j - j')^2 - 1][(j + j' + 1)^2 - 1]\langle \alpha' j' m' | \vec{\mathbf{G}} | \alpha j m \rangle = 0.$$

Therefore the *selection rule for matrix elements of a vector operator* is

$$j' = j \pm 1.$$

For $j' = j$, on the other hand, (5.75) yields

$$\begin{aligned} \langle \alpha' j m' | \vec{\mathbf{G}} | \alpha j m \rangle &= \frac{\langle \alpha' j m' | \vec{\mathbf{J}}\vec{\mathbf{G}} \cdot \vec{\mathbf{J}} | \alpha j m \rangle}{\hbar^2 j(j+1)} \\ &= \frac{1}{\hbar^2 j(j+1)} \sum_{\alpha'' j'' m''} \langle \alpha' j m' | \vec{\mathbf{J}} | \alpha'' j'' m'' \rangle \langle \alpha'' j'' m'' | \vec{\mathbf{G}} \cdot \vec{\mathbf{J}} | \alpha j m \rangle. \end{aligned}$$

But since the α are eigenvalues of a rotationally invariant operator, $\vec{\mathbf{J}}$ is diagonal in α as well as in j, and since $\vec{\mathbf{G}} \cdot \vec{\mathbf{J}}$ is invariant under rotations, it is not only diagonal in m but the matrix element is independent of m. As a result we obtain

$$\langle \alpha' j m' | \vec{\mathbf{G}} | \alpha j m \rangle = \mathcal{F}(\alpha', \alpha, j) \langle j m' | \vec{\mathbf{J}} | j m \rangle, \tag{5.76}$$

where

$$\mathcal{F}(\alpha', \alpha, j) = \frac{\langle \alpha' j m | \vec{\mathbf{G}} \cdot \vec{\mathbf{J}} | \alpha j m \rangle}{\hbar^2 j(j+1)}, \tag{5.77}$$

so that all the specific dependence of the left-hand side of (5.76) upon m and m' is contained in the matrix element of $\vec{\mathbf{J}}$ and does not depend on $\vec{\mathbf{G}}$.

---

[33] **Show this as an exercise.**
[34] **Do this as an exercise.**
[35] **Do it.**

To obtain the selection rule for Δm when $j' \neq j$, we have, of course, $\langle \alpha' j' m' | \mathsf{G}_z | \alpha j m \rangle = 0$ unless $m = m'$, because $[\mathbf{G}_z, \mathbf{J}_z] = 0$. Using (5.5), we find that $[\mathsf{j}_z, \mathbf{G}_x + i\mathbf{G}_y] = \mathbf{G}_x + i\mathbf{G}_y$, and therefore

$$\langle \alpha' j' m' | \mathbf{G}_x + i\mathbf{G}_y | \alpha j m \rangle = (m' - m) \langle \alpha' j' m' | \mathbf{G}_x + i\mathbf{G}_y | \alpha j m \rangle,$$

implying that $\langle \alpha' j' m' | \mathbf{G}_x + i\mathbf{G}_y | \alpha j m \rangle = 0$ unless $m' = m + 1$; similarly we find that $\langle \alpha' j' m' | \mathbf{G}_x - i\mathbf{G}_y | \alpha j m \rangle = 0$ unless $m' = m - 1$. It follows from these two results that the matrix elements of $\mathbf{G}_x$ and $\mathbf{G}_y$ vanish unless $\Delta m = \pm 1$. What is more, because $[(\mathbf{G}_x + i\mathbf{G}_y), \mathsf{j}_+] = 0$, it follows from (5.11) and (5.12) that

$$\frac{\langle \alpha' j' m + 2 | \mathbf{G}_x + i\mathbf{G}_y | \alpha j m + 1 \rangle}{\sqrt{(j' - m - 1)(j' + m + 2)}} = \frac{\langle \alpha' j' m + 1 | \mathbf{G}_x + i\mathbf{G}_y | \alpha j m \rangle}{\sqrt{(j' - m)(j' + m + 1)}}. \quad (5.78)$$

As a specific physical application, consider expectation values of the magnetic moment of an atom, as given by (5.63). Assuming the Hamiltonian of the atom to be rotationally invariant, we have, according to (5.76)

$$\langle \vec{\mu} \rangle = -\frac{e}{2mc} \langle E j m | (\vec{\mathbf{L}} + 2\vec{\mathsf{S}}) | E j m \rangle = \frac{\langle j m | \vec{\mathbf{J}} | j m \rangle}{\hbar j} \mu_{\text{eff}}, \quad (5.79)$$

where the effective "magnetic moment of the atom" is given by

$$\mu_{\text{eff}} = \frac{\langle E j m | \vec{\mu} \cdot \vec{\mathbf{J}} | E j m \rangle}{\hbar(j + 1)}. \quad (5.80)$$

But from $\vec{\mathbf{L}}^2 = (\vec{\mathbf{J}} - \vec{\mathsf{S}})^2$ and $\vec{\mathsf{S}}^2 = (\vec{\mathbf{J}} - \vec{\mathbf{L}})^2$ we find $\vec{\mathsf{S}} \cdot \vec{\mathbf{J}} = \frac{1}{2}(\vec{\mathbf{J}}^2 + \vec{\mathsf{S}}^2 - \vec{\mathbf{L}}^2)$ and $\vec{\mathbf{L}} \cdot \vec{\mathbf{J}} = \frac{1}{2}(\vec{\mathbf{J}}^2 + \vec{\mathbf{L}}^2 - \vec{\mathsf{S}}^2)$, so that by (5.63)

$$\vec{\mu} \cdot \vec{\mathbf{J}} = -\frac{e}{2Mc}(\vec{\mathbf{L}} \cdot \vec{\mathbf{J}} + 2\vec{\mathsf{S}} \cdot \vec{\mathbf{J}}) = -\frac{e}{2Mc}(\frac{3}{2}\vec{\mathbf{J}}^2 + \frac{1}{2}\vec{\mathsf{S}}^2 - \frac{1}{2}\vec{\mathbf{L}}^2)$$

and therefore

$$\mu_{\text{eff}} = -\frac{e}{2Mc} \frac{\langle j m | \frac{3}{2}\vec{\mathbf{J}}^2 + \frac{1}{2}\vec{\mathsf{S}}^2 - \frac{1}{2}\vec{\mathbf{L}}^2 | j m \rangle}{\hbar(j + 1)}.$$

In the so-called *LS coupling scheme* it is a good approximation to assume that all the electronic orbital angular momenta couple to a total $\vec{\mathbf{L}}$ and all the spins couple to a total spin angular momentum $\vec{\mathsf{S}}$, which then couple to a total $\vec{\mathbf{J}} = \vec{\mathbf{L}} + \vec{\mathsf{S}}$. In that approximation therefore

$$\mu_{\text{eff}} = -\mu_0 \, j \left[1 + \frac{j(j + 1) - l(l + 1) + s(s + 1)}{2j(j + 1)} \right]. \quad (5.81)$$

The bracket is called the *Landé g-factor* for LS coupling.

Spherical tensors

As already mentioned for the case of vectors, for quantum-mechanical purposes it is more convenient to classify tensors in terms of their irreducible transformation properties under rotations, rather than in terms of their Cartesian transformation properties, particularly if the components of the tensor are operators. So if we are given, for instance, a Cartesian tensor of rank 2, $\mathcal{T}_{\mu\nu}$, then we know it transforms under a rotation of the coordinate system like $x_\mu x_\nu$, that is, $\mathcal{T}'_{\mu\nu} = \sum_{\kappa\sigma} T_{\kappa\mu} T_{\sigma\nu} \mathcal{T}_{\kappa\sigma}$, where T is given by (5.29). On the other hand, (5.40) tells us that

$$\frac{1}{r} x_\mu = i\sqrt{4\pi/3} \sum_m \mathcal{M}^*_{m\mu} Y_1^m.$$

Therefore $\sum_{\mu\nu} \mathcal{M}_{m_1\mu} \mathcal{M}_{m_2\nu} \mathcal{T}_{\mu\nu}$ transforms like $Y_1^{m_1} \otimes Y_1^{m_2}$, and we have

$$Y_1^{m_1} \otimes Y_1^{m_2} = \sum_{l=0}^{2} \langle l, m_1 + m_2 | l_1, m_1; l_2, m_2 \rangle^* Y_l^{m_1+m_2}.$$

The simple equation $3 \cdot 3 = 5 + 3 + 1$ suggests that a general tensor of rank 2, which has $3 \cdot 3 = 9$ components, can be decomposed into a direct sum of parts that transform according to the angular momenta 2, 1, and 0. These components are obtained by means of (5.43) and the Clebsch-Gordan coefficients as

$$\hat{\mathcal{T}}^j_m = \sum_{m_1 m_2} \langle j, m | 1, m_1; 1, m_2 \rangle \sum_{\mu\nu} \mathcal{M}_{m_1\mu} \mathcal{M}_{m_2\nu} \mathcal{T}_{\mu\nu}, \quad j = 0, 1, 2. \tag{5.82}$$

We can easily recognize the three irreducible pieces of a Cartesion tensor of rank 2 as follows. First, split the tensor (in an invariant manner) into its symmetric and anti-symmetric parts: $\mathcal{T}_{\mu\nu} = \mathcal{T}^{(\mathrm{s})}_{\mu\nu} + \mathcal{T}^{(\mathrm{a})}_{\mu\nu}$, where $\mathcal{T}^{(\mathrm{s})}_{\mu\nu} \stackrel{\mathrm{def}}{=} \frac{1}{2}(\mathcal{T}_{\mu\nu} + \mathcal{T}_{\nu\mu})$ and $\mathcal{T}^{(\mathrm{a})}_{\mu\nu} \stackrel{\mathrm{def}}{=} \frac{1}{2}(\mathcal{T}_{\mu\nu} - \mathcal{T}_{\nu\mu})$. Furthermore, the trace of the tensor, $\mathcal{T}_0 \stackrel{\mathrm{def}}{=} \sum_\mu \mathcal{T}_{\mu\mu}$ is invariant under rotations, which implies that $\mathcal{T}_0$ transforms like Y_0^0 and thus corresponds to $j = 0$; $\mathcal{T}^{(\mathrm{a})}$ has three components, and, as you know, these transform under rotations like the (axial) vector $\mathcal{T}_\mu \stackrel{\mathrm{def}}{=} \sum_{\kappa\sigma} \epsilon_{\mu\kappa\sigma} \mathcal{T}_{\kappa\sigma}$.[36] We are then left with a traceless symmetric 5-component piece $\mathcal{T}^{(\mathrm{st})}_{\mu\nu} \stackrel{\mathrm{def}}{=} \mathcal{T}^{(\mathrm{s})}_{\mu\nu} - \frac{1}{3}\mathcal{T}_0 \delta_{\mu\nu}$, which corresponds to the angular momentum $j = 2$. (This is the reason why the *graviton,* the quantum of the gravitational field, which is described in the general theory of relativity by a symmetric tensor of rank 2, has *spin 2.*)

The principal reason why the classification of tensor[37] operators according to their irreducible transformation properties under rotations is more

[36] $\epsilon_{\mu\kappa\sigma}$ is the totally anti-symmetric tensor with $\epsilon_{123} = 1$.

[37] Including vectors, which are tensors of rank 1.

convenient for quantum mechanics than their classification according to their Cartesian transformation properties is that it allows us to calculate more easily the matrix elements of such operators between eigenstates of the total angular momentum, and in particular, to ascertain when such matrix elements vanish. As we shall discuss in more detail later in Chapter 8, the moduli squared of such matrix elements determine the *transition probabilities* from one of these states to another, caused by a perturbation of a rotationally invariant Hamiltonian by a tensor operator. If the matrix element vanishes, the corresponding transition is *forbidden,* i.e., in the first approximation (or possibly, exactly) its probability is equal to zero; the lists of such forbidden transitions are called selection rules, some of which we have already derived for scalars and Cartesian vectors.

To see the advantage of using spherical rather than Cartesian tensors, consider a matrix element of a tensor operator such as $\langle \alpha' j' m' | \mathcal{T}^l_M | \alpha j m \rangle$, where α denotes all the other quantum numbers (such as the energy, for example) needed to specify the state uniquely. Under a rotation R, implemented in the Hilbert space by the operator $\mathcal{O}_R$, we have

$$\begin{aligned} \langle \alpha' j' m' | \mathcal{T}^{l\prime}_M | \alpha j m \rangle &= \langle \alpha' j' m' | \mathcal{O}_R \mathcal{T}^l_M \mathcal{O}^\dagger_R | \alpha j m \rangle \\ &= \sum_{M'} \langle \alpha' j' m' | \mathcal{T}^l_{M'} | \alpha j m \rangle D^l_{M'M}(R). \end{aligned}$$

However, we also have

$$\mathcal{O}_R \Psi_{\alpha j m} = \sum_n \Psi_{\alpha j n} D^j_{nm}(R),$$

and

$$\mathcal{O}^{-1}_R \Psi_{\alpha j m} = \mathcal{O}^\dagger_R \Psi_{\alpha j m} = \sum_n \Psi_{\alpha j n} D^{j*}_{mn}(R),$$

so that

$$\sum_{M'} \langle \alpha' j' m' | \mathcal{T}^l_{M'} | \alpha j m \rangle D^l_{M'M}(R) = \sum_{nn'} D^{j'}_{m'n'}(R) \langle \alpha' j' n' | \mathcal{T}^l_{M'} | \alpha j n \rangle D^{j*}_{mn}(R). \tag{5.83}$$

On the other hand, we have for the Clebsch-Gordan coefficients

$$\begin{aligned} \langle j' m' | j, m; l, M \rangle &= (\Psi(j', m'), \mathcal{O}_R \mathcal{O}^{-1}_R \Psi(j, m) \otimes \Psi(l, M)) \\ &= \sum_{nn'M'} D^{j'}_{m'n'}(R) \langle j' n' | j, n; l, M' \rangle D^{j*}_{mn}(R) D^{l*}_{MM'}(R), \end{aligned}$$

and therefore, using the unitarity of the D's,

$$\sum_{M'} \langle j' m' | j, m; l, M' \rangle D^l_{M'M}(R) = \sum_{nn'} D^{j'}_{m'n'}(R) \langle j' n' | j, n; l, M \rangle D^{j*}_{mn}(R). \tag{5.84}$$

Comparison of (5.84) with (5.83) shows the two relations between the D matrices to be the same, which means that the coefficients must be proportional; denote the constant of proportionality, called the *reduced matrix element*, by $\langle \alpha' j' \parallel \mathcal{T}^l \parallel \alpha j \rangle$, and you obtain

$$\langle \alpha' j' m' | \mathcal{T}^l_M | \alpha j m \rangle = \langle j', m' | j, m; l, M \rangle \langle \alpha' j' \parallel \mathcal{T}^l \parallel \alpha j \rangle. \tag{5.85}$$

This important result is called the **Wigner-Eckart theorem,** a special case of which was (5.76). It not only tells us that all the dependence on the variables m, m', and M is contained in the Clebsch-Gordan coefficients, but it implies the *selection rules*

$$\Delta j \leq l, \qquad \Delta m = M \leq l, \tag{5.86}$$

meaning that unless $|j - j'| \leq l$ and $m' - m = M$, the matrix element vanishes and the transition is forbidden. (We shall return to this.)

5.3 Time Reversal

5.3.1 The time-reversal operator

What would happen if the direction of time were reversed, as in a video tape run backwards? Since the directions of all motions are reversed, it would be appropriate to call pressing the back button *motion reversal,* but it is customarily designated as "time reversal."

Consider first the dynamical variables of a particle system: it is clear that under time reversal, the position $\vec{\mathbf{q}}$ remains unchanged, while the momentum $\vec{\mathbf{p}}$ changes sign. Analogously for other dynamical variables: the angular momenta $\vec{\mathbf{L}}$ and $\vec{\mathbf{S}}$, as well as the total angular momentum $\vec{\mathbf{J}}$, change sign. If there is an electromagnetic field, the Maxwell equations require that the electric field, as well as the scalar potential, remain invariant, whereas the magnetic field and the vector potential change sign. Supposing the Hamiltonian of the system to be $\mathbf{H}$, we shall call the time-reversed Hamiltonian, obtained by making all the appropriate changes in the dynamical variables and fields, $\mathbf{H}^{\mathrm{TR}}$, so that under time reversal,

$$\mathbf{H} \to \mathbf{H}^{\mathrm{TR}}.$$

Under a reversal of the direction of time, $t \to t' = -t$, the Schrödinger equation

$$i\hbar \frac{\partial \Psi(t)}{\partial t} = \mathbf{H}\Psi(t),$$

becomes

$$-i\hbar \frac{\partial \Psi^{\mathrm{TR}}(t')}{\partial t'} = \mathbf{H}^{\mathrm{TR}} \Psi^{\mathrm{TR}}(t').$$

Let us then define the *time reversal operator* ϑ by

$$\vartheta\Psi(t) \stackrel{\text{def}}{=} \Psi^{\text{TR}}(t), \tag{5.87}$$

so that $\vartheta\vec{\mathbf{q}}\vartheta^{-1} = \vec{\mathbf{q}}$, $\vartheta\vec{\mathbf{p}}\vartheta^{-1} = -\vec{\mathbf{p}}$, etc., and in particular,

$$\mathbf{H}^{\text{TR}} = \vartheta\mathbf{H}\vartheta^{-1}; \tag{5.88}$$

the time-reversed Schrödinger equation therefore reads

$$-i\hbar\vartheta\frac{\partial\Psi(t)}{\partial t} = \vartheta\mathbf{H}\Psi(t). \tag{5.89}$$

Now, since two successive time reversals clearly cancel each other out, ϑ^2 should be a multiple of the unit operator. But remember that $\vartheta\vec{\sigma}\vartheta^{-1} = -\vec{\sigma}$, which means that for the spin, time reversal is equivalent to a rotation by π; therefore ϑ^2 is equivalent to a rotation by 2π, and spinors change sign under such a rotation. Consequently, for states that contain n spin-1/2 degrees of freedom, the action of ϑ^2 results in multiplication my $(-1)^n$:

$$\vartheta^2 = \alpha\mathbf{1}, \qquad \alpha = (-1)^n. \tag{5.90}$$

Furthermore, multiplication of (5.89) by ϑ and comparison with the original Schrödinger equation leads to the conclusion that

$$\vartheta i = -i\vartheta.$$

According to the Wigner theorem mentioned earlier, ϑ has to be either unitary or "antiunitary"; here we see that the time-reversal operator ϑ must be *antiunitary,* so that for all complex numbers c,

$$\vartheta c\Psi = c^*\vartheta\Psi, \tag{5.91}$$

and for any pair of vectors,

$$(\vartheta\Psi, \vartheta\Phi) = (\Phi, \Psi). \tag{5.92}$$

Suppose now that $\mathbf{A}$ is a dynamical variable that is invariant under time reversal, like $\vec{\mathbf{q}}$, for example, and Ψ_A is an eigenvector belonging to the eigenvalue A:

$$\mathbf{A}\Psi_A = A\Psi_A.$$

Then it follows that, since A is real,

$$\vartheta\mathbf{A}\Psi_A = \mathbf{A}\vartheta\Psi_A = A\vartheta\Psi_A,$$

implying that $\vartheta\Psi_A$ is also an eigenvector, with the same eigenvalue. Therefore, if there is no degeneracy,

$$\vartheta\Psi_A = e^{i\gamma}\Psi_A,$$

and the (arbitrary) phase of Ψ_A can always be chosen[38] so that $\gamma = 0$: *if there is no degeneracy, all the eigenvectors of a time-reversal invariant operator can be chosen so that they are simultaneously eigenvectors of ϑ with the eigenvalue 1.* (We shall return to the case of degeneracy below.)

On the other hand, let $\mathbf{B}$ be a dynamical variable that anti-commutes with ϑ, i.e., one that changes sign under time reversal, like $\vec{\mathbf{p}}$, for example, and let Ψ_B be an eigenvector with the eigenvalue B. In that case, we find similarly, that *if B is an eigenvalue, then so must be $-B$, and the eigenvectors can be chosen in such a way that*

$$\vartheta\Psi_B = \Psi_{-B}.$$

These results have an immediate bearing on the behavior of wave functions under time reversal. First, take a configuration-space wave function $\psi(\vec{q}) = (\Psi_{\vec{q}}, \Psi)$. In this case we have

$$\psi^{\mathrm{TR}}(\vec{q}) = (\Psi_{\vec{q}}, \vartheta\Psi) = (\vartheta\Psi_{\vec{q}}, \vartheta\Psi) = (\Psi, \Psi_{\vec{q}}) = \psi^*(\vec{q}).$$

Therefore, if the eigenstates of the operators $\vec{q}$ are appropriately chosen, *time reversal sends every configuration wave function into its complex conjugate, and if it is the wave function of a time-reversal invariant state, it must be real.*

For momentum-space wave functions, with appropriately chosen momentum eigenvectors, we obtain in a similar way

$$\psi^{\mathrm{TR}}(\vec{p}) = (\Psi_{\vec{p}}, \vartheta\Psi) = \psi^*(-\vec{p}), \tag{5.93}$$

which implies that if Ψ is invariant under time reversal, then $\psi(-\vec{p}) = \psi^*(\vec{p})$.

5.3.2 Time-reversal invariance

Next we want to explore the consequences if the Hamiltonian of a system is *time-reversal invariant;* so we now assume that $\vartheta\mathbf{H}\vartheta^{-1} = \mathbf{H}$, or

$$\vartheta\mathbf{H} = \mathbf{H}\vartheta.$$

According to the result shown above, it then follows that all the nondegenerate eigenvectors of $\mathbf{H}$ are invariant under time reversal and their configuration-space eigenfunctions can be chosen to be real. In the same way it follows that the configuration representation of the operator $\mathbf{H}$ itself, the operator that appears in the Schrödinger equation when written in the configuration representation, must also be real. This means, in particular, that if there is a nonlocal potential, i.e., a potential that is not diagonal

[38] **Prove this.**

in configuration space, the integral kernel representing it must not only be Hermitian but real. Any model that includes a complex potential with a non-vanishing imaginary part cannot be time-reversal invariant: it leads to absorption or emission, processes that distinguish one direction of the time from the other.

Degeneracy

Consider now a system whose Hamiltonian is invariant under time reversal,

$$[\vartheta, \mathbf{H}] = 0,$$

and also has the symmetry group $\mathcal{G}$ of operators $\mathcal{O}_{\mathcal{R}}$,

$$[\mathcal{O}_R, \mathbf{H}] = 0,$$

which are themselves time-reversal invariant, $[\mathcal{O}, \vartheta] = 0$. Let $\Psi_\mu^{(i)}$ be an eigenvector of $\mathbf{H}$ with the energy E that belongs to the μ^{th} row if the i^{th} irreducible representation of $\mathcal{G}$,

$$\mathcal{O}_R \Psi_\mu^{(i)} = \sum_\nu \Psi_\nu^{(i)} D_{\nu\mu}^{(i)}(R).$$

Then

$$\vartheta \mathcal{O}_R \Psi_\mu^{(i)} = \mathcal{O}_R \vartheta \Psi_\mu^{(i)} = \sum_\nu (\vartheta \Psi_\nu^{(i)}) D_{\nu\mu}^{(i)*}(R),$$

which says that $\vartheta \Psi_\mu^{(i)}$ belongs to $D^{(i)*}$. According to Section E.2.2 in the appendix, if a representation is of type 3, it is not equivalent to its complex conjugate. Therefore, if the representation $D^{(i)}$ is of type 3, the energy level under consideration must belong also to the inequivalent representation $D^{(i)*}$: thus **the degeneracy is doubled.**

On the other hand, suppose $\Psi^{(i)}$ belongs to an irreducible representation $D^{(i)}$ of type 1 or 2, meaning that there exists a unitary matrix M such that $D^{(i)*}(R) = M D^{(i)}(R) M^{-1}$, and according to the results of Section E.2.2, $\widetilde{M} = cM$, where $c = 1$ if it is of type 1 and $c = -1$ if it is of type 2. Therefore,

$$a_{\mu\nu}^{(i)} \stackrel{\text{def}}{=} (\vartheta \Psi_\mu^{(i)}, \Psi_\nu^{(i)}) = (\vartheta \Psi_\nu^{(i)}, \vartheta^2 \Psi_\mu^{(i)}) = \alpha (\vartheta \Psi_\nu^{(i)}, \Psi_\mu^{(i)}) = \alpha a_{\nu\mu}^{(i)},$$

where α, defined by (5.90), equals ± 1, depending upon whether the number of spin-1/2 degrees of freedom is even or odd. But we know that $\vartheta \Psi_\mu^{(i)}$ belongs to the representation $D^{(i)*}$; so if the level has the normal degeneracy equal to the dimension of $D^{(i)}$, there must exist a unitary matrix M such that $\vartheta \Psi_\mu^{(i)} \stackrel{\text{def}}{=} \bar{\Psi}_\mu^{(i)} = \sum_\nu \Psi_\nu^{(i)} M_{\nu\mu}$. It follows that

$$a_{\mu\nu}^{(i)} = (\bar{\Psi}_\mu^{(i)}, \Psi_\nu^{(i)}) = \sum_\lambda M_{\lambda\mu}^* (\Psi_\lambda^{(i)}, \Psi_\nu^{(i)}) = M_{\nu\mu}^*,$$

from which we conclude that $\alpha = c$.

Thus we have the following **theorem:** *If the representation to which the level belongs is of type 1 and the number n of spin-1/2 degrees of freedom is even, or if the representation is of type 2 (which is possible only if its dimension is even) and n is odd, time-reversal invariance does not increase the degeneracy; in all other cases, it doubles the degeneracy.*

A special case of this result is one in which a given representation is one-dimensional; i.e., on grounds of symmetry under $\mathcal{G}$ the level is non-degenerate. In that case $D \approx D^*$ means, of course, that $D = D^*$ and the representation is real; moreover, it has to be of type 1. Consequently we can infer what is known as **Kramers's theorem:** *If an atomic system is invariant under time reversal (implying that there is no external magnetic field present), then an odd number of electrons makes every energy level at least twofold degenerate.*

5.4 Problems and Exercises

1. Derive the commutation relations of the components of the angular momentum operator $\vec{\mathbf{J}}$ with the components of a dynamical variable of the physical system that transform like a tensor of rank 2.

2. Denote by $|j, m_x\rangle$ a normalized eigenfunction of the $\vec{\mathbf{J}}^2$ and $\mathbf{J}_x$ with the eigenvalues $j(j+1)$ and m_x, respectively. Use (5.13) and (5.14) to calculate $\langle j, m|j, m_x\rangle$.

3. For $j = 1$, calculate the coefficients a, b, c in $|1, m_x\rangle = a|1, m = -1\rangle + b|1, m = 0\rangle + c|1, m = 1\rangle$, where m is an eigenvalue of j_z and m_x is an eigenvalue of j_x.

4. Suppose that $\mathbf{A}$ and $\mathbf{B}$ are either two numerical vectors, or else vectors whose components are operators that commute with $\vec{\sigma}$. Prove that

$$(\mathbf{A} \cdot \vec{\sigma})(\mathbf{B} \cdot \vec{\sigma}) = \mathbf{A} \cdot \mathbf{B} + i\vec{\sigma} \cdot \mathbf{A} \times \mathbf{B}.$$

5. Prove that if $\vec{\mathbf{A}}$ commutes with $\vec{\sigma}$, then

$$\vec{\sigma}\,\vec{\mathbf{A}} \cdot \vec{\sigma} = \vec{\mathbf{A}}\mathbf{1} + i\vec{\mathbf{A}} \times \vec{\sigma}$$

and

$$[\vec{\sigma}, \vec{\mathbf{A}} \cdot \vec{\sigma}] = 2i\vec{\mathbf{A}} \times \vec{\sigma}.$$

6. Show that the eigenstates of $\vec{\mathcal{J}}^2 \stackrel{\text{def}}{=} (\vec{\mathcal{L}} + \frac{1}{2}\vec{\sigma})^2$ with the eigenvalue $(l + \frac{1}{2})(l + \frac{3}{2})$ are eigenstates of $\vec{\mathcal{L}} \cdot \vec{\sigma}$ with the eigenvalue l, and the eigenstates of $\vec{\mathcal{J}}^2$ with the eigenvalue $(l - \frac{1}{2})(l + \frac{1}{2})$ are eigenstates of $\vec{\mathcal{L}} \cdot \vec{\sigma}$ with the eigenvalue $-l - 1$.

7. Show that the two operators defined by (5.71) are projections (i.e., that they are idempotent).

8. Suppose a physical system is in the angular-momentum state $\chi = a\chi_1^1 + b\chi_0^1 + c\chi_{-1}^1$, where $|a|^1 + |b|^2 + |c|^2 = 1$ and χ_1^1, χ_0^1, χ_{-1}^1 are the spin-1 states (5.21). Calculate the expectation value of $\vec{\mathcal{J}}$.

9. Show that the bilinear product of two spinors $\widetilde{\chi}^{1/2} i\sigma_y \phi^{1/2} = \chi_1^{1/2}\phi_2^{1/2} - \chi_2^{1/2}\phi_1^{1/2}$ is invariant under rotations.

10. Suppose that two successive rotations are performed, first by the angle θ about the axis $\widehat{n}$ and then by an angle ϕ about $\widehat{m}$. Show in terms of the spin-1/2 matrices that the rotation matrix for this is the same as first rotating by ϕ and $\widehat{m}$ and then rotating by θ about a new axis $\widehat{n}'$ that is obtained from $\widehat{n}$ by the second rotation.

11. Let $\vec{\mathbf{A}}$ and $\vec{\mathbf{B}}$ be two vector operators. Show directly from their commutation relations with the angular momentum that $\vec{\mathbf{J}}$ commutes with $\vec{\mathbf{A}} \cdot \vec{\mathbf{B}}$.

12. Calculate the spinor matrix T that corresponds to a rotation in terms the *Euler angles.*

13. Let Ψ_{m_x} be the normalized eigenvector of $\vec{\mathcal{J}}^2$ with the eigenvalue 2 (i.e., j=1) and of j_x with the eigenvalue m_x. Then we must be able to write

$$\Psi_{m_x} = a\Phi_{-1} + b\Phi_0 + c\Phi_1,$$

where Φ_m are the simultaneous normalized eigenvectors of j_z with the eigenvalue m and of $\vec{\mathcal{J}}^2$ with the eigenvalue 2. Find a, b, and c.

14. Suppose that $\widehat{n}$ and $\widehat{m}$ are two linearly independent vectors and that the operator $\mathbf{A}$ commutes with $\widehat{n} \cdot \vec{\mathbf{J}}$ as well as $\widehat{m} \cdot \vec{\mathbf{J}}$. Show that this implies that $\mathbf{A}$ commutes with all components of $\vec{\mathbf{J}}$.

15. Let $\vec{\mathbf{s}}$ be the spin angular momentum vector, in units of $\hbar$, of a particle of spin-1 (i.e., $\vec{\mathbf{s}}^2 = s(s+1) = 2$) and define $s_p \stackrel{\text{def}}{=} \widehat{p} \cdot \vec{\mathbf{s}}$, where $\widehat{p}$ is an aritrary unit vector. (a) Show that $s_p^3 = s_p$, and

$$\exp(-i\varphi s_p) = \mathbb{1} - is_p \sin\varphi + s_p^2(\cos\varphi - 1).$$

Write this out explicitly in a representation in which s_p is diagonal. (b) Take $\widehat{p}$ in the z direction and construct the matrices for s_x and s_y in a representation in which s_z is diagonal. (Use the commutation relations with s_z.)

16. Consider a spin-1/2 particle. Show that in the space of states of a given orbital angular momentum l, the operators

$$\frac{(l+1)\mathbb{1} + \vec{\mathcal{L}} \cdot \vec{\sigma}}{2l+1} \quad \text{and} \quad \frac{l\mathbb{1} - \vec{\mathcal{L}} \cdot \vec{\sigma}}{2l+1}$$

are projections onto the states of total angular momentum $j = l + \frac{1}{2}$ and $j = l - \frac{1}{2}$, respectively.

17. Let $\vec{\mathbf{S}}$ be the total spin (in units of $\hbar$) of a system of two nucleons. (a) Show that the operator $\mathbf{Q} \stackrel{\text{def}}{=} (\vec{\mathbf{S}} \cdot \vec{\mathbf{r}})^2/\mathbf{r}^2$ is a projection. (b) Show that the tensor operator $\mathbf{S}_{12} \stackrel{\text{def}}{=} 2(3\mathbf{Q} - \vec{\mathbf{S}}^2)$ satisfies the identity

$$\mathbf{S}_{12}^2 = 4\vec{\mathbf{S}}^2 - 2\mathbf{S}_{12}$$

and that its only possible eigenvalues are 0, 2, and -4.

18. Show that (a) if a 2×2 matrix commutes with all three of the Pauli matrices then it must be a multiple of the unit matrix, and (b) if a 2×2 matrix anti-commutes with all three of the Pauli matrices then it must be the zero matrix.

19. Assume that the Hamiltonian of a particle of spin-1/2 has the form

$$\mathbf{H} = \frac{1}{2}(\mathbf{Q}_0 \mathbb{1} + \vec{Q} \cdot \vec{\sigma}),$$

where $\mathbf{Q}_0$ is an operator that acts on the non-spin degrees of freedom and $\vec{Q}$ is a numerical vector. Let $\vec{P}$ be the polarization vector. Prove that

$$\hbar \frac{d\vec{P}}{dt} = \vec{Q} \times \vec{P}, \qquad \frac{d(\vec{P})^2}{dt} = 0,$$

and that if $\vec{Q}$ is constant, then $\vec{Q} \cdot \vec{P}$ and $(d\vec{P}/dt)^2$ are constants of the motion.

20. Show that if two pure states of a spin-1/2 particle are orthogonal, then the polarization vectors of these two states are equal and opposite.

21. Consider a spinless particle subject to a spherically symmetric Hamiltonian. Suppose it is in an unpolarized mixed state of angular momentum l (i.e., all z-projections of its angular momentum are equally probable). Write down its angular momentum density matrix in the configuration representation.

    Let the configuration representation matrix of the operator $\mathbf{A}$ be

$$\langle \vec{r} | \mathbf{A} | \vec{r}' \rangle = \sum_l A_l(r, r') P_l(\hat{r} \cdot \hat{r}').$$

    Calculate its expectation value (at fixed r and r') in the aforementioed mixed state.

22. Suppose a pair of spin-1/2 particles is produced in a pure triplet state. Calculate the spin density matrix for one of the particles alone.

23. Suppose that $\mathbf{G}_x$ is the x-component of a vector operator. What are the selection rules for the operator $\mathbf{G}_x^2$; i.e., for which differences between j and j' and between m and m' must the matrix elements $\langle j', m' | \mathbf{G}_x^2 | j, m \rangle$ vanish?

24. A transmitter simultaneously sends two correlated (but not necessarily identical) signals to a pair of independent receivers, A and B, that are far distant from each other. Each of the receivers has a switch that is set at either 1 or 2, independently of the other, before the arrival of the signal. Upon receiving the signal, each receiver flashes

either a red or a green light. Each run is recorded in the form, say, "2G1R" indicating "receiver A, set at 2, flashes green; receiver B, set at 1, flashes red." A long sequence of runs is examined and the following is found: (a) neither 1G2G nor 2G1G ever occurs; (b) 1R1R never happens; (c) 2G2G happens one-third of the time when both receivers are set at 2.

Suppose we denote a pair of signals from the transmitter by, say, $[GR, RR]$, meaning "if A is set at 1, flash green, if set at 2, flash red; if B is set at 1, flash red, if set at 2, flash red."

Now notice that (a) implies no messages of the form $[G\cdot, \cdot G]$ or of the form $[\cdot G, G\cdot]$ were sent, and (b) implies no messages of the form $[R\cdot, R\cdot]$ were sent either. But 2G2G can be brought about only by one of the four signals of the form $[\cdot G, \cdot G]$, each of which would lead to a contradiction with (a) or (b). (Show this.) Therefore, (c) contradicts the combination of (a) and (b).

Show how the observed results can be obtained if the two signals are made up of beams of atoms of spin-1/2, each of the pair originating from the decay of a spin-1 molecule in a specific, fixed pure state, and both receivers are equipped with Stern-Gerlach devices that are oriented vertically if set at 1 and horizontally if set at 2. If an SG apparatus finds spin *up* or *right* it flashes red, and if it finds *down* or *left* it flashes green.

6
Stationary Approximation Methods

Since very few quantum problems are amenable to exact solution, it is important to have a variety of approximation methods at our disposal. This chapter will describe several such procedures for stationary states, applicable under various physical circumstances and for different purposes.

6.1 The WKB Approximation

In this section we are going to study the behavior of quantum-mechanical one-particle systems when the wavelength λ of the wave function is short compared to the scale on which the potential changes by a significant fraction, which is analogous to the emergence of geometrical optics from physical optics when the wavelength of light is small compared to the scale on which the index of refraction changes. Since $\lambda = 2\pi\hbar/p$, this means we are talking about high energy. At the same time, the approximation may also be regarded as the regime where $\hbar$ is small, so that we are close to classical physics. In principle, after all, classical physics ought to emerge from quantum physics when the size of Planck's constant can be regarded as negligible. However, as we shall see, the manner in which this happens is not as simple and straightforward as you might expect. The regime that stops just short of going to the classical limit, and which is therefore referred to

as the *semi-classical approximation,* is also called the *WKB approximation* (after G. Wenzel, H. A. Kramers, and L. Brillouin).[1]

Suppose we write a one-particle wave function in the form

$$\psi(\vec{r},t) = Ae^{iS(\vec{r},t)/\hbar}, \tag{6.1}$$

denoting the particle's position in its configuration space conventionally by $\vec{r}$, and where A is a constant. Insertion in the Schrödinger equation then leads to the following equation for the complex function S:

$$-\frac{i\hbar}{2M}\nabla^2 S + \frac{1}{2M}(\nabla S)^2 + V + \frac{\partial S}{\partial t} = 0. \tag{6.2}$$

You will notice that when $\hbar = 0$, this equation goes over into the Hamilton-Jacobi equation of classical mechanics, whose solution S is called *Hamilton's principal function* and in terms of which $\vec{p} = \nabla S$. In the classical limit, therefore, the particle's physically possible trajectories are orthogonal to the surfaces of constant S, that is, to the surfaces of constant phase of the wave function.

If ψ is an eigenfunction of the energy, then

$$\psi(\vec{r},t) = \psi(\vec{r})e^{-iEt/\hbar},$$

so that if ψ is written in the form

$$\psi(\vec{r}) = Ae^{iW(\vec{r})/\hbar},$$

then

$$S(\vec{r},t) = W(\vec{r}) - Et.$$

The resulting equation for W is

$$-\frac{i\hbar}{2M}\nabla^2 W + \frac{1}{2M}(\nabla W)^2 + V = E, \tag{6.3}$$

which shows that as $\hbar \to 0$, W approaches what in classical mechanics is called *Hamilton's characteristic function.*

Let us now take the case in which the configuration space is one-dimensional (the WKB approximation is really applicable only when a problem can be reduced to a one-dimensional one), so that the equation for W becomes

$$-i\hbar W'' + W'^2 + 2M(V - E) = 0. \tag{6.4}$$

This may be either an instance of one-dimensional motion or that of a particle in three dimensions with a central potential after separating out

[1]It is also called the $JWKB$ approximation, the letter J standing for H. Jeffreys.

the angular variables. [In the latter case, V in (6.4) contains the centrifugal term $\hbar^2 l(l+1)/2Mr^2$.] Setting $W = \hbar u$ changes (6.4) to

$$u'^2 - iu'' = k^2, \tag{6.5}$$

where

$$k(x) \stackrel{\text{def}}{=} \frac{2\pi}{\lambda(x)} \stackrel{\text{def}}{=} \sqrt{\frac{2M}{\hbar^2}[E - V(x)]} \tag{6.6}$$

is the *local wave number* and $\lambda(x)$ is the *local wavelength.*

Next, set

$$u(x) \stackrel{\text{def}}{=} \int^x dx'\, g(x') + \frac{1}{2}\log g(x),$$

which implies that

$$\psi(x) = g^{-1/2}\exp[i\int^x dx'\, g(x')], \tag{6.7}$$

and the equation for g becomes

$$g^2 - \frac{3}{4}\left(\frac{g'}{g}\right)^2 + \frac{1}{2}\frac{g''}{g} = k^2(x). \tag{6.8}$$

The somewhat complicated non-linear equation (6.8) is now solved by successive approximations on the assumption that the function $k(x)$ *is slowly varying,* i.e., that

$$\left|\frac{d\lambda}{dx}\right| \ll 2\pi,$$

so that it is meaningful to refer to a "local wavelength." This assumption implies that

$$\left|\frac{d(1/k)}{dx}\right| = \left|\frac{1}{k^2}\frac{dk}{dx}\right| = \left|\frac{1}{2k^3}\frac{d(k^2)}{dx}\right| \ll 2\pi,$$

or,

$$\lambda\left|\frac{d(E-V)/dx}{E-V}\right| \ll 1.$$

In other words, the relative change of the kinetic energy over one wavelength should be small.

When $k = $ const., the solutions of (6.8) obviously are $g = \pm k$. In the first approximation, therefore, we use these values also when k is slowly varying.

A better approximation would be given by

$$g^2 = k^2 + \frac{3}{4}\left(\frac{k'}{k}\right)^2 - \frac{1}{2}\frac{k''}{k},$$

but that would differ from $g^2 = k^2$ by terms of order $(k'/k)^2$ and k''/k. If $k'/k \ll 1$ inside the whole region of interest, as assumed, then k''/k has to be small there as well for $g = \pm k$ to be a good approximation.

The *WKB* approximation is thus given by

$$\psi_{\rm WKB}(x) = A \exp[-\frac{1}{2} \log k(x) \pm i \int^x dx'\, k(x')]$$

or

$$\psi_{\rm WKB}(x) = \frac{A}{\sqrt{k(x)}}\, e^{\pm i \int^x dx'\, k(x')}, \tag{6.9}$$

which in the classically forbidden region, where $E < V$, becomes

$$\psi_{\rm WKB}(x) = \frac{A'}{\sqrt{\kappa(x)}}\, e^{\pm \int^x dx'\, \kappa(x')}, \qquad \kappa(x) \stackrel{\rm def}{=} \sqrt{\frac{2M}{\hbar^2}[V(x) - E]}. \tag{6.10}$$

Suppose, then, that we wish to solve a bound-state problem for a potential well, or a reflection problem on a barrier. In that case it would be necessary to connect the wave function with a continuous logarithmic derivative across a boundary between a classically allowed and a classically forbidden region; but in the vicinity of this boundary, the assumption $k'/k \ll 1$ breaks down, because $k(x)$ vanishes at the boundary. Classically, a point where $k = 0$ is generally a *turning point* of the motion. The principal problem of the *WKB* approximation is therefore to find the *connecting formulas* for the wave function across such a turning point, the vicinity of which must be treated separately.

6.1.1 The connecting formulas

For simplicity, choose x in such a way that the turning point is at the origin, $x = 0$, so that

$$k^2(x) = cx + \dots .$$

Thus $c > 0$ means that the classically allowed region is to the right (Fig. 6.1), so that *the turning point is on the left,* and the Schrödinger equation near the turning point reads, in appropriate units,

$$\psi'' + cx\psi = 0.$$

In terms of $\xi_1 \stackrel{\rm def}{=} \int_0^x dx'\, k(x') = \frac{2}{3}\sqrt{cx^3} + \dots$ for $x > 0$, two linearly independent solutions in the classically allowed region near the turning point are

$$\psi = \text{const.}\sqrt{\frac{\xi_1}{k}} J_{\pm 1/3}(\xi_1) = \text{const.}\sqrt{x} J_{\pm 1/3}(\frac{2}{3}\sqrt{cx^3}),$$

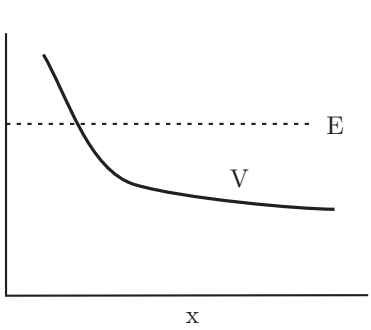

FIGURE 6.1. A turning point to the left of a classically allowed region.

and two solutions in the forbidden region are, in terms of $\xi_2 \stackrel{\text{def}}{=} \int_x^0 dx'\,\kappa(x')$,

$$\psi = \text{const.}\sqrt{\frac{\xi_2}{\kappa}} I_{\pm 1/3}(\xi_2) = \text{const.}\sqrt{|x|} I_{\pm 1/3}(\frac{2}{3}\sqrt{c|x|^3}),$$

where J and I are conventional Bessel functions of real and imaginary arguments, respectively.[2] These functions have the following behavior near the origin:

$$J_{\pm 1/3}(\xi) \stackrel{\xi\to 0}{\longrightarrow} (\frac{1}{2}\xi)^{\pm 1/3}/\Gamma(1 \pm \frac{1}{3}),$$
$$I_{\pm 1/3}(\xi) \stackrel{\xi\to 0}{\longrightarrow} (\frac{1}{2}\xi)^{\pm 1/3}/\Gamma(1 \pm \frac{1}{3});$$

therefore, the solution which for $x > 0$ is $\sqrt{x}J_{1/3}$, becomes

$$\frac{3^{-1/3}c^{1/6}}{\Gamma(4/3)}\, x$$

near $x = 0$, whereas the solution which for $x < 0$ is $\sqrt{|x|}I_{1/3}$ becomes

$$\frac{3^{-1/3}c^{1/6}}{\Gamma(4/3)}\, |x| = -\frac{3^{-1/3}c^{1/6}}{\Gamma(4/3)}\, x.$$

Consequently, $\sqrt{x}J_{1/3}$ fits smoothly with $-\sqrt{|x|}I_{1/3}$, and similarly we find that $\sqrt{x}J_{-1/3}$ fits smoothly with $+\sqrt{|x|}I_{-1/3}$.

At a distance that is many wavelengths away from the turning point, but not so far that the potential differs drastically from the straight line assumed near the turning point, the asymptotic forms of the Bessel functions are[3]

$$J_{\pm 1/3}(\xi) \simeq \sqrt{\frac{2}{\pi\xi}} \cos(\xi \mp \frac{\pi}{6} - \frac{\pi}{4}),$$

---

[2]The asymptotically decreasing solution is known as the *Airy integral*, $Ai(z) \stackrel{\text{def}}{=} \frac{1}{\pi}\int_0^\infty ds\, \cos[(s^3/3) + zs]$, and in terms of this function, $\psi(x) = Ai(-c^{1/3}x)$.

[3]See [Erdélyi], vol. 2, pp. 85/6.

$$I_{1/3}(\xi) \simeq \frac{1}{\sqrt{2\pi\xi}}\, e^{\xi},$$

and

$$I_{+1/3}(\xi) - I_{-1/3}(\xi) \simeq -\sqrt{\frac{2}{\pi\xi}} \sin\frac{\pi}{3}\, e^{-\xi}.$$

Therefore, the solution which far to the left of the turning point is of the form

$$-\sqrt{\frac{2}{\pi\kappa}} \sin\frac{\pi}{3}\, e^{-\int_x^0 dx'\,\kappa(x')},$$

must, many wavelengths to its right, be of the form

$$\begin{aligned} & -\sqrt{\frac{2}{\pi k}}\left[\cos(\xi_1 - \frac{\pi}{4} + \frac{\pi}{6}) + \cos(\xi_1 - \frac{\pi}{4} - \frac{\pi}{6})\right] \\ = \; & -2\sqrt{\frac{2}{\pi k}} \cos\frac{\pi}{6} \cos(\xi_1 - \frac{\pi}{4}) \\ = \; & -2\sqrt{\frac{2}{\pi k}} \sin\frac{\pi}{3} \cos\left(\int_0^x dx'\, k(x') - \frac{\pi}{4}\right). \end{aligned}$$

Consequently the following two *WKB* solutions, the first in the classically forbidden region on the left, and the second in the classically allowed region on the right, are smoothly connected at the turning point:

$$\frac{1}{\sqrt{\kappa(x)}} \exp\left[-\int_x^0 dx'\,\kappa(x')\right] \xrightarrow{\text{ltp}} \frac{2}{\sqrt{k(x)}} \cos\left[\int_0^x dx'\, k(x') - \frac{\pi}{4}\right], \quad (6.11)$$

where the arrow indicates that the formula is stable when used in the direction indicated, and only then. This is because, when used in the opposite direction, a small perturbation of the function on the right would produce an exponentially increasing and thus asymptotically dominant term on the left.

Similarly, we find for arbitrary η that there is a smooth connection between the functions on the right and the left of the turning point,

$$\frac{\sin\eta}{\sqrt{\kappa(x)}} \exp\left[\int_x^0 dx'\,\kappa(x')\right] \xleftarrow{\text{ltp}} \frac{1}{\sqrt{k(x)}} \cos\left[\int_0^x dx'\, k(x') - \frac{\pi}{4} + \eta\right], \quad (6.12)$$

which implies that

$$\frac{1}{\sqrt{\kappa(x)}} \exp\left[\int_x^0 dx'\,\kappa(x') \mp i\frac{\pi}{4}\right] \xleftarrow{\text{ltp}} \frac{1}{\sqrt{k(x)}} \exp\left[\pm i\int_0^x dx'\, k(x')\right]. \quad (6.13)$$

The formula (6.11) shows that an exponentially decreasing, and thus asymptotically small, addition on the left would produce a non-negligible change of the function on the right; therefore, (6.12) and (6.13) should be used only

in the direction of the arrows indicated. These are the connecting formulas for a turning point on the left.

On the other hand, if the classical turning point is on the right, we find by similar arguments[4] that

$$\frac{2}{\sqrt{k(x)}}\cos\left[\int_x^0 dx'\,k(x') - \frac{\pi}{4}\right] \overset{\text{rtp}}{\longleftarrow} \frac{1}{\sqrt{\kappa(x)}}\exp\left[-\int_0^x dx'\,\kappa(x')\right], \quad (6.14)$$

$$\frac{1}{\sqrt{k(x)}}\cos\left[\int_x^0 dx'\,k(x') - \frac{\pi}{4} + \eta\right] \overset{\text{rtp}}{\longrightarrow} \frac{\sin\eta}{\sqrt{\kappa(x)}}\exp\left[\int_0^x dx'\,\kappa(x')\right], \quad (6.15)$$

or

$$\frac{1}{\sqrt{k(x)}}\exp\left[\pm i\int_x^0 dx'\,k(x')\right] \overset{\text{rtp}}{\longrightarrow} \frac{1}{\sqrt{\kappa(x)}}\exp\left[\int_0^x dx'\,\kappa(x') \mp i\frac{\pi}{4}\right]. \quad (6.16)$$

When the energy is near a point at which the slope of the potential vanishes, or where two classical turning points are close to one another, so that the assumption of a straight-line (non-horizontal) potential behavior cannot be maintained over a distance of many wavelengths, these formulas have to be modified. (For further details, see [Fröman].)

6.1.2 Examples

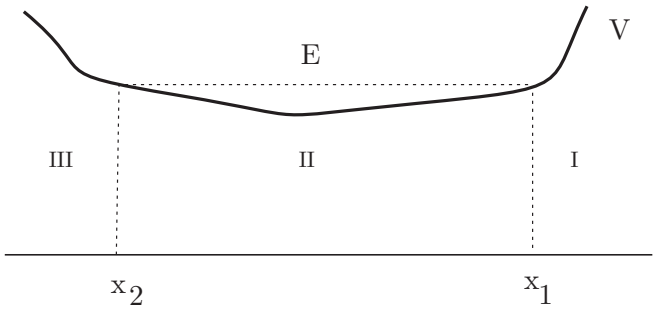

FIGURE 6.2. A potential well.

As a first example, we use the *WKB* approximation to calculate the energy levels in a **potential well** of the form shown in Figure 6.2. In region I the wave function must have the form

$$\psi = \frac{1}{\sqrt{\kappa}}\exp\left[-\int_{x_1}^x dx'\,\kappa(x')\right],$$

[4] **Derive these.**

while in region III it must be of the form

$$\psi = \frac{A}{\sqrt{\kappa}} \exp\left[-\int_x^{x_2} dx'\, \kappa(x')\right].$$

For region II we obtain from the connecting formula (6.14) for the right turning point

$$\psi = \frac{2}{\sqrt{k}} \cos\left[\int_x^{x_1} dx'\, k(x') - \frac{\pi}{4}\right],$$

and for the left turning point from (6.11),

$$\begin{aligned}
\psi &= \frac{2A}{\sqrt{k}} \cos\left[\int_{x_2}^{x} dx'\, k(x') - \frac{\pi}{4}\right] \\
&= \frac{2A}{\sqrt{k}} \cos\left\{-\int_x^{x_1} dx'\, k(x') + \frac{\pi}{4} + [\int_{x_2}^{x_1} dx'\, k(x') - \frac{\pi}{2}]\right\} \\
&= \frac{2A}{\sqrt{k}} \cos\left[\int_x^{x_1} dx'\, k(x') - \frac{\pi}{4} - \eta\right],
\end{aligned}$$

where

$$\eta = \int_{x_2}^{x_1} dx'\, k(x') - \frac{\pi}{2}.$$

In order for this ψ to be equal to the result obtained from the right-hand turning point, it is necesssary and sufficient that

$$\eta = n\pi,$$

where n is an integer. Therefore, the requirement is that $\int_{x_2}^{x_1} dx\, k(x) = (n + \frac{1}{2})\pi$, or

$$2\int_{x_2}^{x_1} dx\, \sqrt{2M[E_n - V(x)]} = (n + \frac{1}{2})h. \tag{6.17}$$

The left-hand side of this equation being the phase integral of the motion, you will recognize the formula as the old Sommerfeld quantization rule, including the somewhat more sophisticated 1/2-term. However, since the connecting formulas hold only at distances that are many wavelengths away from the classical turning point, the result is a good approximation only when $n \gg 1$, in agreement with the fact that the Sommerfeld rule originates in the correspondence-principle limit. Notice that (6.17) implies that if $V \to \infty$ as $x \to \pm\infty$, then there must be infinitely many energy levels.

As a second example, consider **tunneling through a barrier**. The potential is now assumed to have a shape such as that shown in Figure 6.3, and we take the incident wave to come from the left, so that in region I there is no wave traveling to the left. Therefore, the *WKB* solution in region I must be of the form

$$\psi = \frac{A}{\sqrt{k}} \exp\left[i\int_{x_1}^{x} dx'\, k(x')\right] \sim \frac{A}{\sqrt{k}} \exp\left[i\int_{x_1}^{\infty} dx'\, [k(x') - k]\right] e^{ik(x-x_1)}.$$

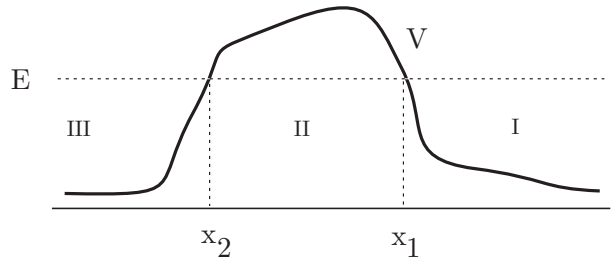

FIGURE 6.3. A potential barrier.

The connecting formula (6.13) tells us that in region II the *WKB* solution must thus be

$$\begin{aligned}\psi &= \frac{A}{\sqrt{\kappa}}\exp\left[\int_x^{x_1} dx'\,\kappa(x') - i\frac{\pi}{4}\right] \\ &= \frac{A}{\sqrt{\kappa}}\exp\left[\int_{x_2}^{x_1} dx'\,\kappa(x') - i\frac{\pi}{4}\right]\exp\left[-\int_{x_2}^{x} dx'\,\kappa(x')\right].\end{aligned}$$

In region III, the smoothly connected *WKB* wave function is, therefore, by (6.14),

$$\begin{aligned}\psi &= \frac{2A}{\sqrt{k}}\exp\left[\int_{x_2}^{x_1} dx'\,\kappa(x')\right]e^{-i\pi/4}\cos\left[\int_x^{x_2} dx'\,k(x') - \frac{\pi}{4}\right] \\ &= \frac{A}{\sqrt{k}}\exp\left[\int_{x_2}^{x_1} dx'\,\kappa(x')\right]\left\{\exp[-i\int_x^{x_2} dx'\,k(x')]\right. \\ &\qquad\left. - i\exp[i\int_x^{x_2} dx'\,k(x')]\right\} \\ &\sim \frac{A}{\sqrt{k}}\exp\left[\int_{x_2}^{x_1} dx'\,\kappa(x')\right]\left\{\exp[-i\int_{-\infty}^{x_2} dx'\,[k(x')-k] - ik(x_2-x)]\right. \\ &\qquad\left. -i\exp[i\int_{-\infty}^{x_2} dx'\,[k(x')-k] + ik(x_2-x)]\right\}.\end{aligned}$$

For large $|x|$, the first term in the braces of the second line goes as e^{ikx} and thus represents the incident wave traveling to the right, while the second term goes as e^{-ikx} and represents a reflected wave traveling to the left; thus the ratio of the transmitted to the incident amplitude is

$$T = e^{-\int_{x_2}^{x_1} dx\,\kappa(x)}e^{ik(x_2-x_1)}e^{i(\int_{-\infty}^{x_2}+\int_{x_1}^{\infty})dx\,[k(x)-k]} = e^{i\int_{-\infty}^{\infty} dx\,[k(x)-k]}.$$

The transmission coefficient, i.e., the ratio of the transmitted to the incident flux, is therefore, in the *WKB* approximation,

$$|T|^2 = \exp\left[-\frac{2}{\hbar}\int_{x_2}^{x_1} dx\,\sqrt{2M[V(x)-E]}\right], \tag{6.18}$$

where the points x_1 and x_2 are the solutions of the equation $V(x) = E$. The ratio of the reflected to the incident amplitude, on the other hand, is

$$R = -i \exp\left[2i \int_{-\infty}^{x_2} dx\, [k(x) - k] + 2ikx_2\right],$$

whose absolute magnitude equals 1, which expresses the fact that in the *WKB* approximation necessarily $|T| \ll 1$. (As the energy approaches the top of the barrier, this is no longer the case, and the *WKB* approximation has to be modified, as indicated earlier.) By conservation of flux, the reflection coefficient has to be

$$|R|^2 = 1 - |T|^2,$$

but since the approximation includes only the leading term, it simply yields $|R|^2 = 1$.

6.1.3 Scattering in the WKB approximation

Let us next look at the manner in which quantum scattering approaches classical scattering in the limit as $\hbar \to 0$. It will turn out that this approach is rather complicated.

Assuming that the potential is spherically symmetric, the first step, of course, is to perform a partial-wave analysis and consider a fixed angular-momentum quantum number l. The relevant potential then includes the

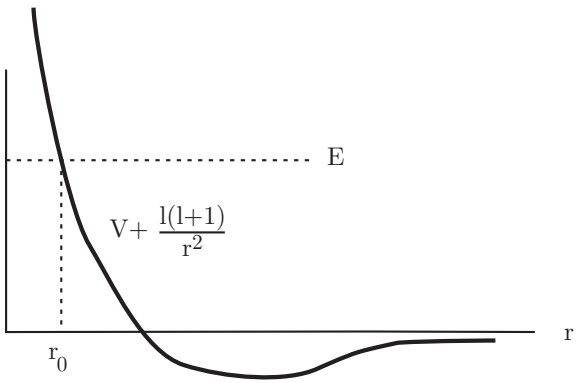

FIGURE 6.4. The effective potential at a given $l > 0$.

centrifugal term,

$$V_{\text{eff}}(r) \stackrel{\text{def}}{=} V(r) + \frac{\hbar^2 l(l+1)}{2Mr^2},$$

as shown in Figure 6.4. Therefore, define

$$\kappa_l(r) \stackrel{\text{def}}{=} \sqrt{\frac{2M}{\hbar^2}[V(r) - E] + \frac{l(l+1)}{r^2}}, \quad k_l(r) \stackrel{\text{def}}{=} \sqrt{\frac{2M}{\hbar^2}[E - V(r)] - \frac{l(l+1)}{r^2}}. \tag{6.19}$$

The point $r = r_0$ is defined by $k_l(r_0) = \kappa_l(r_0) = 0$. In the classically forbidden region near the origin, the radial wave function is supposed to be regular; therefore, in the *WKB* approximation it must there be of the form

$$\psi_l = \frac{1}{\sqrt{\kappa_l}} \exp\left[-\int_r^{r_0} dr'\, \kappa_l(r')\right].$$

The connecting formula (6.11) tells us that for $r > r_0$ it must thus be

$$\begin{aligned}
\psi_l &= \frac{2}{\sqrt{k_l(r)}} \cos\left[\int_{r_0}^{r} dr'\, k_l(r') - \frac{\pi}{4}\right] \\
&= \frac{2}{\sqrt{k_l(r)}} \sin\left\{\int_{r_0}^{r} dr'\, [k_l(r') - k] + \frac{\pi}{4} + (r - r_0)k\right\} \\
&\sim_{r\to\infty} \frac{2}{\sqrt{k}} \sin\left\{kr + \frac{\pi}{4} - kr_0 + \int_{r_0}^{\infty} dr'\, [k_l(r') - k]\right\},
\end{aligned}$$

since $k_l(r) \sim k$ as $r \to \infty$. As a result, we can conclude that the phase shift in the *WKB* approximation is given by

$$\delta_l^{\mathrm{WKB}} = \frac{\pi}{4} + \frac{\pi}{2} l - kr_0 + \int_{r_0}^{\infty} dr'\, [k_l(r') - k] = \frac{\pi}{2}(l + \frac{1}{2}) - \int_{r_0}^{\infty} dr'\, r' \frac{d}{dr'}\, k_l(r'). \tag{6.20}$$

Since the dominant term in $k_l(r)$ near the origin is the centrifugal term, the behavior of the *WKB* wave function $\psi_l(r)$ near $r = 0$ is

$$\psi \to \text{const.}\, r^{1/2} \exp[\sqrt{l(l+1)}\log r] = \text{const.}\, r^{\frac{1}{2} + \sqrt{l(l+1)}},$$

whereas we know that the behavior of the exact radial wave function is as r^{l+1}. Therefore, replacing $l(l+1)$ by $(l + \frac{1}{2})^2$ and using

$$k_l(r) \stackrel{\mathrm{def}}{=} \sqrt{\frac{2M}{\hbar^2}[E - V(r)] - \frac{(l + \frac{1}{2})^2}{r^2}}. \tag{6.21}$$

in (6.20), rather than (6.19), can be expected to improve the approximation.

The argument leading to the expression (6.20) for the phase shift is based on the assumption that for a particle of angular momentum l there exists a distance r_0 of closest approach, classically speaking, which is assured for $l > 0$, because of the centrifugal repulsion. For $l = 0$, however, such a point r_0 will generally not exist. Fortunately, this hiatus does not present any real obstacle, since in the semi-classical regime many phase shifts contribute to the scattering and the $l = 0$ term matters little, as we shall elaborate now.

The *WKB* approximation to the scattering amplitude does not consist simply of the use of the *WKB* phase shifts. The second important step is to rely on the fact that, since at high energies many l-values enter into the scattering amplitude, the large l-values will carry most of the burden of

the result. Or, looked at in another way: classically, each deflection angle is associated with a specific value of the angular momentum, which means, as $\hbar \to 0$, a fixed *large* l-value. The idea, therefore, is to do two things with the partial-wave expansion of the scattering amplitude: (a) to replace the Legendre polynomials in (4.50) by their asymptotic values for large l, and (b) to replace the series in (4.50) by an integral.

In step (a), the appropriate approximations are, for $\sin\theta > 1/l$,

$$P_l(\cos\theta) \simeq \sqrt{\frac{2}{\pi(l+\frac{1}{2})\sin\theta}}\,\cos[(l+\frac{1}{2})\theta - \frac{\pi}{4}], \tag{6.22}$$

and for $\sin\theta < 1/l$,

$$P_l(\cos\theta) \simeq (\cos\theta)^l\, J_0[(l+\frac{1}{2})\sin\theta], \tag{6.23}$$

(where J_0 is the Bessel function of order 0) which are fairly accurate even when l is not very large.

Step (b) is justified by the fact that many l-values make significant contributions to the partial-wave sum, and the *WKB* phase shift is a relatively slowly varying function of l. Restricting ourselves to scattering angles θ that are not too small, we then obtain from (a) and (b), using (6.22),

$$\begin{aligned} f(k,\theta) &= \frac{1}{2ik}\sum_l (2l+1)(e^{2i\delta_l}-1)P_l(\cos\theta) \\ &\approx \frac{1}{ik}\frac{1}{\sqrt{2\pi\sin\theta}}\int_0^\infty dl\,\sqrt{l+\frac{1}{2}}\left[e^{i[2\delta_l+(l+\frac{1}{2})\theta-\frac{\pi}{4}]}\right. \\ &\qquad\qquad \left. + e^{-i[2\delta_l+(l+\frac{1}{2})\theta+\frac{\pi}{4}]}\right] \\ &= \frac{1}{ik}\frac{1}{\sqrt{2\pi\sin\theta}}\int_0^\infty dl\,\sqrt{l+\frac{1}{2}}\left[e^{i[2\delta_l+(l+\frac{1}{2})\theta+2\pi nl-\frac{\pi}{4}]}\right. \\ &\qquad\qquad \left. + e^{-i[2\delta_l+(l+\frac{1}{2})\theta+2\pi nl+\frac{\pi}{4}]}\right], \end{aligned}$$

where the "1" has been dropped because it contributes only for $\theta = 0$, and the terms $2\pi nl$, with n an arbitrary integer, have been added in the exponentials without changing the result. Now writing $J \stackrel{\text{def}}{=} \hbar(l+\frac{1}{2})$, so that

$$\delta_l^{\text{WKB}} = \frac{1}{\hbar}\left[\frac{\pi}{2}J - \int_{r_0}^\infty dr\,\frac{d}{dr}\sqrt{2M(E-V)-J^2r^{-2}}\right], \tag{6.24}$$

the approximation to the scattering amplitude can be written in the form

$$f(k,\theta) \approx \frac{1}{\sqrt{\hbar}}\frac{1}{ip}\frac{1}{\sqrt{2\pi\sin\theta}}\int_0^\infty dJ\,J^{1/2}\left(e^{i\varphi_+}+e^{i\varphi_-}\right),$$

with

$$\varphi_\pm \stackrel{\text{def}}{=} [2\hbar\delta_l^{\text{WKB}} \pm J\theta \mp \frac{\pi}{4} + 2\pi n(J - \frac{1}{2}\hbar)]/\hbar.$$

The final step in the approximation procedure is to evaluate the above integrals by the method of stationary phase. For $\hbar \to 0$, as the terms $e^{i\varphi_\pm}$ become more and more oscillatory, thus leading to cancellations, the principal contribution to the integral comes from those J-values for which one of the phases is stationary, that is, locally independent of J, i.e., those for which

$$2\frac{\partial\delta_l^{\text{WKB}}}{\partial l} \pm \theta + 2\pi n = 0, \tag{6.25}$$

where the equation with the +sign leads to the leading contribution from φ_+, and the equation with the −sign, from φ_-.

Recall now that the classical deflection function Θ for a particle, which determines the asymptotic deviation θ of a particle's final trajectory from its incoming direction[5] by the formula

$$\theta = \mp[\Theta(J) + 2\pi n],$$

where n is an integer such that $0 < \theta < \pi$, is given in terms of the angular momentum J by

$$\Theta = \pi - 2\int_{r_0}^{\infty} dr\, r^{-2}[2M(E - V)J^{-2} - r^{-2}]^{-1/2}.$$

This deflection function can be expressed in terms of the *WKB* phase shift, as given by (6.24), in the form

$$\Theta(J) = 2\hbar\frac{\partial\delta_l^{\text{WKB}}}{\partial J} = 2\frac{\partial\delta_l^{\text{WKB}}}{\partial l}. \tag{6.26}$$

Thus (6.25) shows that in the limit as $\hbar \to 0$, *the quantum-mechanical scattering amplitude is concentrated precisely at the classical scattering angle.* If, for a given value of the angle θ, we denote the stationary value of l by l_0 and expand the phases $\varphi_\pm$ about l_0, we obtain

$$\varphi_\pm(l) = 2\delta_l^{\text{WKB}} \pm (l + \frac{1}{2})\theta \mp \frac{\pi}{4} + 2\pi nl = \varphi_0 + \frac{1}{2}(l - l_0)^2\varphi_0'' + \ldots$$

and hence

$$f(k,\theta) \approx \frac{1}{ik}\sqrt{\frac{l_0 + \frac{1}{2}}{2\pi\sin\theta}}e^{i\varphi_0}\int_{-\infty}^{\infty} dl\, e^{\frac{i}{2}(l-l_0)^2\varphi_0''} = \frac{1}{ik}\sqrt{\frac{l_0 + \frac{1}{2}}{-i\varphi_0''\sin\theta}}e^{i\varphi_0},$$

[5] See, for example, [Newton 82], Sec. 5.1.

where the integral has been extended to infinity with insignificant error because almost all of its value comes from contributions in the region of stationary phase anyhow. Insertion of

$$\varphi_0'' = 2\frac{\partial^2 \delta_l^{\mathrm{WKB}}}{\partial l^2} = \frac{\partial \Theta}{\partial (l+\frac{1}{2})}$$

therefore leads to the limit of the scattering cross section

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= |f|^2 = \frac{1}{k^2}\frac{l+\frac{1}{2}}{\sin\theta}\left|\frac{\partial(l+\frac{1}{2})}{\partial\theta}\right| \\ &= \frac{1}{2ME}\frac{J}{\sin\theta}\left|\frac{\partial J}{\partial\theta}\right| = \frac{b}{\sin\theta}\left|\frac{\partial b}{\partial\theta}\right|, \end{aligned}$$

where $b = J/p$ is the classical impact parameter. *This is precisely the classical result.*

The scattering amplitude, on the other hand, has, of course, no classical analogue, and in most circumstances its phase has no directly observable consequences. However, this is not so if there exists more than one value of l for which the phase of either φ_+ or φ_- is stationary. In that case, the quantum scattering cross section is the modulus squared of the sum of the various amplitudes obtained from the stationary points, and there will be interference effects that differ from the classical result, which would be the sum of the cross sections for the various impact parameters leading to the same scattering angle if Θ is not a monotone function of the impact parameter. These interference effects do not disappear as $\hbar \to 0$ but, instead, become highly oscillatory. (The angular widths of the interference fringes are of the order of the ratio between the de Broglie wavelength of the scattered particles and the difference between two impact parameters that lead to the same deflection angle.) As the phase of the scattering amplitude turns out to be proportional to $k = \sqrt{2ME}/\hbar$, this means at the same time that energy differences that macroscopically are considered small lead to large phase differences and observable interferences analogous to the diffraction effects that are the remnant deviations of physical optics from geometrical optics. Specific physical applications of these effects have been found, for example, in α-particle scattering by nuclei.[6]

There is an important physical lesson in this result: classical physics does not directly emerge as the limit $\hbar \to 0$ of quantum physics; that limit here simply does not exist. Classical results are valid for the description of the scattering of marbles and billiard balls because the interference effects would be noticeable only if the energies of the missiles could be controlled far more precisely than is macroscopically feasible; they are washed out by the practical experimental errors of observations of such objects.

[6]For details, see the papers [Ford 59a] and [Ford 59b].

6.2 Bound-State Perturbation Theory

The next topic will be perturbation theory as applicable to the bound-state solutions of the time-independent Schrödinger equation. The question to be answered is, if the Hamiltonian of a system is altered by a small amount, what are the corresponding small changes in the energy levels and in the energy eigenvectors?

Assume, then, that the Hamiltonian is of the form

$$\mathbf{H} = \mathbf{H}_0 + \lambda \mathbf{H}', \tag{6.27}$$

where the unperturbed Hamiltonian $\mathbf{H}_0$ is now not just the kinetic energy but includes some known interaction. For simplicity, assume that the spectrum of $\mathbf{H}_0$ consists of discrete points only, and the eigenvalues E_n^0, $n = 1, 2, \ldots$, as well as the (normalized) eigenfunctions $|E_n^0\rangle$ of $\mathbf{H}_0$ are known. Furthermore, assume that both the eigenvalues and the eigenvectors of $\mathbf{H}$ can be expanded in a power series in λ, of which we are interested only in the first few terms.[7]

As we saw in Section 4.4.1, the eigenvalues and eigenvectors of $\mathbf{H}$ can be recognized as the poles and residues of the *resolvent* operator,

$$(E - \mathbf{H})^{-1} = \sum_n \frac{|E_n\rangle \langle E_n|}{E - E_n}, \tag{6.28}$$

which is nothing but a re-expression of (4.125) in operator form. Consequently the equation

$$\langle E_m^0 | (E - \mathbf{H})^{-1} | E_{m'}^0 \rangle = \sum_n \frac{\langle E_m^0 | E_n \rangle \langle E_n | E_{m'}^0 \rangle}{E - E_n} \tag{6.29}$$

may be used to recognize the eigenvalues of $\mathbf{H}$ from the positions of the poles of the function of E on the left, and the transformation functions $\langle E_m^0 | E_n \rangle$ from the corresponding residues. Once the numbers $\langle E_m^0 | E_n \rangle$ are known, so are the new, perturbed, normalized eigenvectors by their expansion

$$|E_n\rangle = \sum_m \langle E_m^0 | E_n \rangle \, |E_m^0\rangle \tag{6.30}$$

on the basis of the unperturbed eigenvectors.

Let us isolate the terms up to λ^2 in the operator[8]

$$(E - \mathbf{H})^{-1} = (E - \mathbf{H}_0 - \lambda \mathbf{H}')^{-1} = (E - \mathbf{H}_0)^{-1} \left[\mathbb{1} - \lambda \mathbf{H}' (E - \mathbf{H}_0)^{-1} \right]^{-1};$$

[7] The question of whether the series converges or perhaps is a non-convergent asymptotic series will be left open.

[8] Remember that $(\mathsf{AB})^{-1} = \mathsf{B}^{-1}\mathsf{A}^{-1}$.

that is,

$$\begin{aligned}(E-\mathbf{H}_0-\lambda\mathbf{H}')^{-1} &= (E-\mathbf{H}_0)^{-1}+\lambda(E-\mathbf{H}_0)^{-1}\mathbf{H}'(E-\mathbf{H}_0)^{-1}\\ &\quad+\lambda^2(E-\mathbf{H}_0)^{-1}\mathbf{H}'(E-\mathbf{H}_0)^{-1}\mathbf{H}'(E-\mathbf{H}_0)^{-1}\\ &\quad+\lambda^3[(E-\mathbf{H}_0)^{-1}\mathbf{H}']^3(E-\mathbf{H}_0-\lambda\mathbf{H}')^{-1}, \quad (6.31)\end{aligned}$$

and from now on ignore the last term. The neglect of such higher-order terms in powers of λ relies on the assumption that the operator norm $\| \lambda(E-\mathbf{H}_0)^{-1}\mathbf{H}' \|$ is small compared to 1, and this requires not only that λ be sufficiently small for a given $\mathbf{H}'$, but also that E not be too close to one of the eigenvalues of $\mathbf{H}_0$. Perturbation theory has to be treated very cautiously if the spacing of the eigenvalues of $\mathbf{H}_0$ is very close and a shifted eigenvalue of $\mathbf{H}$ gets near the next eigenvalue of $\mathbf{H}_0$. (This case will be considered separately in Section 6.2.2.)

6.2.1 The nondegenerate case

Assuming for the moment—subject to later correction—that there is no degeneracy, writing

$$H'_{nm} \stackrel{\text{def}}{=} \langle E_n^0|\mathbf{H}'|E_m^0\rangle,$$

and keeping only terms up to order λ^2, we obtain from (6.31)

$$\begin{aligned}\langle E_n^0|(E-\mathbf{H})^{-1}|E_m^0\rangle &= \frac{\delta_{nm}}{E-E_n^0}+\lambda\frac{H'_{nm}}{(E-E_n^0)(E-E_m^0)}\\ &+ \lambda^2\sum_{k\neq n}\frac{H'_{nk}H'_{km}}{(E-E_n^0)(E-E_k^0)(E-E_m^0)}+\lambda^2\frac{H'_{nn}H'_{nm}}{(E-E_n^0)^2(E-E_m^0)}+\ldots\\ &= \frac{\mathcal{N}_{nm}(E)}{\mathcal{D}_n(E)}, \quad (6.32)\end{aligned}$$

where the numerator $\mathcal{N}_{nm}$ and the denominator $\mathcal{D}_n$ are given by

$$\begin{aligned}\mathcal{N}_{nm}(E) &\stackrel{\text{def}}{=} \delta_{nm}\left(1-\lambda\frac{H'_{nn}}{E-E_n^0}-\lambda^2\sum_{k\neq n}\frac{|H'_{nk}|^2}{(E_n^0-E_k^0)(E-E_n^0)}\right)\\ &\quad+\lambda\frac{H'_{nm}}{E-E_m^0}+\lambda^2\sum_{k\neq n}\frac{H'_{nk}H'_{km}}{(E-E_k^0)(E-E_m^0)}+\ldots \quad (6.33)\\ \mathcal{D}_n(E) &\stackrel{\text{def}}{=} E-E_n^0-\lambda H'_{nn}-\lambda^2\sum_{k\neq n}\frac{|H'_{nk}|^2}{E_n^0-E_k^0}+\ldots. \quad (6.34)\end{aligned}$$

The terms belonging to $\mathcal{N}_{nm}$ are unambiguously distinguished from those of $\mathcal{D}_n$ by the fact that $\mathcal{D}_n$ is independent of m and of the form $\mathcal{D}_n(E)=E-E_n^0-\lambda a+\lambda^2 b+\ldots$,

while $\mathcal{N}_{nm}$ varies with m and remains finite in the vicinity of $E = E_n^0$. So, first take $n = m$ to identify $\mathcal{D}_n$ and $\mathcal{N}_{nn}$:

$$\begin{aligned}
&\frac{1}{E - E_n^0} + \lambda \frac{H'_{nn}}{(E - E_n^0)^2} \\
&+\lambda^2 \sum_{k \neq n} \frac{|H'_{nk}|^2}{(E - E_n^0)^2 (E - E_k^0)} + \lambda^2 \frac{(H'_{nn})^2}{(E - E_n^0)^3} + \dots \\
= \quad & \frac{1}{E - E_n^0} \left[1 + \lambda \frac{H'_{nn}}{(E - E_n^0)} + \lambda^2 \frac{(H'_{nn})^2}{(E - E_n^0)^2} + \lambda^2 \sum_{k \neq n} \frac{|H'_{nk}|^2}{(E - E_n^0)(E_n^0 - E_k^0)} + \dots \right] \\
& \times \left[1 + \lambda^2 \sum_{k \neq n} \frac{|H'_{nk}|^2}{E - E_n^0} \left(\frac{1}{E - E_k^0} - \frac{1}{E_n^0 - E_k^0} \right) + \dots \right] \\
= \quad & \left[E - E_n^0 - \lambda H'_{nn} - \lambda^2 \sum_{k \neq n} \frac{|H'_{nk}|^2}{E_n^0 - E_k^0} + \dots \right]^{-1} \\
& \times \left[1 - \lambda^2 \sum_{k \neq n} \frac{|H'_{nk}|^2}{(E - E_k^0)(E_n^0 - E_k^0)} + \dots \right];
\end{aligned}$$

then multiply the left-hand side of (6.32) by this $\mathcal{D}_n$ to calculate $\mathcal{N}_{nm}$ for $n \neq m$.

From the denominator function $\mathcal{D}_n(E)$ we can conclude that the pole occurs at the shifted eigenvalue $E = E_n$, where

$$E_n = E_n^0 + \lambda H'_{nn} + \lambda^2 \sum_{k \neq n} \frac{|H'_{nk}|^2}{E_n^0 - E_k^0} + \dots, \tag{6.35}$$

while from the numerator function $\mathcal{N}_{nn}(E_n)$ for $m = n$, evaluated at $E = E_n$, it follows that

$$|\langle E_n^0 | E_n \rangle|^2 = 1 - \lambda^2 \sum_{k \neq n} \frac{|H'_{nk}|^2}{(E_k^0 - E_n^0)^2} + \dots,$$

or, to within an arbitrary phase,

$$\langle E_n^0 | E_n \rangle = 1 - \frac{1}{2} \lambda^2 \sum_{k \neq n} \frac{|H'_{nk}|^2}{(E_k^0 - E_n^0)^2} + \dots. \tag{6.36}$$

For $m \neq n$, the numerator function becomes, to order λ^2,

$$\begin{aligned}
\mathcal{N}_{nm}(E_n) &= \langle E_n^0 | E_n \rangle \langle E_n | E_m^0 \rangle \\
&= \lambda \frac{H'_{nm}}{E_n - E_m^0} + \lambda^2 \sum_{k \neq n} \frac{H'_{nk} H'_{km}}{(E_n - E_k^0)(E_n - E_m^0)} + \dots
\end{aligned}$$

from which it follows (since we found that $\langle E_n^0 | E_n \rangle$ differs from 1 only by a term of order λ^2) that for $m \neq n$,

$$\langle E_m^0 | E_n \rangle = \lambda \frac{H'_{mn}}{E_n - E_m^0} + \lambda^2 \sum_{k \neq n} \frac{H'_{mk} H'_{kn}}{(E_n - E_k^0)(E_n - E_m^0)} + \dots$$

$$= \lambda \frac{H'_{mn}}{E_n^0 - E_m^0}$$
$$+\lambda^2 \left[\sum_{k \neq n} \frac{H'_{mk} H'_{kn}}{(E_n^0 - E_k^0)(E_n^0 - E_m^0)} - \frac{H'_{mn} H'_{nn}}{(E_n^0 - E_m^0)^2} \right] + \dots . \qquad (6.37)$$

Equations (6.35) and (6.37) are the eigenvalue shifts and the expansion coefficients of the eigenfunctions to second-order perturbation theory, provided there is no degeneracy and the unperturbed levels are not too closely spaced.

6.2.2 Closely spaced levels

In order to handle the more complicated case in which two unperturbed levels, call them E_1^0 and E_2^0, are so closely spaced that there is a possibility that a small perturbation may actually make them cross, let us write the resolvent operator $(E - \mathbf{H})^{-1}$ in a convenient form by introducing the mutually orthogonal projection operators

$$\mathsf{P}_1 \stackrel{\text{def}}{=} |E_1^0\rangle\langle E_1^0|, \qquad \mathsf{Q}_1 \stackrel{\text{def}}{=} \mathbb{1} - \mathsf{P}_1 = \sum_{n \neq 1} |E_n^0\rangle\langle E_n^0|$$

and defining

$$\mathsf{Q}' \stackrel{\text{def}}{=} \mathsf{Q}_1 (E - \mathbf{H}_0)^{-1} \mathsf{Q}_1 = \sum_{n \neq 1} \frac{|E_n^0\rangle\langle E_n^0|}{E - E_n^0}.$$

A bit of straight-forward algebra[9] leads to the following expression:

$$(E - \mathbf{H}_0 - \lambda \mathbf{H}')^{-1} = \mathsf{Q}'(\mathbb{1} - \lambda \mathbf{H}' \mathsf{Q}')^{-1}$$
$$+ \quad (\mathbb{1} - \lambda \mathsf{Q}' \mathbf{H}')^{-1} \mathsf{P}_1 (E - E_1^0 - \mathcal{M})^{-1} \mathsf{P}_1 (\mathbb{1} - \lambda \mathbf{H}' \mathsf{Q}')^{-1}, \quad (6.38)$$

where

$$\mathcal{M} \stackrel{\text{def}}{=} \lambda \mathsf{P}_1 (\mathbb{1} - \lambda \mathbf{H}' \mathsf{Q}')^{-1} \mathbf{H}' \mathsf{P}_1.$$

Next, we separate out the second level by means of the projection

$$\mathsf{P}_2 \stackrel{\text{def}}{=} |E_2^0\rangle\langle E_2^0|$$

and define

$$\mathsf{Q}'' \stackrel{\text{def}}{=} \mathsf{Q}' - \frac{\mathsf{P}_2}{E - E_2^0}.$$

[9] **Do it!**

Regarding $E - E_2^0$ as of the same order as $\lambda|H'_{22}|$, one then obtains, to first order,

$$\mathcal{M} = \lambda \mathsf{P}_1 \mathbf{H}' \mathsf{P}_1 + \frac{\lambda^2}{E - E_2^0} \mathsf{P}_1 \mathbf{H}' \mathsf{P}_2 \left(\mathbb{1} - \frac{\lambda}{E - E_2^0} \mathsf{P}_2 \mathbf{H}' \mathsf{P}_2 \right)^{-1} \mathsf{P}_2 \mathbf{H}' \mathsf{P}_1 + \ldots .$$

Now, according to (6.38), the position of the pole of $(E - \mathbf{H})^{-1}$ near E_1^0 is obtained, to first order, from the equation

$$E = E_1^0 + \langle E_1^0 | \mathcal{M} | E_1^0 \rangle = E_1^0 + \lambda H'_{11} + \frac{\lambda^2 |H'_{12}|^2}{E - E_2^0 - \lambda H'_{22}},$$

which means, with $E'_1 \stackrel{\text{def}}{=} E_1^0 + \lambda H'_{11}$ and $E'_2 \stackrel{\text{def}}{=} E_2^0 + \lambda H'_{22}$,

$$(E - E'_1)(E - E'_2) = \lambda^2 |H'_{12}|^2,$$

the solutions of which are (assuming $E_1^0 < E_2^0$),

$$\begin{aligned} E_a &= \frac{1}{2}[E_1^0 + E_2^0 + \lambda(H'_{11} + H'_{22})] \\ &- \frac{1}{2}\sqrt{[E_1^0 - E_2^0 + \lambda(H'_{11} - H'_{22})]^2 + 4\lambda^2 |H'_{12}|^2} < E'_1, \end{aligned}$$

$$\begin{aligned} E_b &= \frac{1}{2}[E_1^0 + E_2^0 + \lambda(H'_{11} + H'_{22})] \\ &+ \frac{1}{2}\sqrt{[E_1^0 - E_2^0 + \lambda(H'_{11} - H'_{22})]^2 + 4\lambda^2 |H'_{12}|^2} > E'_2. \end{aligned}$$

Take the case when $H'_{11} > 0$, $H'_{22} < 0$, and $H'_{12} \neq 0$. As shown in Figure 6.5, what is happening here is that the two levels, which in ordinary first-order perturbation theory would cross as λ increases, because E'_1 is rising linearly and E'_2 is declining, are apparently *repelled* by one another; for large enough λ, E'_b, which started out decreasing from E_2^0, will rise like E'_1, and E'_a, which began by rising from E_1^0, will decrease like E'_2, and the levels never actually intersect. Thus, what at first sight looks like a crossing of the two levels with increasing λ, on closer inspection reveals itself as no crossing at all, but a switching of the identities of the two levels, with a small remaining gap.

6.2.3 The case of degeneracy

The case in which there is degeneracy in the unperturbed spectrum and a perturbation removes it—that is, it splits a single energy level into several distinct ones—is the physically most interesting one, because minute splittings are much easier to observe experimentally than simple small shifts.

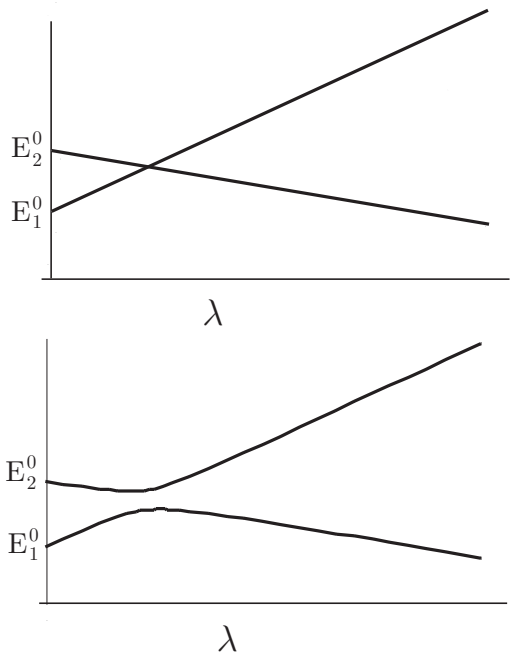

FIGURE 6.5. The upper figure shows the change of two energy levels as a function of λ according to ordinary first-order perturbation theory; the lower figure shows the way they actually change when the proximity of the levels is taken into account.

It is therefore important to note that, if there is degeneracy in the unperturbed spectrum, the procedure given above breaks down. The underlying reason for this is not hard to understand: on the degenerate subspace (the eigenspace of $\mathbf{H}_0$ at the degenerate eigenvalue E_n^0) the (normalized) basis vectors, as eigenvectors of $\mathbf{H}_0$, are not uniquely determined (to within a phase factor); any orthogonal linear combination of them would serve as well. But in order for the λ-expansion to make sense, the eigenvectors have to be continuous functions of λ. So if the perturbation removes the degeneracy, envision starting with the perturbed eigenfunctions and continuously decreasing λ to 0, thus turning the perturbation off; the eigenvectors will then approach a specific set of eigenvectors on the degenerate subspace, the "correct linear combinations" of the original basis vectors, which, in the absence of the perturbation, are arbitrarily oriented. If an "incorrect linear combination" had been chosen when attempting to start the perturbation expansion, these basis vectors would not be continuously connected to the perturbed eigenvectors, and the perturbation theory would fail at the lowest order of λ that removes the degeneracy. Let's see how this, in fact, manifests itself.

To begin with, consider first-order perturbation theory. While the first-order energy shift $\lambda H'_{nn}$ of the n^{th} level seems to present no difficulties if $E^0_n = E^0_{n'}$, you will notice from (6.36) and (6.37) that there is going to be trouble because of vanishing energy-denominators. To remove the problem, we make use of the freedom to rotate the eigenvectors $|E^0_{n_j}\rangle$ on the degenerate subspace and define a new set of orthonormal eigenvectors by

$$|E^0_{n_l}\rangle\!\rangle \stackrel{\text{def}}{=} \sum_j M_{jl}|E^0_{n_j}\rangle, \tag{6.39}$$

where the matrix M is unitary. The new matrix elements of the perturbation on the degenerate subspace are then

$$\widehat{H'_{n_i n_j}} \stackrel{\text{def}}{=} \langle\!\langle n_i|\mathbf{H}'|n_j\rangle\!\rangle = (M^{-1}H'M)_{ij}, \tag{6.40}$$

and, since $\mathbf{H}'$ is Hermitian, M may be chosen so as to *diagonalize* the perturbation matrix H' on the E^0_n-eigenspace of $\mathbf{H}_0$. The effect of using $\widehat{H'_{nm}}$ instead of H'_{nm} in the expansions, beginning with (6.32), is then that the summations exclude all the terms for which $E^0_k = E^0_n$, and in the formula that replaces (6.37) for $\langle\!\langle E^0_{n_i}|E_{n_j}\rangle$, $i \neq j$, the linear term in λ is missing. The numbers $\widehat{H'_{nn}}$ that appear in (6.35) as the first-order energy shifts are now the *eigenvalues* of the perturbation matrix on the degenerate subspace, and they are calculated in the usual manner as the zeros of the secular equation for the matrix.

If the eigenvalues of the perturbation matrix H' on the degenerate subspace are not all equal, so that the diagonal matrix elements $\widehat{H'_{n_i n_i}}$ are not all the same, then (6.35) shows that *the degeneracy is (at least in part) removed in first order,* that is, whereas several unperturbed energies are degenerate, their first-order perturbed values are no longer all equal. For the calculation to work, the *correct linear combinations* of unperturbed eigenfunctions have to be chosen, namely, those that diagonalize the perturbation matrix on the degenerate subspace, and these are the eigenfunctions of $\mathbf{H}_0$ that are the continuous limits of the eigenfunctions of $\mathbf{H}_0 + \lambda\mathbf{H}'$ as $\lambda \to 0$.

An example. Denote the eigenvalues of the unperturbed Hamiltonian by E^0_n, $n = 0, 1, 2, \ldots$, and let the matrix of a perturbation $\mathbf{H}'$ on the basis of the eigenstates of $\mathbf{H}_0$ be H'_{nm}. Assume that E^0_0 is nondegenerate and that E^0_1 is twofold degenerate. We then have to find the eigenvalues of the matrix

$$\lambda \begin{pmatrix} H'_{11} & H'_{12} \\ H'_{21} & H'_{22} \end{pmatrix},$$

which means we have to solve the secular equation

$$\begin{vmatrix} \lambda H'_{11} - \varepsilon & \lambda H'_{12} \\ \lambda H'_{21} & \lambda H'_{22} - \varepsilon \end{vmatrix} = 0,$$

or

$$\varepsilon^2 - \varepsilon\lambda(H'_{11} + H'_{22}) + \lambda^2(H'_{11}H'_{22} - |H'_{12}|^2) = 0.$$

Therefore, the degeneracy is removed in first order and the two energy shifts are, in first order perturbation theory,

$$E_1^{\pm} - E_1^0 = \lambda\left[\frac{1}{2}(H'_{11} + H'_{22}) \pm \frac{1}{2}\sqrt{(H'_{11} - H'_{22})^2 + 4|H'_{12}|^2}\right],$$

leading to a split of the originally degenerate level into two levels with an energy difference given by

$$\Delta = \lambda\sqrt{(H'_{11} - H'_{22})^2 + 4|H'_{12}|^2}.$$

The only way for this split to vanish, that is, for the degeneracy not to be removed in first order, is for $H'_{11} = H'_{22}$ and $H'_{12} = 0$, which makes the matrix H' equal to $H'_{11}\mathbf{1}$.

Degeneracy removal in second order

If, after the diagonalization of the matrix H' on the degenerate subspace, the diagonal values $\widehat{H'_{n_i n_i}}$ are all equal, then the degeneracy is not removed in first order. At the same time, this means that all the eigenvalues of the matrix of H' on that subspace are the same, which implies that H' there is a multiple of the unit matrix, and the diagonalization was not necessary because it must have been diagonal to begin with. Equation (6.37) shows in that case an analogous problem arises in second order for $\langle E^0_{n'}|E_n\rangle$ if $E^0_{n'} = E^0_n$. If the degeneracy is removed in first order, then $E_n - E^0_{n'}$ is of order λ and the first line of (6.37) shows that the λ^2-term becomes linear in λ, with no ensuing difficulty. However, if the degeneracy is not removed in first order, then $E_n - E^0_{n'}$ in the denominator is of order λ^2, and the λ^2-term becomes independent of λ, thus remaining nonzero as $\lambda \to 0$.

The remedy is analogous to the previous case. It is now necessary to diagonalize the matrix

$$\mathcal{M}_{n'n} = \sum_{k\neq n} \frac{H'_{n'k}H'_{kn}}{E^0_n - E^0_k} = \langle E^0_{n'}|\mathbf{H}'\mathsf{Q}_n(E^0_n - \mathbf{H}_0)^{-1}\mathsf{Q}_n\mathbf{H}'|E^0_n\rangle,$$

on the degenerate subspace, where Q_n is the projection on the orthogonal complement of the eigenspace of $\mathbf{H}_0$ at E^0_n, and the basis functions that do this diagonalization are the "proper linear combinations" of the original basis functions that fit continuously with the perturbed eigenfunctions. (Since, by assumption, the matrix H' on the degenerate subspace is a multiple of the unit matrix, this change of basis has no effect on it.)

6.2.4 Second-order shifts from first-order states

Suppose the wave function has been calculated to *first order* in λ, that is, according to (6.37) and (6.36),

$$\langle E_k^0|E_n\rangle = \langle E_k^0|E_n^1\rangle = \lambda\frac{H'_{kn}}{E_k^0 - E_n^0}, \qquad k \neq n,$$

while $\langle E_n^0|E_n^1\rangle = 1$. We then find that to second order in powers of λ

$$\lambda\langle E_n^1|\mathbf{H}'|E_n^1\rangle = \lambda H'_{nn} + 2\lambda^2 \sum_{k\neq n} \frac{|H'_{kn}|^2}{E_n^0 - E_k^0} + \dots$$

and

$$\begin{aligned}
\langle E_n^1|\mathbf{H}_0|E_n^1\rangle &= E_n^0 + \lambda^2 \sum_{k\neq n} \frac{E_k^0|H'_{kn}|^2}{(E_n^0 - E_k^0)^2} + \dots \\
&= E_n^0 \left(1 + \lambda^2 \sum_{k\neq n} \frac{|H'_{kn}|^2}{(E_n^0 - E_k^0)^2}\right) \quad - \quad \lambda^2 \sum_{k\neq n} \frac{|H'_{kn}|^2}{E_n^0 - E_k^0} \quad + \dots \\
&= \left(1 + \lambda^2 \sum_{k\neq n} \frac{|H'_{kn}|^2}{(E_n^0 - E_k^0)^2}\right) \left(E_n^0 - \lambda^2 \sum_{k\neq n} \frac{|H'_{kn}|^2}{E_n^0 - E_k^0}\right) + \dots,
\end{aligned}$$

so that

$$\begin{aligned}
\langle E_n^1|\mathbf{H}_0 + \lambda\mathbf{H}'|E_n^1\rangle \quad = \quad & \left(1 + \lambda^2 \sum_{k\neq n} \frac{|H'_{kn}|^2}{(E_n^0 - E_k^0)^2}\right) \\
& \times \left(E_n^0 + \lambda H'_{nn} + \lambda^2 \sum_{k\neq n} \frac{|H'_{kn}|^2}{E_n^0 - E_k^0}\right) + \dots.
\end{aligned}$$

Furthermore,[10]

$$\langle E_n^1|E_n^1\rangle = 1 + \lambda^2 \sum_{k\neq n} \frac{|H'_{kn}|^2}{(E_n^0 - E_k^0)^2} + \dots.$$

Comparison with (6.35) shows that therefore, if we define the functional

$$\mathfrak{E}_{\mathbf{H}}(\Psi) \stackrel{\text{def}}{=} \frac{(\Psi, \mathbf{H}\Psi)}{(\Psi, \Psi)}, \tag{6.41}$$

then to second order [writing $E_n \stackrel{\text{def}}{=} E_n^{(2)} + O(\lambda^3)$]

$$E_n^{(2)} = \mathfrak{E}_{\mathbf{H}}(\Psi_{E_n^1}), \tag{6.42}$$

[10]The first-order state vectors are normalized to first order, but not to second order!

which means that the energy is given correctly to *second order* as the expectation value of $\mathbf{H}$ in the eigenstate as calculated correctly to *first order,* just as the energy is given correctly to first order as the expectation value of the Hamiltonian in an eigenstate that is correct to zeroth order (see (6.35)). The underlying reason for this will become aparent in Section 6.3.

6.2.5 Effects of symmetry

Suppose that the Hamiltonian $\mathbf{H}_0$ has a symmetry group $\mathcal{G}$ (for example, $\mathbf{H}_0$ is rotationally invariant) and the degeneracy of the level E_n^0 is *normal,* that is, E_n^0 belongs to an irreducible representation of $\mathcal{G}$ (in the rotationally invariant case, each level belongs to a specific value of l, and there is just the m degeneracy). If $\mathbf{H}'$ has the same symmetry group, then $\mathbf{H}$ must have *at least* the same degeneracy; $\mathbf{H}'$ cannot remove this degeneracy in any order. This follows, first of all, from the fact that $\mathcal{G}$ is then a symmetry group of $\mathbf{H} = \mathbf{H}_0 + \lambda\mathbf{H}'$, but it can also be seen in each order of perturbation theory, because Schur's lemma E.3 in Appendix E tells us that all matrix elements H'_{nk} that connect different irreducible representations vanish, and that on each degenerate subspace (i.e., for each irreducible representation), the matrix $\{H'_{nn'}\}$ is a multiple of the unit matrix.

If $\mathbf{H}_0$ has accidental degeneracy as well (say, a level of one angular momentum l happens to have the same energy as a level of another angular momentum l'), this degeneracy can be removed by $\mathbf{H}'$, and using the basis vectors belonging to irreducible representations of $\mathcal{G}$ is automatically the correct choice to make perturbation theory work.

Next, suppose that $\mathbf{H}'$ has less symmetry than $\mathbf{H}_0$ (say, it has only axial symmetry). In that case, the symmetry group $\mathcal{G}'$ of $\mathbf{H}$ is a subgroup of $\mathcal{G}$, and every representation of $\mathcal{G}$ is also a representation of $\mathcal{G}'$. However, if it is irreducible as a representation of $\mathcal{G}$, it need not be irreducible as a representation of $\mathcal{G}'$. Let E_n belong to the κ-dimensional representation Γ^κ, irreducible for $\mathcal{G}$, and suppose that Γ^κ contains the two irreducible representations $\Gamma_1^{\kappa_1}$ and $\Gamma_2^{\kappa_2}$ of $\mathcal{G}'$: $\Gamma^\kappa = \Gamma_1^{\kappa_1} \bigotimes \Gamma_2^{\kappa_2}$. Then the originally κ-fold degenerate level E_n^0 can be split into two levels, one of degeneracy κ_1 and the other of degeneracy κ_2, and again, the basis vectors according to the irreducible representations $\Gamma_1^{\kappa_1}$ and $\Gamma_2^{\kappa_2}$ of $\mathcal{G}'$ are the correct choice, because on these bases H' is diagonal.

There is, however, another possibility: there may be accidental degeneracy in $\mathbf{H}_0$ in which the states of the level transform *under the same* irreducible representation several times over. In such a case, the arguments given are insufficient, and it becomes necessary to do additonal work to find the "correct linear combinations."

6.2.6 Relativistic correction to hydrogenic levels

The relativistic Hamiltonian of a particle of electric charge e subject to a scalar potential ϕ differs from its nonrelativistic form; since the relativistic energy-momentum relation for such a particle is

$$(E - e\phi)^2 = M^2c^4 + c^2\vec{p}^2,$$

Schrödinger already introduced a relativistic wave equation that has the same form as the nonrelativistic equation that bears his name, but with the Hamiltonian operator defined by

$$(\mathbf{H} - e\phi)^2 = M^2c^4 + c^2\vec{\mathbf{p}}^2. \tag{6.43}$$

Rather than fully treating the resulting modified equation, which is subject to some special difficulties (we shall discuss relativistic wave equations in Chapter 10), it is sometimes useful and sufficient to apply a first-order relativistic correction to the kinetic energy,

$$\mathbf{T} = \sqrt{\vec{\mathbf{p}}^2c^2 + M^2c^4} - Mc^2 = \frac{\vec{\mathbf{p}}^2}{2M} - \frac{(\vec{\mathbf{p}}^2/2M)^2}{2Mc^2} + \ldots, \tag{6.44}$$

which means that the first-order correction to the nonrelativistic Hamiltonian $\mathbf{H}_0$ of a hydrogenic atom is given by

$$\mathbf{H}'_{\text{rel}} = -\frac{(\vec{\mathbf{p}}^2/2M)^2}{2Mc^2} = -\frac{1}{2Mc^2}(\mathbf{H}_0 + \frac{Ze^2}{r})^2. \tag{6.45}$$

In first-order perturbation theory, the energy shift is thus

$$\Delta E = -\frac{1}{2Mc^2}\left[E^{02} + 2Ze^2E^0\langle r^{-1}\rangle + Z^2e^4\langle r^{-2}\rangle\right], \tag{6.46}$$

which means we need to calculate the expectation values $\langle r^{-2}\rangle$ and $\langle r^{-1}\rangle$ in hydrogenic states.

The radial Hamiltonian for a hydrogenic atom is given by

$$\mathbf{H}_0 = \frac{1}{2M}\left[\mathbf{p}_r^2 + \frac{\hbar^2 l(l+1)}{r^2}\right] - \frac{Ze^2}{r}.$$

To calculate $\langle r^{-2}\rangle$, employ the Hellmann-Feynman formula (4.38), differentiating $\mathbf{H}_0$ with respect to the parameter l,[11]

$$\frac{\partial E^0}{\partial l} = \left\langle \frac{\partial \mathbf{H}_0}{\partial l} \right\rangle = \frac{\hbar^2}{M}(l + \frac{1}{2})\langle r^{-2}\rangle.$$

On the other hand, according to (4.70), the energy eigenvalues of a hydrogenic atom are

$$E_n^0 = -\frac{Z^2e^2}{2n^2a_0}, \qquad n = l + 1 + n_r,$$

---

[11] As already noted earlier, in the radial Hamiltonian, l is a parameter that can be changed continuously; the requirement that l be an integer arises from the angular part of the Hamiltonian.

where

$$a_0 \stackrel{\text{def}}{=} \frac{\hbar^2}{Me^2}, \tag{6.47}$$

is the Bohr radius, so that

$$\frac{\partial E^0}{\partial l} = \frac{\partial E^0}{\partial n} = \frac{Z^2 e^2}{n^3 a_0},$$

with the result

$$\langle r^{-2} \rangle = \frac{Z^2}{n^3 a_0^2} \frac{1}{l + \frac{1}{2}}. \tag{6.48}$$

In order to calculate $\langle r^{-1} \rangle$, differentiate $\mathbf{H}_0$ with respect to Z and again apply the Hellmann-Feynman formula:

$$\left\langle \frac{\partial \mathbf{H}_0}{\partial Z} \right\rangle = -e^2 \langle r^{-1} \rangle = \frac{\partial E^0}{\partial Z} = -\frac{Ze^2}{n^2 a_0};$$

therefore

$$\langle r^{-1} \rangle = \frac{Z}{n^2 a_0}. \tag{6.49}$$

Inserting the results (6.48) and (6.49) in (6.46) yields the following relativistic shift in the n^{th} energy level of a hydrogenic atom

$$\Delta E_n^{\text{rel}} = \frac{Z^4 \alpha^4}{2n^3} \left(\frac{3}{4n} - \frac{1}{l + \frac{1}{2}} \right) Mc^2, \tag{6.50}$$

in which

$$\alpha \stackrel{\text{def}}{=} \frac{e^2}{\hbar c} \tag{6.51}$$

is the *fine-structure constant*, whose dimensionless value is $\alpha \simeq 1/137.037$.

It should be noted that we applied perturbation theory and managed to calculate small corrections to the eigenvalues, even though the perturbation is not "small" in an operator sense; it is, however, small for these particular states. Note also that relativity destroys the l-degeneracy of the nonrelativistic hydrogen atom. (This is true, however, only if the spin of the electron is neglected, as we shall see later.)

6.2.7 The Stark effect

Let us apply stationary perturbation theory to a physical system consisting of a hydrogen atom in a weak, uniform external electric field of magnitude $\mathcal{E}$. In this case, the unperturbed Hamiltonian is

$$\mathbf{H}_0 = -\frac{\hbar^2}{2M} \nabla^2 - \frac{e^2}{r},$$

and if the z-axis is chosen in the direction of the electric field, the interaction with the electric field is given by

$$\mathbf{H}' = e\phi = -e\mathcal{E}\mathbf{z}.$$

This perturbation, of course, is not small in the operator sense; it is unbounded. Not only that, but the total potential energy tends to $-\infty$ in one direction, so that the system cannot have any bound states at all! In the presence of the electric field, the original bound states, in fact, all become unstable states, with a finite lifetime. Nevertheless, perturbation theory works, and the shifts from the original energies of the stable states to those of the new, unstable ones will be small if the electric field is sufficiently weak.

In first order, the shift in the energy of the n^{th} level is simply

$$\Delta E_n = H'_{nn} = -e\mathcal{E}\int d^3r\, z|\psi_n|^2$$

if ψ_n is the n^{th} eigenfunction of the unperturbed Hamiltonian. Consider first the *ground state.* In that case we must have $l = m = 0$, and the eigenfunction is spherically symmetric; as a result, the integral on the right vanishes, since the integrand is anti-symmetric. *There is no first-order Stark effect in the ground state of the hydrogen atom.*

The first excited level is fourfold degenerate, with the angular momentum quantum numbers $(l, m) = (0, 0), (1, 0), (1, 1), (1, -1)$. Because $\mathbf{z}$ and $\mathbf{L}_z$ commute, the matrix elements of $\mathbf{z}$ connecting different values of m necessarily vanish: $0 = \langle m|\mathbf{z}\mathbf{L}_z - \mathbf{L}_z\mathbf{z}|m'\rangle = (m' - m)\langle m|\mathbf{z}|m'\rangle$. Therefore, there are nonvanishing off-diagonal matrix elements for the first two of the above listed angular momenta only, and these are

$$\begin{aligned} H'_{2l0,2l'0} &= -e\mathcal{E}\int d^3r\, \psi^*_{2l0} z \psi_{2l'0} \\ &= -e\mathcal{E}2\pi\sqrt{\frac{2}{2l+1}}\sqrt{\frac{2}{2l'+1}} \\ &\quad \times \int_0^\infty dr r^2 \int_{-1}^1 d\cos\theta\, R_l(r)R_{l'}(r) r\cos\theta P_l(\cos\theta)P_{l'}(\cos\theta). \end{aligned}$$

Since $\int_{-1}^1 d\cos\theta\, \cos\theta P_l^2(\cos\theta) = 0$, the diagonal elements of this 2×2 submatrix are zero, and the off-diagonal elements are[12]

$$H'_{200,210} = H'_{210,200} = -\frac{2\pi\mathcal{E}}{\sqrt{3}} a_0 \frac{2}{3}\frac{1}{2!3!^{3/2}4}\int_0^\infty d\rho\, \rho^4 L_2^1(\rho)L_3^3(\rho)e^{-\rho} = 3ea_0\mathcal{E},$$

where a_0 is the Bohr radius given by (6.47). So the perturbation matrix is

$$H' = 3ea_0\mathcal{E}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

[12]The two relevant Laguerre polynomials are $L_3^3(\rho) = -6$ and $L_2^1(\rho) = 2(\rho - 2)$. **You should verify this.**

whose eigenvalues are $\pm 3ea_0\mathcal{E}$. As a result, the two degenerate levels $(2,1,0)$ and $(2,0,0)$ are split, one shifted upward by $3ea_0\mathcal{E}$, the other downward by the same amount, while there still remain the two degenerate, unshifted levels $(2,1,1)$ and $(2,1,-1)$. The latter two will, of course, be shifted in second order, but for a weak electric field, this second-order shift will be much smaller than the first order.

Now recall that the interaction energy of an electric point dipole $\vec{a}$ in an electric field $\vec{E}$ is $-\vec{a}\cdot\vec{E}$. Therefore, the hydrogen atom in a state with $n=2$ acts as though it had an electric dipole moment of magnitude $-3ea_0$, which can be oriented parallel, anti-parallel, or perpendicular to $\vec{E}$. In the ground state, no such permanent electric dipole moment can exist; there can then be only an induced dipole moment, so that the interaction energy is proportional to $\mathcal{E}^2$: this is the quadratic Stark effect, which manifests itself in second order. Note that the electric dipole moment of a charged system is given in terms of its charge density ρ by $\int d^3r\,\vec{r}\rho$, which quantum mechanically becomes $e\int d^3r\,\vec{r}|\psi|^2$. Because $\mathbf{H}_0$ is rotationally invariant, this would vanish on parity grounds, if it weren't for the degeneracy; therefore in its first excited state the hydrogen atom can have a permanent electric dipole moment, a spontaneous breaking of its symmetry, only because of the degeneracy of that state!

6.3 Variational Methods

In this section we shall describe an approximation procedure that is particularly useful for precise numerical calculations.

6.3.1 Bound states

Suppose that E is an eigenvalue of $\mathbf{H}$ and that Ψ_E is a corresponding eigenvector, $\mathbf{H}\Psi_E = E\Psi_E$; then the functional $\mathfrak{E}_{\mathbf{H}}(\Psi)$ defined by (6.41), which is simply the expression for the expectation value of $\mathbf{H}$ in an eigenstate, is such that $\mathfrak{E}_{\mathbf{H}}(\Psi_E) = E$. Furthermore, this expression for the eigenvalue has the remarkable property of *being stationary with respect to small variations* in Ψ in a neighborhood of an eigenvector Ψ_E of $\mathbf{H}$. To see this, calculate

$$\begin{aligned}\delta\mathfrak{E}_{\mathbf{H}}(\Psi) &= \frac{(\delta\Psi,\mathbf{H}\Psi)+(\Psi,\mathbf{H}\delta\Psi)}{(\Psi,\Psi)} - (\Psi,\mathbf{H}\Psi)\frac{(\delta\Psi,\Psi)+(\Psi,\delta\Psi)}{(\Psi,\Psi)^2}\\ &= \frac{(\delta\Psi,[(\Psi,\Psi)\mathbf{H}-(\Psi,\mathbf{H}\Psi)]\Psi)+([(\Psi,\Psi)\mathbf{H}-(\Psi,\mathbf{H}\Psi)]\Psi,\delta\Psi)}{(\Psi,\Psi)^2},\end{aligned}$$

for arbitrary small $\delta\Psi$, to first order in $\|\,\delta\Psi\,\|$, and this vanishes when $\Psi=\Psi_E$. What is more, it follows from the second line that if $\delta\mathfrak{E}_{\mathbf{H}}(\Psi)=0$

for arbitrary small $\delta\Psi$, then $(\Psi,\Psi)\mathbf{H}\Psi - (\Psi,\mathbf{H}\Psi)\Psi = 0$,[13] which implies that Ψ is an eigenvector of $\mathbf{H}$ with the eigenvalue $E = (\Psi,\mathbf{H}\Psi)/(\Psi,\Psi)$. Therefore, we have the following:

Theorem 6.1 *The functional* $\mathfrak{E}_{\mathbf{H}}(\Psi)$ *defined by (6.41) is stationary near* $\Psi = \Psi_E$ *if and only if* Ψ_E *is an eigenvector of* $\mathbf{H}$ *with the eigenvalue* $E = \mathfrak{E}_{\mathbf{H}}(\Psi_E)$.

In view of this theorem, the *point spectrum of* $\mathbf{H}$ *may be defined as the set of those values of* $\mathfrak{E}_{\mathbf{H}}(\Psi)$ *at which this functional is stationary.*

We can say more: if E_0 is the energy of the ground state, i.e., the smallest eigenvalue of $\mathbf{H}$, then for all Ψ,

$$\mathfrak{E}_{\mathbf{H}}(\Psi) \geq E_0,$$

as is easily seen[14] by expanding Ψ on the basis of the eigenstates of $\mathbf{H}$. So at $\Psi = \Psi_{E_0}$, the stationary point of $\mathfrak{E}$ is actually a *global minimum* with the value E_0. This implies that if we have a *pretty good approximation* Ψ to the ground state Ψ_{E_0}, then the value of $\mathfrak{E}_{\mathbf{H}}(\Psi)$ is a *very good approximation* to the ground-state energy. We can put it this way: if the trial function Ψ^{approx} differs from the exact eigenfunction Ψ^{exact} by an unknown error-vector of order ε,

$$\Psi^{\text{approx}} = \Psi^{\text{exact}} + \varepsilon\Psi',$$

then the eigenvalue calculated as the expectation value of $\mathbf{H}$ in the state Ψ^{approx},

$$\mathfrak{E}_{\mathbf{H}}(\Psi^{\text{approx}}) = E^{\text{approx}} = E^{\text{exact}} + \varepsilon^2 c + \ldots,$$

is correct with an error of order ε^2. *This explains why the perturbation of a state vector, calculated by means of (6.41) and the use of a state vector that is correct to first order, leads to an eigenvalue that is correct to second order, and why the zeroth-order state is sufficient to calculate the eigenvalue to first order.* Moreover, the thus calculated approximate eigenvalue always lies *above* E_0. As a result, the perturbation result (6.35) can be sharpened for the ground state,

$$E_0^{\text{exact}} \leq E_0^0 + \lambda H'_{00} - \lambda^2 \sum_{k\neq 0} \frac{|H'_{0k}|^2}{E_k^0 - E_0^0} + \ldots. \tag{6.52}$$

Suppose the functional $\mathfrak{E}_{\mathbf{H}}(\Psi)$ is restricted to vectors Ψ that are orthogonal to the ground state Ψ_0. In that case the value of the so restricted functional is never less than the energy of the first excited state,[15]

$$\mathfrak{E}_{\mathbf{H}}(\Psi)|_{(\Psi,\Psi_0)=0} \geq E_1,$$

[13] First you conclude that its real part vanishes, and then, replacing $\delta\Psi$ by $i\delta\Psi$, that its imaginary part vanishes.

[14] **Show it.**

[15] **Show it.**

and it attains its global minimum there.

The important lesson to be learned is that for numerical purposes, the employment of a stationary expression such as

$$E = \langle \mathbf{H} \rangle = (\Psi_E, \mathbf{H}\Psi_E)/(\Psi_E, \Psi_E),$$

has great advantages over others, such as

$$E = (\Phi, \mathbf{H}\Psi_E)/(\Phi, \Psi_E),$$

with an arbitrary Φ, which is equally correct but not stationary near the exact state Ψ_E. After all, every numerically given eigenfunction is inevitably an approximation, and the calculation of the energy by means of the stationary formula minimizes the resulting error.

To use the variational method as a practical calculational tool, one usually starts with a family of trial functions $\Psi(\alpha)$ that depend on one or more parameters α and that are preferably chosen so as to resemble (say, in the configuration representation) in some rough way the shape of the anticipated eigenfunction. One then calculates $f(\alpha) \stackrel{\text{def}}{=} \mathfrak{E}_{\mathbf{H}}(\Psi(\alpha))$ and finds the value α_0 of α at which $f(\alpha)$ has a minimum; the best approximation to the ground-state energy obtainable from this chosen family of trial functions is then $E_0^{\text{approx}} = f(\alpha_0)$. Furthermore, if $\Psi(\alpha_0)$ is a good approximation to the exact eigenstate, then E_0^{approx} is an even better approximation to the exact ground-state energy E_0; moreover, we can be sure that $E_0^{\text{exact}} \leq E_0^{\text{approx}}$. The approximation can be made better and better by using trial functions with more and more parameters. This is, in fact, how most eigenvalues of complicated Hamiltonians of physical systems are calculated in practice in order to be compared with experimental results.

6.3.2 The helium atom

Consider the Hamiltonian of the helium atom,

$$\mathbf{H} = -\frac{\hbar^2}{2M}(\nabla_1^2 + \nabla_2^2) - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{12}} = \mathbf{H}_0 + \mathbf{H}', \tag{6.53}$$

where

$$\mathbf{H}_0 \stackrel{\text{def}}{=} -\frac{\hbar^2}{2M}(\nabla_1^2 + \nabla_2^2) - \frac{2e^2}{r_1} - \frac{2e^2}{r_2}, \qquad \mathbf{H}' \stackrel{\text{def}}{=} \frac{e^2}{r_{12}}.$$

Since $E_H^{(Z)} \stackrel{\text{def}}{=} Z^2 E_{\text{H}} = -Z^2 e^2/2a_0$, with $a_0 = \hbar^2/Me^2$, is the ground-state energy of a hydrogenic atom of nuclear charge Z, the ground-state energy of $\mathbf{H}_0$, in which the two electrons don't interact with one another, is $E_0 = 8E_{\text{H}}$, and the normalized wave function is the product of the two hydrogen wave functions, given by (see Section 4.2.5)

$$\varphi = \frac{8}{\pi a_0^3} e^{-2(r_1+r_2)/a_0}. \tag{6.54}$$

The first-order energy shift owing to the interaction $\mathbf{H}'$ between the two electrons is therefore

$$\Delta E = \frac{64e^2}{\pi^2 a_0^6} \int d^3r_1 d^3r_2 \frac{e^{-4(r_1+r_2)/a_0}}{|\vec{r}_1 - \vec{r}_2|}, \tag{6.55}$$

an integral that I will ask you do as a homework problem. The result is

$$\Delta E = \frac{5}{2}|E_{\rm H}| = \frac{5}{4}\frac{e^2}{a_0},$$

so that the shifted ground-state energy is given by

$$E_0' = (-8 + \frac{5}{2})|E_{\rm H}| = -\frac{11}{4}\frac{e^2}{a_0} = -2.75\frac{e^2}{a_0}. \tag{6.56}$$

Since the interaction energy, as you will notice, is not a small fraction of the energy of the system, first-order perturbation theory is not likely to be a very good approximation. Therefore, we are going to use the variational method to improve it, using the functional (6.41), whose value at φ of (6.54) is just (6.56). In place of (6.54), we employ as normalized trial function

$$\varphi_Z = \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0}, \tag{6.57}$$

which would be appropriate for an atom of nuclear charge Z instead of 2. With this wave function, you find[16]

$$\langle -\frac{\hbar^2}{2M}\nabla^2 \rangle = Z^2 \frac{e^2}{2a_0}, \qquad \langle \frac{1}{r} \rangle = \frac{Z}{a_0}, \qquad \langle \frac{1}{r_{12}} \rangle = \frac{5}{8}\frac{Z}{a_0},$$

so that

$$\langle \mathbf{H} \rangle = \frac{e^2}{a_0}(Z^2 - 4Z + \frac{5}{8}Z) = \frac{e^2}{a_0}(Z^2 - \frac{27}{8}Z).$$

Using Z as the variational parameter, this expression is to be minimized, with the result[17]

$$Z_{\rm min} = \frac{27}{16}, \qquad \langle \mathbf{H} \rangle_{\rm min} = -\left(\frac{27}{16}\right)^2 \frac{e^2}{a_0} = -2.85\frac{e^2}{a_0}, \tag{6.58}$$

which is to be compared to the experimental value, $E_0^{\rm exp} = -2.904e^2/a_0$. Thus the variational procedure has obviously improved the agreement. The physical interpretation of the fact that a Z value of 27/16 leads to a better result than $Z = 2$ is that each of the orbital electrons shields the Coulomb potential seen by the other, so that the *effective* nuclear charge that binds

[16] **Verify these equations.**
[17] **Verify this.**

them both is $Z_{\text{eff}} = 27/16 < 2$. More complicated trial functions have been used to calculate better approximations to the ground-state energy, and there is quite satisfactory agreement between theoretical and experimental results. (In Section 9.1 we shall return to other aspects of the helium atom, such as auto-ionization and inelastic scattering.)

6.3.3 Scattering

The variational method may also be applied to scattering problems. Consider the calculation of phase shifts for a central potential $V(r)$. Any regular solution of the radial Schrödinger equation for given $l \geq 0$ and $k > 0$ can be so normalized that its asymptotic form for large r is given by

$$\varphi(r) \sim a_\varphi \cos(kr - \frac{\pi}{2}l) + k^{-1} \sin(kr - \frac{\pi}{2}l); \tag{6.59}$$

the number a_φ then determines the phase shift by [see (4.41)]

$$a_\varphi = k^{-1} \tan \delta_l. \tag{6.60}$$

In other words, if a regular solution φ of

$$\mathbf{H}_l \varphi = k^2 \varphi, \tag{6.61}$$

with

$$\mathbf{H}_l \stackrel{\text{def}}{=} -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + \frac{2M}{\hbar^2} V(r), \tag{6.62}$$

is normalized so that

$$\lim_{r\to\infty} [k\varphi(r) \sin(kr - \frac{\pi}{2}l) + \varphi'(r) \cos(kr - \frac{\pi}{2}l)] = 1, \tag{6.63}$$

and a_φ is defined by

$$a_\varphi \stackrel{\text{def}}{=} \lim_{r\to\infty} [\varphi(r) \cos(kr - \frac{\pi}{2}l) - k^{-1}\varphi'(r) \sin(kr - \frac{\pi}{2}l)], \tag{6.64}$$

then (6.60) determines the phase shift. If such a solution is inserted in the functional defined by

$$\mathfrak{f}(\varphi) \stackrel{\text{def}}{=} \lim_{r\to\infty} [\varphi(r) \cos(kr - \frac{\pi}{2}l) - k^{-1}\varphi'(r) \sin(kr - \frac{\pi}{2}l)] - \int_0^\infty dr\, \varphi[\mathbf{H}_l - k^2]\varphi, \tag{6.65}$$

the integral vanishes, with the result that $\mathfrak{f}(\varphi) = a_\varphi$. Moreover, if we require that the function φ be regular at the origin and have the asymptotic form (6.59) with an unknown a_φ, then you will easily verify[18] by two integrations by parts, that its first variation near a regular solution of (6.61) vanishes:

$$\delta\mathfrak{f}(\varphi) = 0.$$

[18] **Do it.**

This means that the phase shift δ_l can be calculated by inserting an exact regular solution φ of (6.61) with the asymptotic normalization (6.63) in the functional, which then has the value

$$\mathfrak{f}(\varphi) = a_\varphi = k^{-1} \tan \delta_l .$$

On the other hand, if φ is an arbitrary trial function (not necessarily a solution of (6.61)) that vanishes at the origin, has the asymptotic normalization (6.63), and is a good approximation to a regular solution of (6.61), then the calculated value of $\mathfrak{f}(\varphi)$ will be a very good approximation to the value of a_φ that determines the exact phase shift δ_l by (6.60). For numerical calculations of the phase shift from a numerical solution of (6.61) it is therefore much preferable and likely to yield more accurate results to employ (6.65) rather than simply (6.64).

6.4 Problems and Exercises

1. Use the *WKB* approximation to show that an attractive potential in three dimensions that falls off like r^{-n} for large r has an infinite number of bound states if $n \leq p$. What is the value of p obtained in this way?

2. A particle of zero angular momentum is subject to the potential

$$V(r) = \begin{cases} -V_0 & \text{if} \quad r < r_0, \\ \alpha/r & \text{if} \quad r > r_0, \end{cases}$$

where $\alpha > 0$ and $V_0 > 0$. Use the *WKB* approximation to calculate the transmission probability if the particle is inside the well and has the energy E. (The reult is known as Gamow's formula.) From this, estimate the probability per unit time for the particle to get out, and hence the lifetime of the state.

3. Show that the *WKB* approximation gives the correct energy eigenvalues for all states of the harmonic oscillator.

4. Consider the energy levels of a double potential well with a high barrier in the center, as shown in Figure 3.6. As we saw in Chapter 3, the levels are almost degenerate. By regarding the splitting of the levels as giving rise to "beats" that describe the passing back and forth of a particle from one well to the other, calculate the approximate energy split by means of the transmission coefficient in the WKB approximation.

5. Calculate the transmissioin coefficient for the penetration of the barrier potential $V = A - \frac{1}{2}\alpha^2 x^2$ at an energy $E < A$ in the WKB approximation.

6. Show that for any two vectors $\vec{a}$ and $\vec{b}$,

$$3 \int d\Omega \, \vec{a} \cdot \hat{r} \, \vec{b} \cdot \hat{r} = \vec{a} \cdot \vec{b}.$$

7. Do the integral in (6.55), using the expansion (D.9).

8. Consider a particle of mass M free to move along the x-axis, subject to the potential

$$V = \frac{4\sqrt{2M}}{\hbar} A|x|, \qquad A > 0.$$

(a) Use the variational method to find a reasonable approximation to the ground-state energy. Is the exact ground-state energy likely to be greater than your result, or smaller?

(b) By a similar method, find an approximation to the lowest energy of a state of odd parity.

9. A rotationally invariant Hamiltonian of a spinless particle is subject to a perturbation $\mathbf{H}'$ that is the y-component of a vector. In first-order perturbation theory, calculate the energy shifts of the s and p bound states, assuming that there is no accidental degeneracy.

10. Consider a rigid spherically symmetric system whose center is fixed and whose moment of inertia about that center is I. Write down its kinetic energy and add the perturbation $\mathbf{H}' = \alpha(\mathbf{L}_x\mathbf{L}_y + \mathbf{L}_y\mathbf{L}_x)$. Calculate the energies of the first four levels in first-order perturbation theory. Is this answer an approximation or is it exact?

11. Consider a hydrogen atom inserted in a constant electric field. What are the good quantum numbers for the atom in that situation? Use ordinary first-order perturbation theory for a weak electric field to calculate the first-order energy shifts of all the levels of the atom. Is this answer reliable?

12. Calculate the energy eigenvalues for the anharmonic oscillator, whose Hamiltonian is given by

$$\mathbf{H} = \mathbf{p}^2/2M + \frac{1}{2}M\omega^2\mathbf{x}^2 + \alpha\mathbf{x}^4$$

for small α to first order in α.

13. Calculate the second-order energy shifts in a one-dimensional harmonic oscillator subject to the perturbation $\mathbf{H}' = \alpha\mathbf{x}^3$. Does this system actually have bound states? Comment and explain.

14. Use a trial function of the form $\phi(\vec{r}) = e^{-\alpha r}$ to approximate the ground state of a particle of mass M and charge e in the Coulomb potential $-e/r$. Is the result a good approximation?

15. Two one-dimensional harmonic oscillators of the same mass and classical frequencies are weakly coupled; the Hamiltonian is given by

$$\mathbf{H} = \frac{\mathbf{p}_1^2 + \mathbf{p}_2^2}{2M} + \frac{1}{2}M\omega^2(\mathbf{x}_1^2 + \mathbf{x}_2^2) + \alpha\mathbf{x}_1\mathbf{x}_2.$$

Calculate the three lowest energy eigenvalues to first order in α.

16. The proton is actually not a point charge but has a charge distribution whose r.m.s. radius is $\sim 0.8 \times 10^{-13}$cm. Show that electronic s-states are much more sensitive to the finite nuclear size than are states of higher angular momentum, and by how much. (What is the physical reason for this?) Estimate the energy shifts for the 1s and 2s levels in

hydrogen by assuming that the proton's charge is uniformly spread over a sphere, and compare your result with the fine-structure splitting. How big are these effects for a muon bound to a proton? What about a bound system consisting of a proton and an antiproton?

17. A particle of mass M is bound by a potential that has the form $V(r) = -Ar^{-2}\sin(\pi r/R)$ for $r < R$, and $V = 0$ for $r > R$, with $4MA/\hbar^2 = 10^3$. Use the variational method with a trial function $e^{-\alpha r}$ to get a good limit on the lowest energy eigenvalue.

18. Using the variational method and the trial function $\phi(x) = e^{-\alpha^2 x^2}$ find an approximation to the ground-state energy of a particle of mass M in one dimension, subject to the potential $V(x) = \frac{1}{2}M\omega^2 x^2$. Is the result a good approximation?

19. Consider a rotating system subject to the Hamiltonian

$$\mathbf{H}_0 = \vec{\mathbf{L}}^2/2I + A\mathbf{L}_z^2$$

with $AI < 2$ and a small perturbation $\mathbf{H}' = \alpha\mathbf{L}_x^2$. Find the energies of the four lowest levels in first-order perturbation theory.

7
Static Magnetic Fields

7.1 The Larmor Precession

A magnetic field $\vec{B}(\vec{r})$ can be generated by the vector potential

$$\vec{A}(\vec{r}) = \frac{1}{4\pi}\nabla \times \int d^3r' \, \frac{\vec{B}(\vec{r}')}{|\vec{r} - \vec{r}'|}, \tag{7.1}$$

so that $\vec{B} = \nabla \times \vec{A}$. As we saw in Section 2.4, this vector potential manifests itself in the Hamiltonian of a particle of charge e in the form

$$\mathbf{H} = \frac{1}{2M}\left(\vec{\mathbf{p}} - \frac{e}{c}\vec{A}\right)^2 + V. \tag{7.2}$$

If the field $\vec{B}$ is *uniform*, the vector potential can be taken to be

$$\vec{A} = \frac{1}{2}\vec{B} \times \vec{r};$$

therefore in this case

$$\begin{aligned}
(\vec{\mathbf{p}} - \frac{e}{c}\vec{A})^2 &= \vec{\mathbf{p}}^2 + \frac{e^2}{c^2}\vec{A}^2 - \frac{e}{c}(\vec{\mathbf{p}} \cdot \vec{A} + \vec{A} \cdot \vec{\mathbf{p}}) \\
&= \vec{\mathbf{p}}^2 + \frac{e^2}{c^2}\vec{A}^2 - \frac{e}{2c}(\vec{\mathbf{p}} \cdot \vec{B} \times \vec{\mathbf{r}} + \vec{B} \times \vec{\mathbf{r}} \cdot \vec{\mathbf{p}}) \\
&= \vec{\mathbf{p}}^2 + \frac{e^2}{c^2}\vec{A}^2 - 2M\mu_0 \vec{B} \cdot \vec{\mathcal{L}},
\end{aligned}$$

where $\vec{\mathcal{L}} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}/\hbar$ is the orbital angular momentum operator in units of $\hbar$, and

$$\mu_0 \stackrel{\text{def}}{=} \frac{e\hbar}{2Mc}.$$

Thus the Hamiltonian of a particle of charge e in the presence of a static, uniform magnetic field and no other forces becomes

$$\mathbf{H} = \frac{\vec{\mathbf{p}}^2}{2M} + \frac{e^2}{2Mc^2}\vec{A}^2 - \mu_0 \vec{\mathcal{L}} \cdot \vec{B}, \tag{7.3}$$

or, for a sufficiently weak magnetic field (so that the Lorentz force on the moving particle is small compared to the centrifugal force, $ev\mathcal{B}/c \ll Mv^2/r$), more simply

$$\mathbf{H} = \frac{\vec{\mathbf{p}}^2}{2M} - \mu_0 \vec{\mathcal{L}} \cdot \vec{B}, \tag{7.4}$$

the last term being the energy of the effective magnetic moment of the moving charged particle in a magnetic field. As mentioned in Chapter 5, the factor by which this magnetic moment differs from the angular momentum is called the *gyromagnetic ratio*, and its magnitude μ_0 for an electron is the *Bohr magneton*, whose numerical value is

$$\mu_0 = e\hbar/2Mc = 0.927 \times 10^{-20} \text{ergs/gauss}.$$

Equation (2.19) allows us to write down the equation of motion for the angular momentum in the form

$$i\hbar \frac{d}{dt}\vec{\mathcal{L}} = [\vec{\mathcal{L}}, \mathbf{H}] = -\mu_0 [\vec{\mathcal{L}}, \vec{\mathcal{L}} \cdot \vec{B}]$$

because the first term in (7.4) is invariant under rotations and hence commutes with $\vec{\mathcal{L}}$. The commutation relations (2.32) therefore yield the equation of motion

$$\frac{d\vec{\mathcal{L}}}{dt} = \frac{\mu_0}{\hbar}\vec{\mathcal{L}} \times \vec{B}, \tag{7.5}$$

just as classically. The result is a precession of the angular momentum vector about the magnetic field[1] with the *Larmor frequency*

$$\omega_L \stackrel{\text{def}}{=} \frac{e\mathcal{B}}{2Mc} \tag{7.6}$$

if $\mathcal{B}$ is the magnitude of the magnetic field.

Let us find the energy eigenvalues of the Hamiltonian (7.2). In the configuration representation, the time-independent Schrödinger equation has the form

$$\left[-\frac{\hbar^2}{2M}(\nabla - i\frac{e}{\hbar c}\vec{A})^2 + V\right]\psi = E\psi. \tag{7.7}$$

[1] **Show this.**

Consider the simplest case, a uniform magnetic field $\mathcal{B}\widehat{e}_z$ in the z-direction, and no other potential. Such a magnetic field can be generated by the vector potential

$$A_x = -\frac{1}{2}\mathcal{B}y, \qquad A_y = \frac{1}{2}\mathcal{B}x, \qquad A_z = 0.$$

In cylindrical coordinates, the Schrödinger equation then reads[2]

$$\left[-\frac{\partial^2}{\partial\rho^2} - \frac{1}{\rho}\frac{\partial}{\partial\rho} - \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial z^2} + \frac{\rho^2}{4R_c^4} + i\frac{1}{R_c^2}\frac{\partial}{\partial\phi}\right]\psi = \frac{2ME}{\hbar^2}\psi,$$

where

$$R_c \stackrel{\text{def}}{=} \sqrt{\frac{\hbar c}{e\mathcal{B}}}$$

is the *cyclotron radius*. The assumption that ψ has the form $\psi = f(z)F(\rho,\phi)$, as in the usual separation of variables, leads to $f(z) = e^{iKz}$. Subsequently expanding the ϕ-dependence of $F(\rho,\phi)$ in a Fourier series

$$F(\rho,\phi) = \sum_{m=-\infty}^{\infty} e^{im\phi}R_m(\rho),$$

yields the following differential equation for $R_m(\rho)$:

$$\left[-\frac{\partial^2}{\partial\rho^2} - \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{m^2}{\rho^2} + \frac{\rho^2}{4R_c^4}\right]R_m(\rho) = \left(\frac{2ME}{\hbar^2} - K^2 + \frac{m}{R_c^2}\right)R_m(\rho). \tag{7.8}$$

For comparison, consider the Schrödinger equation of an isotropic harmonic oscillator of mass M in two dimensions:

$$\left[-\frac{\hbar^2}{2M}\nabla^2 + \frac{1}{2}M\omega_L^2(x^2+y^2)\right]\psi^{\text{osc}} = E\psi^{\text{osc}},$$

with ω_L as defined by (7.6). In polar coordinates, after the same separation of variables as before, this becomes

$$\left[-\frac{\partial^2}{\partial\rho^2} - \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{m^2}{\rho^2} + \frac{M^2}{\hbar^2}\omega_L^2\rho^2\right]R_m^{\text{osc}} = \frac{2ME}{\hbar^2}R_m^{\text{osc}}. \tag{7.9}$$

The spectrum of the two-dimensional harmonic oscillator is easily obtained from that of the one-dimensional one, because in Cartesian coordinates the eigenfunctions are simply products of those in one dimension; that is, they are of the form

$$\psi_{n_1,n_2}(x,y) = H_{n_1}(\sqrt{M\omega_L/\hbar}x)H_{n_2}(\sqrt{M\omega_L/\hbar}y)\exp\left[-\frac{M\omega_L}{2\hbar}(x^2+y^2)\right],$$

[2] **Show this.**

where the H_n are Hermite polynomials, and the energy eigenvalues are the sums of those for the two one-dimensional oscillators

$$E_n^{\rm osc} = \hbar\omega_L(n_1 + \frac{1}{2} + n_2 + \frac{1}{2}), \qquad n_1, n_2 = 0, 1, 2, \ldots .$$

Since H_n is a polynomial of order n, the highest powers of $\cos\phi$ and $\sin\phi$ in $\psi_{n_1,n_2}(x, y)$ are $\cos^{n_1}\phi$ and $\sin^{n_2}\phi$; this implies that (a) for a given value of m, as in (7.8), we must have $n_1 = n_2$, so that the polynomials in $\cos\phi$ and $\sin\phi$ can add up to $e^{im\phi} = (\cos\phi + i\sin\phi)^m$; and (b) as a function of ϕ, for a given value of $n = n_1 = n_2$ there are no Fourier components of order higher than $|m| = 2n$; in other words, for a given value of m, we must have $2n \geq |m|$. It follows that the eigenvalues of (7.9) are $2ME_n^{\rm osc}/\hbar^2$, with

$$E_n^{\rm osc} = \hbar\omega_L\, 2(n + \frac{1}{2}), \qquad n = |m|, |m| + 1, |m| + 2, \ldots .$$

However, since $M\omega_L/\hbar = 1/(2R_c^2)$, the operator in (7.8) is identical to that in (7.9); therefore, its eigenvalues are

$$\frac{2M}{\hbar^2}E_n - K^2 + \frac{m}{R_c^2} = \frac{2M}{\hbar^2}\hbar\omega_L(2n + 1) = R_c^{-2}(2n + 1),$$

implying that the energies of the charged particle in the magnetic field are

$$E_n = \hbar\omega_c(n + \frac{1}{2}) + \frac{\hbar^2K^2}{2M}, \qquad n = 1, 2, \ldots , \tag{7.10}$$

where $\omega_c \stackrel{\rm def}{=} e\mathcal{B}/Mc = 2\omega_L$ is the *cyclotron frequency*, and for each given value of n, the allowed values of m are $|m| < 2n$. The term $\hbar^2K^2/2M$ is, of course, nothing but the part of the kinetic energy owing to the motion in the direction of the magnetic field, which is unaffected by that field, just as classicially. So, the circular projection of the classical motion of a charged particle on the plane perpendicular to the magnetic field is quantized, with an energy spectrum equal to that of a simple harmonic oscillator of the cyclotron frequency, while the component of the motion in the direction of the magnetic field retains its continuous spectrum.

It is interesting to note that the problem can also be solved by adopting a different gauge for the vector potential, in which

$$A_x = -\mathcal{B}y, \qquad A_y = A_z = 0,$$

so that (7.7) becomes

$$\left[-\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} - \frac{2iy}{R_c^2}\frac{\partial}{\partial x} + \frac{y^2}{R_c^4}\right]\psi = \frac{2ME}{\hbar^2}\psi. \tag{7.11}$$

The *ansatz*

$$\psi \stackrel{\rm def}{=} e^{iKz+ikx}f(y),$$

(which is justified by the fact that (7.11) does not contain the variables x and z, so that the Hamiltonian commutes with the x and z components of the momentum, as a result of which these momentum components are conserved) leads to the differential equation[3] for f

$$-f'' + \frac{(y + kR_c^2)^2}{R_c^4} f = \left(\frac{2ME}{\hbar^2} - K^2 \right) f, \tag{7.12}$$

which is identical to the equation of a simple one-dimensional harmonic oscillator with the cylcotron frequency ω_c, but with its center displaced to $y_0 \stackrel{\text{def}}{=} -kR_c^2$. The resulting energy eigenvalues agree with (7.10), and the (non-normalized) eigenfunctions are given by

$$\psi_n = e^{iKz+ikx} H_n(\frac{y - y_0}{R_c}) e^{-(y-y_0)^2/2R_c^2}, \tag{7.13}$$

in terms of Hermite polynomials (see Appendix D.3). The (real) value of k being arbitrary, the center y_0 of the motion is arbitrary as well, and there is an infinite degeneracy.

7.2 The Aharanov-Bohm Effect

As you know, classically it is only the magnetic field that can have any physical effects, not the vector potential, which is subject to gauge transformations. However, in quantum mechanics, this is not so, as the following will demonstrate.

Suppose there is a magnetic field $\vec{B}$ that is confined to a restricted region Ω in space (such as the interior of a solenoid), and the corresponding vector potential is (to within a gauge transformation) given by $\vec{A}$, so that $\vec{B} = \nabla \times \vec{A}$. In the region $\overline{\Omega}$ outside Ω, this vector potential does not necessarily vanish, but since there $\nabla \times \vec{A} = 0$, it follows that the line integral $\int^{\vec{r}} d\vec{r}' \cdot \vec{A}$ is independent of the path of integration,[4] and $\vec{A}$ is a well-defined function of $\vec{r}$ such that $\vec{A} = \nabla \int^{\vec{r}} d\vec{r}' \cdot \vec{A}$. Consequently, you easily verify that

$$\nabla - i\frac{e}{c}\vec{A} = \exp\left[i\frac{e}{c} \int^{\vec{r}} d\vec{r}' \cdot \vec{A} \right] \nabla \exp\left[-i\frac{e}{c} \int^{\vec{r}} d\vec{r}' \cdot \vec{A} \right],$$

which implies that if the boundary conditions associated with (7.7) force the wave function ψ to vanish in the region where $\vec{B} \neq 0$, that is, the particle is kept away from the magnetic field, the new wave function

$$\psi'(\vec{r}) \stackrel{\text{def}}{=} \psi(\vec{r}) \exp\left[-i\frac{e}{c} \int^{\vec{r}} d\vec{r}' \cdot \vec{A} \right]$$

[3] **Show this.**
[4] **Prove this.**

satisfies the Schrödinger equation

$$\left[-\frac{\hbar^2}{2M}\nabla^2 + V\right]\psi' = E\psi';$$

the vector potential has been *tranformed away,* making itself felt, however, in the phase of the wave function.

Suppose, then, that a coherent beam of charged particles is split, with one half passing to the left and the other half passing to the right of an infinitely long solenoid with a magnetic field inside, with no penetration of the particles possible. As the split beam is recombined, their two halves have different phases, because the line integrals for the two wave functions are different, and they will show interference effects from the phase difference,

$$\begin{aligned} & \left|\exp\left[-i\frac{e}{c}\int_{\text{path I}}^{\vec{r}} d\vec{r}\,'\cdot\vec{A}\right] - \exp\left[-i\frac{e}{c}\int_{\text{path II}}^{\vec{r}} d\vec{r}\,'\cdot\vec{A}\right]\right|^2 \\ = & \left|\exp\left[-i\frac{e}{c}\left(\int_{\text{path I}}^{\vec{r}} - \int_{\text{path II}}^{\vec{r}}\right) d\vec{r}\,'\cdot\vec{A}\right] - 1\right|^2 \\ = & \left|\exp\left[-i\frac{e}{c}\oint d\vec{r}\,'\cdot\vec{A}\right] - 1\right|^2 = 4\sin^2\left(\frac{e}{2c}\Phi\right), \end{aligned}$$

where Φ is the total magnetic flux in the interior of the solenoid. The resulting observable interference fringes can be experimentally manipulated by changing the current in the solenoid and hence the magnetic flux Φ. This shows that the vector potential has detectable effects in a region of space in which the magnetic field vanishes. Even though the charged particles are totally excluded from the region of the magnetic field, the fact that, according to (7.1), there has to be a nonvanishing vector potential exterior to the field leads to consequences which have in fact been experimentally observed. This is called the *Aharanov-Bohm effect.*

7.3 Charged Particles with Spin

A charged spin-1/2 particle of mass M, such as an electron, automatically has a magnetic moment

$$\vec{\mu} = \mu_0\vec{\sigma},$$

where μ_0 is the Bohr magneton. The magnitude of this magnetic moment, as we shall see in Chapter 10, requires the relativistic Dirac equation for an explanation; in the non-relativistic Schrödinger equation, its inclusion by Pauli was ad hoc. Moreover, that the magnitude of the magnetic moment of the electron is exactly one Bohr magneton is an approximation, albeit

an excellent one, corrections to which owe their existence to the effects of quantum electrodynamics.

The nonrelativistic Hamiltonian of an electron in the presence of a static magnetic field thus includes the "Pauli term" and is of the form

$$\mathbf{H} = \frac{1}{2M}(\vec{\mathbf{p}} - \frac{e}{c}\vec{A})^2 + V - \mu_0 \vec{\sigma} \cdot \vec{B}, \tag{7.14}$$

which for a uniform magnetic field $\vec{B}$ becomes

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_1, \qquad \mathbf{H}_0 \stackrel{\text{def}}{=} \frac{\vec{\mathbf{p}}^2}{2M} + V, \tag{7.15}$$

$$\mathbf{H}_1 \stackrel{\text{def}}{=} \frac{e^2}{2Mc^2}\vec{A}^2 - \mu_0(\vec{\sigma} + \vec{\mathcal{L}}) \cdot \vec{B} = \frac{e^2}{2Mc^2}\vec{A}^2 - \mu_0(2\vec{\mathsf{S}} + \vec{\mathcal{L}}) \cdot \vec{B}, \tag{7.16}$$

where $\vec{\mathsf{S}}$ is the spin angular momentum (in units of $\hbar$) of the particle.

Just as for the orbital angular momentum, we can now derive the equation of motion for the spin angular momentum, namely,

$$\frac{d\vec{\mathsf{S}}}{dt} = \frac{e}{Mc}\vec{\mathsf{S}} \times \vec{B}, \tag{7.17}$$

which leads to a precession about the magnetic field with the *cyclotron frequency* $\omega_c = e\mathcal{B}/Mc$, i.e., twice that of the orbital precession ω_L.

7.3.1 Spin-orbit coupling and the fine structure

The first of the consequences of effective internal magnetic fields in atoms is a relativistic phenomenon that arises from the fact that a magnetic moment moving in the static electric field of the nucleus sees an effective magnetic field

$$\vec{B}' = -\frac{1}{2}\frac{\vec{v}}{c} \times \vec{E}, \tag{7.18}$$

which exerts a torque on it.

In order to prove (7.18), consider the torque on the magnetic moment of an electron moving in the static electric field produced by a point charge. To begin with, the effective electric dipole moment $\vec{\mathfrak{p}}$ of a magnetic dipole $\vec{\mu}$ moving with the velocity $\vec{v}$ is, to first order in v/c,

$$\vec{\mathfrak{p}} = \frac{\vec{v}}{c} \times \vec{\mu},$$

so that the torque on it exerted by an electric field $\vec{E}$ is given by

$$\vec{\mathfrak{p}} \times \vec{E} = (\frac{\vec{v}}{c} \times \vec{\mu}) \times \vec{E}.$$

Hence the equation of motion for the intrinsic angular momentum $\vec{\mathbf{S}}$ of an electron, which has a magnetic moment $\vec{\mu} = \frac{e}{Mc}\vec{\mathbf{S}}$, and whose motion is governed by the Newtonian

equation $M\dot{\vec{v}} = e\vec{E}$, is given by

$$\begin{aligned}
\frac{d\vec{\mathbf{S}}}{dt} &= (\frac{\vec{v}}{c} \times \vec{\mu}) \times \vec{E} \\
&= \frac{1}{2}(\frac{\vec{v}}{c} \times \vec{\mu}) \times \vec{E} - \frac{1}{2}(\vec{\mu} \times \frac{\vec{v}}{c}) \times \vec{E} \\
&= -\frac{1}{2}(\vec{E} \times \frac{\vec{v}}{c}) \times \vec{\mu} - \frac{1}{2}[(\vec{\mu} \times \vec{E}) \times \frac{\vec{v}}{c} + (\vec{\mu} \times \frac{\vec{v}}{c}) \times \vec{E}] \\
&= -\frac{1}{2}(\vec{E} \times \frac{\vec{v}}{c}) \times \vec{\mu} + \frac{M}{2ce}\frac{d}{dt}[\frac{\vec{v}}{c} \times (\vec{\mu} \times \vec{v})] - \frac{M}{2ce}\frac{\vec{v}}{c} \times (\dot{\vec{\mu}} \times \vec{v}) \\
&= -\frac{1}{2}(\vec{E} \times \frac{\vec{v}}{c}) \times \vec{\mu} + \frac{1}{2}\frac{d}{dt}[\frac{\vec{v}}{c} \times (\vec{\mathbf{S}} \times \frac{\vec{v}}{c})] - \frac{1}{2}\frac{\vec{v}}{c} \times (\dot{\vec{\mathbf{S}}} \times \frac{\vec{v}}{c}),
\end{aligned}$$

or

$$\frac{d}{dt}[\vec{\mathbf{S}} - \frac{1}{2}(\frac{\vec{v}}{c} \times \vec{\mathbf{S}}) \times \frac{\vec{v}}{c}] = \frac{1}{2}\vec{\mu} \times (\vec{E} \times \frac{\vec{v}}{c}) - \frac{1}{2}\frac{\vec{v}}{c} \times (\dot{\vec{\mathbf{S}}} \times \frac{\vec{v}}{c}),$$

which implies that to first order in v/c

$$\frac{d\vec{\mathbf{S}}}{dt} = \frac{1}{2}\vec{\mu} \times (\vec{E} \times \frac{\vec{v}}{c}) = \vec{\mu} \times \vec{B}',$$

where the effective magnetic field is given by (7.18).

The unexpected factor of $\frac{1}{2}$, which owes its existence to the precession of the electron's spin, is known as the *Thomas factor* after L.H. Thomas. (Remember that the magnetic field seen by an observer moving with the velocity $\vec{v}$ with respect to a frame in which there is only an electric field $\vec{E}$ is given by $\vec{B} = -\frac{\vec{v}}{c} \times \vec{E}$.) As a result, the interaction energy of the electron's magnetic moment with the Coulomb field in an atom is

$$\begin{aligned}
\mathbf{H}'_{\mathrm{LS}} &= -\vec{\mu} \cdot \vec{B}' = -\frac{1}{2}\vec{\mu} \cdot (\vec{E} \times \frac{\vec{v}}{c}) \\
&= -\frac{\mu_0}{2c}\vec{\sigma} \cdot \frac{-Ze\vec{\mathbf{r}}}{\mathbf{r}^3} \times \frac{\vec{\mathbf{p}}}{M} = \frac{\mu_0 Ze}{2Mc\mathbf{r}^3}\vec{\sigma} \cdot \vec{\mathbf{r}} \times \vec{\mathbf{p}} \\
&= Z\mu_0^2 \frac{\vec{\sigma} \cdot \vec{\mathcal{L}}}{\mathbf{r}^3}, \qquad (7.19)
\end{aligned}$$

in which $\vec{\mathcal{L}}$ is the orbital angular momentum in units of $\hbar$. Since this interaction energy depends both on the spin and the orbital angular momentum of the electron, it is called *spin-orbit coupling.* Thus the Hamiltonian of an electron in a hydrogenic atom is given by

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}'_{\mathrm{LS}}, \quad \mathbf{H}_0 = \frac{\vec{\mathbf{p}}^2}{2M} - \frac{Ze^2}{\mathbf{r}}. \qquad (7.20)$$

Let us, then, calculate the first-order energy shift due to the perturbation (7.19). Since $\vec{\sigma} \cdot \vec{\mathcal{L}}$ is a scalar operator and therefore invariant under active rotations of the system, it commutes with $\vec{\mathbf{J}}$, and the perturbation is diagonal in the eigenvalues of $\vec{\mathbf{L}}^2, \vec{\mathbf{S}}^2$, and $\mathbf{J}_z$, as is $\mathbf{H}_0$; therefore, degenerate perturbation theory works, and the first-order energy shift is given by

$$\Delta E = Z\mu_0^2 \langle \vec{\sigma} \cdot \vec{\mathcal{L}} \mathbf{r}^{-3} \rangle, \qquad (7.21)$$

where $\vec{\sigma}\cdot\vec{\mathcal{L}}$ can be replaced by its eigenvalue $(\vec{\sigma}\cdot\vec{\mathcal{L}})'$ in the state under consideration. Because $\vec{\mathbf{J}}^2 = (\vec{\mathbf{L}}+\vec{\mathbf{S}})^2$, we find that $j(j+1) = l(l+1) + \frac{3}{4} + (\vec{\mathcal{L}}\cdot\vec{\sigma})'$, and therefore,

$$(\vec{\mathcal{L}}\cdot\vec{\sigma})' = \left\{ \begin{array}{cc} l & \text{if } j = l+\frac{1}{2}, \\ -(l+1) & \text{if } j = l-\frac{1}{2} \end{array} \right\} = \frac{2l(l+1)(j-l)}{j+\frac{1}{2}}. \tag{7.22}$$

We still need the expectation value $\langle\vec{\sigma}\cdot\vec{\mathcal{L}}\mathbf{r}^{-3}\rangle$ in hydrogenic states.

The radial hamiltonian for a hydrogenic atom is given by

$$\mathbf{H}_0 = \frac{1}{2M}\left[\mathbf{p}_r^2 + \frac{\hbar^2 l(l+1)}{\mathbf{r}^2}\right] - \frac{Ze^2}{\mathbf{r}},$$

so that

$$\frac{i}{\hbar}[\mathbf{p}_r, \mathbf{H}_0] = -\frac{\hbar^2 l(l+1)}{M\mathbf{r}^3} + \frac{Ze^2}{\mathbf{r}^2},$$

which allows us to evaluate $\langle l(l+1)r^{-3}\rangle$; because $\langle[\mathbf{p}_r, \mathbf{H}_0]\rangle = 0$ it follows from (7.22) that[5]

$$\langle\vec{\sigma}\cdot\vec{\mathcal{L}}\mathbf{r}^{-3}\rangle = \frac{2(j-l)}{j+\frac{1}{2}}\langle l(l+1)\mathbf{r}^{-3}\rangle = \frac{2Z(j-l)}{a_0(j+\frac{1}{2})}\langle\mathbf{r}^{-2}\rangle, \tag{7.23}$$

where a_0 is the Bohr radius, defined in (6.47). Using the result (6.48), we therefore obtain

$$\langle\vec{\sigma}\cdot\vec{\mathcal{L}}\mathbf{r}^{-3}\rangle = \frac{Z^3}{n^3 a_0^3}\frac{2(j-l)}{(l+\frac{1}{2})(j+\frac{1}{2})}. \tag{7.24}$$

Inserting Eq. (7.24) into Eq. (7.21) yields the following spin-orbit shift:

$$\Delta E_n^{\mathrm{LS}} = \frac{Z^4\mu_0^2}{a_0^3 n^3}\frac{2(j-l)}{(l+\frac{1}{2})(j+\frac{1}{2})}. \tag{7.25}$$

Therefore, the *fine structure* of the spectrum, which is the combination of the spin-orbit shift (7.25) and the relativistic shift of (6.50), in the n^{th} energy level of a hydrogenic atom, is given by

$$\Delta E_n^{\mathrm{LS\,rel}} = \frac{Z^4\alpha^2}{n^3}\left[\frac{3}{4n} - \frac{1}{j+\frac{1}{2}}\right]\mathrm{Ry}, \tag{7.26}$$

in which

$$\mathrm{Ry} \stackrel{\mathrm{def}}{=} \frac{e^2}{2a_0} = \frac{1}{2}\alpha^2 Mc^2 \tag{7.27}$$

is the *Rydberg* unit, whose approximate numerical value, expressed as a frequency, is 330,000 MHz and in terms of which the unperturbed energies of a hydrogenic atom are given by

$$E_n^0 = -\frac{Z^2}{n^2}\,\mathrm{Ry}.$$

---

[5]Notice that the factor $l(l+1)$ vanishes for $l=0$, while $\langle\mathbf{r}^{-3}\rangle$ is infinite in an s-state; the combination $\langle l(l+1)\mathbf{r}^{-3}\rangle$ is nevertheless finite and well defined.

The most remarkable feature of (7.26) is that, whereas the spin-orbit and the straight relativistic energy shifts by themselves destroy the originally present l-degeneracy, their combination conspires to reinstate it!

The ground state of hydrogen, that is, the $1s$-state, is shifted by $-\alpha^2/4$ Ry, while the first excited state, the degenerate $2s, 2p$-state, is split into two levels, the lower one still degenerate, $2s_{\frac{1}{2}}$ and $2p_{\frac{1}{2}}$, and the upper one $2p_{\frac{3}{2}}$, with an energy difference of $\alpha^2/16$ Ry, which corresponds to a frequency of the radiation emitted in a transition of about $11,000$ MHz or a wavelength of approximately 3 cm. The degeneracy of the $2s_{\frac{1}{2}}$ and $2p_{\frac{1}{2}}$ states was finally found in 1951 to be broken by 1062 MHz in an experiment by Lamb and Retherford, a phenomenon, called the *Lamb shift*, that is explained by quantum electrodynamics.

7.3.2 The Zeeman effect

Consider an electron bound by a central potential $V(r)$, in the presence of a uniform external magnetic field $\vec{B}$ in the z direction that is weak enough for the term proportional to $\vec{A}^2$ in (7.16) to be negligible, and yet strong enough to dominate over the internal magnetic field (and the straight relativistic effect), so that the effective Hamiltonian is given by

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}'_{\mathcal{B}}, \qquad \mathbf{H}_0 \stackrel{\text{def}}{=} \frac{\vec{\mathbf{p}}^2}{2M} - \frac{Ze^2}{\mathbf{r}}, \qquad \mathbf{H}'_{\mathcal{B}} \stackrel{\text{def}}{=} -\mu_0 \mathcal{B}(2\mathsf{S}_z + \mathcal{L}_z) \quad (7.28)$$

while $\mathbf{H}'_{\text{LS}}$ and $\mathbf{H}'_{\text{rel}}$ can be neglected. The order of magnitude of the effective magnetic field (7.18) seen by the magnetic moment of an electron in a hydrogen atom is about 10^5 gauss.[6] Therefore, if the atom finds itself in an external magnetic field that is strong compared to this figure, the spin-orbit interaction can be considered negligible. The resulting shifts in the energy levels is called the *Paschen-Back* effect.

The Paschen-Back effect

Because the g-factors of the orbital and spin magnetic moments are not equal, the Hamiltonian (7.28) does not commute with $\mathbf{J}_z$ or with $\vec{\mathbf{J}}^2$, so that angular momentum is not conserved, but a complete set of commuting observables is given by $\vec{\mathbf{L}}^2, \vec{\mathbf{S}}^2, \mathbf{L}_z$, and $\mathbf{S}_z$.

In first-order perturbation theory, the shifted energy levels are given by

$$E_n = E_n^0 - \mu_0 \mathfrak{m}_{ls} \mathcal{B}, \text{ where } \mathfrak{m}_{ls} \stackrel{\text{def}}{=} m_l + 2m_s, \quad (7.29)$$

which shows that the m_l- and m_s-degeneracy that is there when $\mathcal{B} = 0$ is removed.[7] For the hydrogen atom, the $1s$-state is split into two levels,

[6] **Show this.**

[7] **Why can we use non-degenerate perturbation theory?**

one shifted up by $\mu_0\mathcal{B}$ and the other down by the same amount, resulting in an energy difference of $2\mu_0\mathcal{B}$ between them; the originally eight-fold degenerate state $2s, 2p$ is split into five levels, each a distance $\mu_0\mathcal{B}$ apart:

unpert.	pert.	$\mathfrak{m}_{ls}$	n	l	m_l	m_s	n	l	m_l	m_s
	———	1	1	0	0	1/2				
$1s$ ——-	$2\mu_0\mathcal{B}\updownarrow$									
	———	−1	1	0	0	−1/2				
	———	2	2	1	1	1/2				
	———	1	2	1	0	1/2	2	0	0	1/2
$2s,\ 2p$ ——	———	0	2	1	1	−1/2	2	1	−1	1/2
	———	−1	2	1	0	−1/2	2	0	0	−1/2
	———	−2	2	1	−1	−1/2				

The weak-field Zeeman effect

Next consider the Hamiltonian of an electron bound by a nucleus in the presence of an external field that is so weak that the internal effects leading to the fine structure are dominant. In that case, the starting point is an unperturbed Hamiltonian that includes $\mathbf{H}'_{\rm LS}$ as well as $\mathbf{H}'_{\rm rel}$, and $\mathbf{H}'_{\mathcal{B}}$ is the perturbation. For the unperturbed Hamiltonian l and m (where $m\hbar$ is the eigenvalue of $\vec{\mathbf{J}}_z$) are then good quantum numbers. Since $\mathbf{H}'_{\mathcal{B}}$ is diagonal in these quantum numbers, so that degenerate perturbation theory works, we simply need to calculate $\langle\mathbf{H}'_{\mathcal{B}}\rangle$ in such a state. Now, since $\mathbf{H}'_{\mathcal{B}} = -\mu_0\mathcal{B}(\mathcal{L}_z + \frac{1}{2}\sigma_z)$, we have

$$\langle\mathbf{H}'_{\mathcal{B}}\rangle = -\mu_0\mathcal{B}(m + \frac{1}{2}\langle\sigma_z\rangle),$$

and we need $\langle\sigma_z\rangle_{ljm}$, which can be calculated by means of (5.73) and (5.74), with the result,

$$\langle\sigma_z\rangle_{l,j=l\pm\frac{1}{2}} = \frac{l+\frac{1}{2}\pm m}{2l+1} - \frac{l+\frac{1}{2}\mp m}{2l+1} = \pm\frac{2m}{2l+1}.$$

Consequently, the energy shift is given by

$$\Delta E = -m\frac{l+\frac{1}{2}\pm\frac{1}{2}}{l+\frac{1}{2}}\mu_0\mathcal{B},$$

which means that the weak-field Zeemann effect is

$$\Delta E = -m\frac{j+\frac{1}{2}}{l+\frac{1}{2}}\mu_0\mathcal{B}. \tag{7.30}$$

The effective magnetic dipole moment of an atom in a weak magnetic field may be conveniently defined by

$$\mu \stackrel{\rm def}{=} -\frac{\partial E}{\partial\mathcal{B}},$$

which may be written as

$$\mu \stackrel{\text{def}}{=} mg\mu_0,$$

so that, according to (7.30), the effective g-factor of the atom is given by

$$g = \frac{j+\frac{1}{2}}{l+\frac{1}{2}}. \tag{7.31}$$

Note that this has the value $g = 2$ for $l = 0$ and the limit $g = 1$ as $l \to \infty$: when $l = 0$, the magnetic moment is entirely due to the spin of the electron, whereas in the limit as $l \to \infty$, the spin contribution is negligible and the magnetic moment is all due to the orbital motion.

The intermediate Zeeman effect

If the external magnetic field is neither weak nor strong compared to the effective internal one (but still not strong enough for the $\vec{A}^2$-term in the Hamiltonian to have to be included), then the unperturbed Hamiltonian must be taken to be $\mathbf{H}_0$ of (7.15) and the perturbation is $\mathbf{H}'_{\text{LS}} + \mathbf{H}'_{\text{rel}} + \mathbf{H}'_{\mathcal{B}}$. If j, l, m (again, $\hbar m$ is the eigenvalue of $\vec{\mathbf{J}}_z$) are used as quantum numbers, then there is j-degeneracy in the unperturbed levels, and, after writing

$$\mathbf{H}'_{\mathcal{B}} = -\mu_0\mathcal{B}(\mathcal{J}_z + \frac{1}{2}\sigma_z),$$

degenerate perturbation theory requires the diagonalization of σ_z on the degenerate subspace, which has $j = l \pm \frac{1}{2}$. The calculation of the energy shift in this case will be left as an **exercise**; the result is

$$\begin{aligned}\Delta E_n \quad = \quad & \left\{ \frac{1}{l(l+1)} \left[-m\eta_n + \frac{1}{2(2l+1)} \pm \frac{1}{2}\sqrt{1+\eta_n^2 - \frac{4m\eta_n}{(2l+1)}} \right] \right. \\ & \left. + \frac{3}{4n} - \frac{1}{l+\frac{1}{2}} \right\} \xi_n, \qquad \text{for } l \neq 0, \end{aligned} \tag{7.32}$$

and

$$\Delta E_n = 1 - 2m\mu_0\mathcal{B} + \left[\frac{3}{4n} - 2 \right] \xi_n \qquad \text{for } l = 0, \tag{7.33}$$

where

$$\xi_n \stackrel{\text{def}}{=} \frac{Z^4\alpha^4}{2n^3}Mc^2, \qquad \eta_n \stackrel{\text{def}}{=} \frac{\mu_0 l(l+1)}{\xi_n}\mathcal{B}.$$

In contrast to the weak-field case, in which the $\pm$ referred to $j = l \pm \frac{1}{2}$, the $\pm$ here cannot be so identified; both levels contain a superposition of the two. As a consequence of (7.32) and (7.33), the effective magnetic moment of the atom is given by

$$\mu = \mu_0 \left\{ m \mp \frac{1}{2} \frac{\eta_n - \frac{m}{l+\frac{1}{2}}}{\sqrt{1+\eta_n^2 - \frac{2m\eta_n}{l+\frac{1}{2}}}} \right\}, \qquad \text{for } l \neq 0, \tag{7.34}$$

which depends on $\mathcal{B}$ and is thus an *induced* magnetic moment; and

$$\mu = 2m\mu_0 \qquad \text{for } l = 0. \tag{7.35}$$

7.3.3 The hyperfine structure

In studying the structure of hydrogenic atoms we have been assuming that their nuclei are simple point charges producing an electrostatic field described by a Coulomb potential. In fact, however, atomic nuclei also have magnetic moments and therefore produce magnetic fields. In this section we shall investigate the effects of these fields.

The order of magnitude of the magnetic moment of a nucleus is roughly that of the *nuclear magneton*,

$$\mu_{\rm NM} \stackrel{\rm def}{=} \frac{e\hbar}{2M_N c} = \frac{M}{M_N}\mu_0,$$

if M is the mass of the electron and M_N that of the nucleus. For hydrogen, the nuclear magneton is therefore about 1846 times smaller than the Bohr magneton, and correspondingly, the effect on the hydrogen spectrum can be expected to be much smaller than the fine structure. The actual, experimentally measured value μ_N of the proton's magnetic moment, however, differs from $\mu_{\rm NM}$, as already mentioned in Chapter 5, and it is this value that is relevant for our calculations.

The vector potential produced by a magnetic moment $\vec{\mu}_N$ fixed at the origin is

$$\vec{A} = \frac{\vec{\mu}_N \times \vec{r}}{r^3} = -\vec{\mu}_N \times \nabla \frac{1}{r}.$$

If again the $\vec{A}^2$ term is neglected, the addition to the hydrogen Hamiltonian caused by the nuclear magnetic moment will be

$$\begin{aligned}
\mathbf{H}'_{\rm NM} &= \frac{e}{Mc}\vec{\mathbf{p}}\cdot\vec{A} - \mu_0\vec{\sigma}\cdot\vec{B} \\
&= -\frac{e}{Mc}\vec{\mathbf{p}}\cdot\frac{\vec{\mu}_N\times\vec{\mathbf{r}}}{\mathbf{r}^3} + \mu_0\vec{\sigma}\cdot\nabla\times(\vec{\mu}_N\times\nabla\mathbf{r}^{-1}) \\
&= -\frac{e}{Mc}\frac{\vec{\mu}_N\cdot\vec{\mathbf{r}}\times\vec{\mathbf{p}}}{\mathbf{r}^3} + \mu_0\vec{\sigma}\cdot\nabla\times(\vec{\mu}_N\times\nabla\mathbf{r}^{-1}) \\
&= -\mu_0\frac{\vec{\mu}_N\cdot\vec{\mathcal{L}}}{\mathbf{r}^3} + \mu_0\vec{\sigma}\times\nabla\cdot(\vec{\mu}_N\times\nabla)\mathbf{r}^{-1},
\end{aligned}$$

but since

$$\vec{\sigma}\times\nabla\cdot(\vec{\mu}_N\times\nabla)\mathbf{r}^{-1} = (\vec{\sigma}\cdot\vec{\mu}_N\nabla^2 - \vec{\sigma}\cdot\nabla\vec{\mu}_N\cdot\nabla)\mathbf{r}^{-1},$$

this becomes

$$\mathbf{H}'_{\rm NM} = -\mu_0\frac{\vec{\mu}_N\cdot\vec{\mathcal{L}}}{\mathbf{r}^3} + \mu_0\left[\frac{\vec{\sigma}\cdot\vec{\mu}_N}{\mathbf{r}^3} - 3\frac{\vec{\sigma}\cdot\vec{\mathbf{r}}\,\vec{\mu}_N\cdot\vec{\mathbf{r}}}{\mathbf{r}^5}\right] - 4\pi\mu_0\,\vec{\sigma}\cdot\vec{\mu}_N\delta^3(\vec{\mathbf{r}}). \tag{7.36}$$

The physical meaning of the three terms in this expression is that the first term describes the interaction of the nuclear magnetic moment with the orbital moment, the second, which can also be written in the form

$$\mu_0 \left[\frac{\vec{\sigma} \cdot \vec{\mu}_N}{\mathbf{r}^3} - 3 \frac{\vec{\sigma} \cdot \vec{\mathbf{r}} \, \vec{\mu}_N \cdot \vec{\mathbf{r}}}{\mathbf{r}^5} \right] = -\mu_0 \vec{\sigma} \cdot \nabla \vec{\mu}_N \cdot \nabla \mathbf{r}^{-1},$$

describes that of the nuclear magnetic moment with the intrinsic magnetic dipole moment of the electron, and the third is a contact term, an effect of the possibility of the electron's passing right through the nucleus.

We will calculate the level splitting that results from this perturbation in the 1s-state only, because the absence of fine structure in that case makes the effect most easily observable.

For $l = 0$, the wave function is spherically symmetric, and for such functions we have

$$\int d^3r \, f(r) [\vec{\sigma} \cdot \nabla \, \vec{\mu} \cdot \nabla - \frac{1}{3} \vec{\sigma} \cdot \vec{\mu} \nabla^2] r^{-1} = 0. \tag{7.37}$$

To prove (7.37) split the integral into two parts, the first extended over the exterior of a sphere about the origin of radius ϵ, the second over its interior. For the first, we have $\nabla^2 r^{-1} = 0$ and[8]

$$I_1 = \int_{r>\epsilon} d^3r \, f(r) \vec{\sigma} \cdot \nabla \, \vec{\mu} \cdot \nabla r^{-1} = \int_{r>\epsilon} d^3r \, f(r) r^{-3} (3\vec{\sigma} \cdot \hat{\mathbf{r}} \, \vec{\mu} \cdot \hat{\mathbf{r}} - \vec{\sigma} \cdot \vec{\mu}) = 0.$$

For the second part,

$$\begin{aligned} I_2 &= \int_{r<\epsilon} d^3r \, \nabla \cdot [f(r)(\vec{\mu}\vec{\sigma} \cdot \nabla r^{-1} - \frac{1}{3} \vec{\sigma} \cdot \vec{\mu} \nabla r^{-1})] \\ &- \int_{r<\epsilon} d^3r \, [\nabla f(r)] \cdot (\vec{\mu}\vec{\sigma} \cdot \nabla r^{-1} - \frac{1}{3} \vec{\sigma} \cdot \vec{\mu} \nabla r^{-1})] \\ &= f(\epsilon) \int d\Omega \, (\frac{1}{3} \vec{\sigma} \cdot \vec{\mu} - \vec{\mu} \cdot \hat{\mathbf{r}} \vec{\sigma} \cdot \hat{\mathbf{r}}) - \int_0^\epsilon dr \frac{\partial}{\partial r} f(r) \int d\Omega (\frac{1}{3} \vec{\sigma} \cdot \vec{\mu} - \vec{\mu} \cdot \hat{\mathbf{r}} \vec{\sigma} \cdot \hat{\mathbf{r}}) = 0. \end{aligned}$$

Therefore for an $l = 0$ state, $\vec{\sigma} \cdot \nabla \, \vec{\mu} \cdot \nabla r^{-1}$ may be replaced by $\frac{1}{3} \vec{\sigma} \cdot \vec{\mu} \nabla^2 r^{-1} = -(4\pi/3) \vec{\sigma} \cdot \vec{\mu} \, \delta^3(\vec{r})$, and as a result we have for s-states in the configuration representation

$$\mathbf{H}'_{\mathrm{NM}} = -\frac{8\pi}{3} \mu_0 \vec{\sigma} \cdot \vec{\mu}_N \delta^3(\vec{r}). \tag{7.38}$$

Let the nuclear spin be $\mathfrak{i}$, its spin-vector operator $\vec{\mathfrak{I}}$ (both $\mathfrak{i}$ and $\vec{\mathfrak{I}}$ in units of $\hbar$), and write

$$\vec{\mu}_N \stackrel{\text{def}}{=} \frac{|\mu_N| \vec{\mathfrak{I}}}{\mathfrak{i}},$$

---

[8] **I will ask you to prove the last step in the following equation as a homework problem.**

so that (7.38) becomes

$$\mathbf{H}'_{\mathrm{NM}} = -\frac{8\pi}{3}\frac{|\mu_0||\mu_N|}{\mathfrak{i}}\delta^3(\vec{r})\vec{\sigma}\cdot\vec{\mathfrak{l}}. \tag{7.39}$$

In order to calculate the first-order perturbation of the energy we therefore need the eigenvalues of $\vec{\sigma}\cdot\vec{\mathfrak{l}}$.

Since the atom is assumed to be in an s-state, the total angular momentum of the system (in units of $\hbar$) is given by

$$\vec{\mathfrak{F}} = \vec{\mathfrak{l}} + \vec{\mathfrak{J}} = \vec{\mathfrak{l}} + \frac{1}{2}\vec{\sigma},$$

which implies that

$$\vec{\mathfrak{F}}^2 = \vec{\mathfrak{l}}^2 + \vec{\mathfrak{l}}\cdot\vec{\sigma} + \frac{3}{4}.$$

We obtain the eigenvalues of $\vec{\mathfrak{l}}\cdot\vec{\sigma}$ from this just as we found those of $\vec{\mathcal{L}}\cdot\vec{\sigma}$ in (7.22):

$$f(f+1) = \mathfrak{i}(\mathfrak{i}+1) + \frac{3}{4} + (\vec{\sigma}\cdot\vec{\mathfrak{l}})'$$

if the eigenvalues of $\vec{\mathfrak{F}}^2$ are $f(f+1)$, from which it follows that the eigenvalues of $\vec{\mathfrak{l}}\cdot\vec{\sigma}$ are

$$(\vec{\mathfrak{l}}\cdot\vec{\sigma})' = \frac{2\mathfrak{i}(\mathfrak{i}+1)(f-\mathfrak{i})}{f+\frac{1}{2}}, \qquad f = \mathfrak{i}\pm\frac{1}{2}.$$

Therefore in an $l=0$ state of the atom, the hyperfine structure produced by first-order perturbation theory is a doublet,

$$\langle\mathbf{H}'_{\mathrm{NM}}\rangle = \frac{8\pi}{3}\frac{|\mu_0|\,|\mu_N|}{\mathfrak{i}}|\psi(0)|^2\frac{2\mathfrak{i}(\mathfrak{i}+1)(f-\mathfrak{i})}{f+\frac{1}{2}},$$

which is split by the energy difference

$$\Delta E = \frac{8\pi}{3}\left(\frac{e\hbar}{2Mc}\right)\left(\frac{e\hbar}{2M_Nc}\right)\frac{2\mathfrak{i}+1}{\mathfrak{i}}\mu_N^*\,|\psi(0)|^2, \tag{7.40}$$

where $\mu_N^* \overset{\text{def}}{=} \mu_N(2M_nc/e\hbar)$ is the magnitude of the nuclear magnetic moment in units of the nuclear magneton. In the ground state of a hydrogenic atom, (4.71) leads to

$$|\psi(0)|^2 = 4\left(\frac{Z}{a_0}\right)^3\frac{1}{4\pi},$$

which yields the split

$$\Delta E = \frac{4}{3}\mu_N^*\frac{2\mathfrak{i}+1}{\mathfrak{i}}\frac{M}{M_N}Z^3\alpha^2\,\mathrm{Ry}. \tag{7.41}$$

For the proton, $\mathfrak{i} = \frac{1}{2}$ and $\mu_N^* = 2.7925$, which gives the numerical value of the frequency corresponding to the hyperfine split for the ground state of hydrogen as 1416 MHz, compared to an experimental value of 1420 MHz. The discrepancy is accounted for by the fact that the electron has an anomalous magnetic moment, about 0.1% higher than μ_0 (for the discovery of which Polykarp Kush was awarded the Nobel prize in 1955). Since μ_N^* is experimentally measured in units of the magnetic moment of the electron, this also increases μ_N^* by about 0.1%, resulting in a total increase in the theoretically predicted value of the hyperfine split by 0.2%, in satisfactory agreement with observation. Note that the frequency of 1420 MHz corresponds to a wavelength of 21 cm for the radiation emitted in a transition from the upper to the lower of the two levels; this is the famous 21 cm signature-line of the hydrogen spectrum.

7.4 Problems and Exercises

1. Derive (7.32) and (7.33).
2. Show that (7.32) and (7.33) go over into (7.29) and (7.30), respectively, as $\eta \to \infty$ and $\eta \to 0$.
3. A spin-1/2 particle of charge e and mass M is placed in a constant magnetic field $\mathcal{H}$. If its spin initially points in the z direction and $\mathcal{H}$ is in the x direction, calculate the direction of its spin at the time t.

This page intentionally left blank

8
Time-Dependent Hamiltonians

In this chapter we shall study problems in which the Hamiltonian depends explicitly on the time, so that energy is not conserved. Three different kinds of approximations may be applicable in such a situation. The first applies when the time-dependent part of the Hamiltonian is weak compared to the time-independent part, for which the solutions of the Schrödinger equation are assumed to be known. In that case, an analogue to the perturbation theory discussed in Chapter 6 can be used. The second kind of approximation is used when the change is very *rapid*, like the sudden switching on and off of an external field. The third is applicable when the Hamiltonian, on the contrary, changes very slowly; for reasons to be discussed it is called the *adiabatic approximation.* The approximation methods appropriate to these three kinds of situations will be taken up in that order.

8.1 Perturbation Theory

The Hamiltonian in the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi(t) = \mathbf{H}(t)\Psi(t)$$

is now assumed to consist of two parts,

$$\mathbf{H}(t) = \mathbf{H}_0 + \mathbf{H}'(t),$$

and the eigenvalues and eigenstates of the time-independent $\mathbf{H}_0$ are assumed to be known. The question we would like to answer is this: if the

system is in a state $\Psi(t)$ and a measurement of $\mathbf{H}_0$ is performed, what is the probability of obtaining the result E_n (where E_n is an eigenvalue of $\mathbf{H}_0$)? The answer, of course, is that this probability is the magnitude squared of the amplitude $(\Psi_{E_n}, \Psi(t))$. Furthermore, since $\Psi(t)$ can be expanded in the form

$$\Psi(t) = \sum_n (\Psi_{E_n}, \Psi(t))\Psi_{E_n} = \sum_n (\Psi_{E_n}(t), \Psi(t))\Psi_{E_n}(t),$$

in which $\Psi_{E_n}(t) = \Psi_{E_n} \exp(-iE_n t/\hbar)$, the inner products $\langle E_n | \rangle_t \stackrel{\text{def}}{=} (\Psi_{E_n}(t), \Psi(t))$ determine everything we might want to know about $\Psi(t)$.

Use of the Schrödinger equation gives us

$$i\hbar \frac{\partial}{\partial t} \langle E_n | \rangle_t = \langle E_n | \mathbf{H} - \mathbf{H}_0 | \rangle_t,$$

which means that $\langle E_n | \rangle_t$ satisfies the differential equation

$$i\hbar \frac{\partial}{\partial t} \langle E_n | \rangle_t = \langle E_n | \mathbf{H}'(t) | \rangle_t. \tag{8.1}$$

Since

$$\langle E_n | \rangle_t = e^{\frac{i}{\hbar} E_n t} (\Psi_{E_n}, \Psi(t)) = (\Psi_{E_n}, e^{\frac{i}{\hbar} \mathbf{H}_0 t} \Psi(t)),$$

Eq. (8.1) may also be thought of as originating from the equation

$$i\hbar \frac{\partial}{\partial t} \underline{\Psi}(t) = \underline{\mathbf{H}}'(t) \underline{\Psi}(t), \tag{8.2}$$

where

$$\underline{\Psi}(t) \stackrel{\text{def}}{=} e^{\frac{i}{\hbar} \mathbf{H}_0 t} \Psi(t), \qquad \underline{\mathbf{H}}'(t) \stackrel{\text{def}}{=} e^{\frac{i}{\hbar} \mathbf{H}_0 t} \mathbf{H}'(t) e^{-\frac{i}{\hbar} \mathbf{H}_0 t}.$$

The underlined quantities are the state vectors and operators in the *interaction picture,* (also sometimes called the *Dirac picture*). As the Schrödinger equation (8.2) shows, it stands between the Schrödinger picture, in which the state vectors carry all the time dependence and the dynamical operators are time-independent (unless they have an intrinsic time dependence), and the Heisenberg picture, in which the state vectors are constant and the dynamical operators depend on the time. Here both the state vectors and the dynamical variables vary with time, but the time dependence of the state vectors is governed by the interaction alone (i.e., by the part of the Hamiltonian that is regarded as the interaction in the particular case at hand), and the time dependence of the dynamical operators is independent of the interaction:

$$\frac{d\underline{\mathbf{O}}}{dt} = \frac{\partial \underline{\mathbf{O}}}{\partial t} + \frac{i}{\hbar} [\mathbf{H}_0, \underline{\mathbf{O}}].$$

Assuming that initially, at the time $t = -\infty$, the system was in the eigenstate Ψ_{E_i} of $\mathbf{H}_0$, the appropriate initial condition to go with the Schrödinger equation is

$$\lim_{t \to -\infty} \| \Psi(t) - \Psi_{E_i}(t) \| = 0, \tag{8.3}$$

and hence, $\langle E_f|\,\rangle_t \to \delta_{fi}$ as $t \to -\infty$, assuming that both E_f and E_i are point eigenvalues. The differential equation (8.1) with the initial condition (8.3) can therefore be replaced by the integral equation

$$\begin{aligned}\langle E_f|\,\rangle_t &= \delta_{fi} + \frac{1}{i\hbar}\int_{-\infty}^{t} dt'\,\langle E_f|\mathbf{H}'(t')|\,\rangle_{t'} \\ &= \delta_{fi} + \sum_k \frac{1}{i\hbar}\int_{-\infty}^{t} dt'\,H'_{fk}(t')\langle E_k|\,\rangle_{t'} e^{i\omega_{fk}t'}, \qquad (8.4)\end{aligned}$$

where

$$H'_{fk}(t) \stackrel{\text{def}}{=} \langle E_f|\mathbf{H}'(t)|E_k\rangle, \qquad \omega_{fk} \stackrel{\text{def}}{=} (E_f - E_k)/\hbar.$$

If the interaction $\mathbf{H}'$ is weak (or for sufficiently small λ if $\mathbf{H}'$ is replaced by $\lambda\mathbf{H}'$), a first approximation to the solution of (8.4) is obtained by replacing $\langle E_k|\,\rangle_{t'}$ in the integral by its initial value δ_{fi}, so that

$$\langle E_f|\,\rangle_t \simeq \delta_{fi} + \frac{1}{i\hbar}\int_{-\infty}^{t} dt'\,H'_{fi}(t')e^{i\omega_{fi}t'}. \qquad (8.5)$$

Suppose now that the time dependence of $\mathbf{H}'$ consists simply of being turned on at the time $t = 0$ and off at $t = T$, and having the constant value $\mathbf{H}'$ from 0 to T. In that case, for $t > T$,

$$\langle E_f|\,\rangle_t \simeq \delta_{fi} - H'_{fi}\frac{e^{i\omega_{fi}T} - 1}{\hbar\omega_{fi}}$$

so that for $f \neq i$ the probability that the system will be found to have the energy E_f at the time $t \geq T$ is

$$P_{fi}(t) \simeq 4|H'_{fi}|^2\frac{\sin^2(\frac{1}{2}\omega_{fi}T)}{\hbar^2\omega_{fi}^2}. \qquad (8.6)$$

Though the right-hand side of this approximate equation is not necessarily less than 1, the approximation is, of course, good only if it is, in fact, much less than 1.

For a fixed time interval T and a fixed initial energy E_i, the function of E_f, the energy to which the system has made a transition, defined by (8.6), is plotted in Figure 8.1. It is sharply peaked at $\omega_{fi} = 0$, that is, at the energy $E_f = E_i$, where energy is conserved, with a maximum value proportional to T^2 and a width $D = 2\pi\hbar/T$. Thus, the length of time T during which the perturbation acts is related to the "energy uncertainty" ΔE by $T\Delta E \simeq h$, in accordance, in some sense, with the uncertainty principle. As we shall see, when the perturbation is slowly turned on and off, the system tends to remain in the "same state" whose energy slowly shifts; in general, however, the perturbation causes transitions from one level to another. If higher orders of perturbation theory are taken into

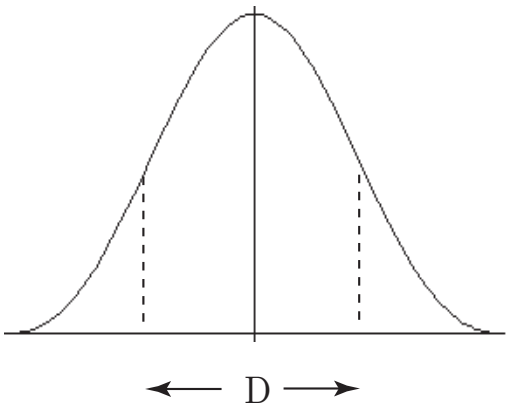

FIGURE 8.1. The transition probability as a function of the energy.

account, there are many intermediate transitions back and forth, which are called *virtual*.

In contrast to (8.6), we might have expected that the transition probability from one level to another is proportional to the time T, so that it makes sense to define a transition probability per unit time. Such a result, however, can be obtained only if the levels in the target area, i.e., near E_f in the region of the peak in Figure 8.1, are closely spaced, and if in this region H'_{fi} is essentially independent of f. In that case we would integrate and obtain for the area under the peak

$$\int_0^{2\pi/T} d\omega \, \frac{\sin^2(\omega T/2)}{\omega^2} = T \int_0^{2\pi} dx \, \frac{\sin^2 \frac{1}{2}x}{x^2} \propto T.$$

In a situation in which there are many closely spaced states near the final state with the energy E_f, all with nearly the same energy E_i as the initial state, let us assume that there are $\rho(f)$ states per unit energy interval (the *density of final states*) and that $\rho(f)$ and $|H'_{fi}|$ are slowly varying functions of f. Then the probability of a transition to any one of the states in the peak of Figure 8.1 is

$$\begin{aligned} P &= \sum_f |\langle E_f | \rangle_T|^2 = \int dE_f \, \rho(f) |\langle E_f | \rangle_T|^2 \\ &= \int_{\text{peak}} dE_f \, \rho(f) \frac{4}{\hbar^2} |H'_{fi}|^2 \, \frac{\sin^2(\frac{1}{2}\omega_{fi} T)}{\omega_{fi}^2}. \end{aligned}$$

The assumption that $\rho(f)$ and $|H'_{fi}|$ are slowly varying functions of f in the peak allows us to take both outside the integral, after which the integral is extended from $-\infty$ to ∞, since its value comes almost entirely from the peak anyway, with the result

$$P = \frac{4}{\hbar} \rho(f) |H'_{fi}|^2 \int_{-\infty}^{\infty} d\omega \, \frac{\sin^2(\frac{1}{2}\omega T)}{\omega^2} = \frac{2\pi}{\hbar} T \rho(f) |H'_{fi}|^2, \qquad (8.7)$$

which finally gives us a constant transition probability per unit time,

$$W_{fi} = \frac{2\pi}{\hbar}\rho(f)|H'_{fi}|^2, \tag{8.8}$$

a formula known as *Fermi's golden rule.*

Two criteria have to be satisfied for this result to be valid. On one hand, the assumption that both ρ and H'_{fi} be slowly varying over the width $\mathcal{D} = 2\pi/T$ (in frequency space) of the peak in Figure 8.1 means that we need

$$\frac{\partial H'\rho}{\partial\omega}\mathcal{D} \ll H'\rho,$$

or

$$T \gg 2\pi\frac{\partial|H'|\rho/\partial\omega}{|H'|\rho};$$

on the other hand, the probability (8.7) has to be small compared to 1:

$$\frac{2\pi}{\hbar}T\rho|H'|^2 \ll 1.$$

Consequently the time T must be neither too small nor too large:

$$2\pi\frac{\partial|H'|\rho/\partial\omega}{|H'|\rho} \ll T \ll \frac{\hbar/2\pi}{\rho|H'|^2},$$

and this is possible only if

$$\frac{\partial|H'|^2\rho}{\partial\omega} \ll \frac{\hbar}{(2\pi)^2}.$$

Notice the important fact, which should always remain part of your physical interpretation of the quantum-mechanical formalism, that *the modulus squared of the matrix element H'_{fi} of the perturbation $\mathbf{H}'$ is a measure of the transition probability* (and hence of the spectral intensity, in the case of transitions in which a photon is emitted) from the state Ψ_{E_i} to the state Ψ_{E_f}. This implies in particular that if $H'_{fi} = 0$, i.e., if $\mathbf{H}'$ has no non-vanishing matrix elements connecting the two levels, then $W_{fi} = 0$: the transition cannot go, to first order. To put it another way: the perturbation induces transitions only between states that it connects. This is the basis of the physical relevance of the *selection rules* discussed earlier, and such a view of the matrix elements H'_{fi} is also at the heart of the physical interpretation usually attached to such perturbation-theoretic formulas as (6.35): the perturbation induces "virtual" (energy-nonconserving) transitions, from the level n to the level k and back, and, in higher order, from n to k to m and back, etc., and the probability amplitudes for all these processes are to be added up.

Since W_{fi} as given by (8.8) is the probability per unit time of a transition from the level i to the level f, the probability of the opposite transition, from f to i, is given by

$$W_{if} = \frac{2\pi}{\hbar}\rho(i)|H'_{if}|^2.$$

Thus, apart from the factor of $\rho(i)$, the density of final states, W_{if} and W_{fi} *are equal*, because $\mathbf{H}'$ is Hermitian. This is known as the *reciprocity theorem;* it is a quantum analogue of the classical reversibility of the equations of motion.

The first-order perturbation formula (8.5) is applicable not only when the perturbation is weak, but also when, however strong, it acts only for a very short time. This means, specifically, that the duration T of its action is such that $T \ll 1/\omega_{fi}$ for the level f nearest to i, and, in addition, $|H'_{fi}|T/\hbar \ll 1$. In that case (8.5) becomes for $f \neq i$,

$$P_{fi} = \frac{1}{\hbar^2}\left|\int_0^T dt\, H'_{fi}(t)\right|^2, \tag{8.9}$$

which is called the *sudden approximation.*

8.1.1 Application to exponential decay

Suppose at the time t there are $N_m(t)$ systems of the same kind occupying the m^{th} energy level of the Hamiltonian $\mathbf{H}_0$. Then the decrease in this population during the time dt, owing to transitions to the group n of (closely spaced) levels (assuming there are no others accessible), is[1]

$$-dN_m(t) = dt\, N_m(t)W_{nm}.$$

The solution of this equation,

$$N_m(t) = N_m(0)e^{-W_{nm}t},$$

is the familiar exponential decay law of the initial state, and it gives a "half-life" (in first-order perturbation theory),

$$\tau_m = \frac{1}{W_{nm}} = \frac{\hbar}{2\pi}\frac{1}{\rho(n)|H'_{nm}|^2}, \tag{8.10}$$

so that $N_m(\tau_m)/N_m(0) = 1/e$. An equivalent way of deriving it is to cut the time t into p small intervals of length t/p and to express $N_m(t)$ in the form

$$N_m(t) = N_m(0)\lim_{p\to\infty}\left(1 - W_{nm}\frac{t}{p}\right)^p = N_m(0)e^{-W_{nm}t}.$$

[1]If other groups are accessible, W_{nm} has to be replaced here by $\sum_n W_{nm}$.

There is, however, something quite puzzling about this result, so easily derived. The expression used for the transition probability after the time t, based on a constant transition probability per unit time, $P_{fi} = W_{fi}t$, must surely break down when $W_{fi}t > 1$. Yet we assume that the exponential decay law holds for times when $W_{fi}t > 1$; otherwise it would tell us very little. The explanation is most clearly seen in the second manner of deriving the exponential law given above. The underlying assumption is that the system is examined over and over again after short intervals; it is not left alone to develop on its own without having measurements performed. That this makes a difference, even if the outcomes of the intermediate measurements are ignored, is characteristic of all probabilistic theories; it is not a special, peculiar property of quantum mechanics.[2]

Therefore, the exponential decay law is an outcome of an assumption of almost constant observation (or of other coherence-destroying interactions, such as with the environment) of a system; it is not an exact law of quantum mechanics. In fact, a system in an unstable state, left to its own devices, decays exponentially neither initially nor asymptotically. In many circumstances, however, it does follow an exponential law for many lifetimes after an initial transient period, but the demonstration of this is much more complicated than that given above.[3]

In addition, the specific relation (8.10) between the lifetime and the matrix element H'_{nm} of the perturbation, is, of course, a first-order perturbation theory result; it ignores all the decays that can take place via intermediate states and, in particular, it ignores the possibility of back-reactions

---

[2]Consider the analogous situation in statistical mechanics. What is the probability P_{BA} of a given system, originally at the time $t = t_1$ in the state A, defined by a coarse grain $G_{t_1}(A)$ of volume $\Omega(G_{t_1}(A))$ in phase space, to be found in the state B, defined by a coarse grain $G(B)$, at a later time t_2? In order to calculate it, we have to study the Hamiltonian flow from A to B and count how many of the systems in $G_{t_1}(A)$, assumed uniformly distributed there, end up in $G(B)$; the result is

$$P_{BA} = \frac{\Omega(G_{t_1,t_2}(A) \cap G(B))}{\Omega(G_{t_1}(A))},$$

if $G_{t_1,t_2}(A)$ denotes the collection of systems at the time t_2 originally in $G_{t_1}(A)$. Similarly, for the time $t_3 > t_2$,

$$P_{CB} = \frac{\Omega(G_{t_2,t_3}(B) \cap G(C))}{\Omega(G_{t_2}(B))},$$

and the probability of finding the system at the time t_3 in the state C, given that it was in the state A at the time t_1 and was found in the state B at the time t_2, is $P_{BA}P_{CB}$. Now, if the result of the intermediate measurement is to be ignored, we have to sum over all the possible grains $G(B)$ that the system could have visited at the time t_2. But in general $\sum_B P_{BA}P_{CB} \neq P_{CA}$, because the states B that the system could visit at t_2 from A do not exhaust the phase space, and the grains it does visit are not necesssarily uniformly populated by the flow from A.

[3]See [Newton 61].

to the initial state. In many practical situations, such as the decay of an excited atomic energy level, that back-reaction is quite negligible because of the much larger phase-space volume available to the final states than in the initial state.

8.1.2 Application to scattering

Fermi's golden rule may also be applied to the calculation of scattering cross sections. For this purpose, imagine the particle system to be confined to a box of sidelength L, with periodic boundary conditions. Then the number of free states in the wave-vector interval d^3k' is given by

$$\rho(k')dE_{k'} = \rho(k')\frac{\hbar^2 k' dk'}{M} = d^3k'(L/2\pi)^3 = k'^2 dk'\, d\Omega(L/2\pi)^3,$$

from which we conclude that the density of final states is

$$\rho = \left(\frac{L}{2\pi}\right)^3 \frac{Mk'}{\hbar^2} d\Omega. \tag{8.11}$$

The number of particles of velocity v in the final solid-angle $d\Omega$, per unit time and incident flux v/L^3, is

$$\frac{d\sigma}{d\Omega}d\Omega = \frac{W}{v/L^3} = W\frac{L^3 M}{\hbar k},$$

so that by (8.8),

$$\frac{d\sigma}{d\Omega} = \frac{Mk'}{\hbar^2}\left(\frac{L}{2\pi}\right)^3 \frac{ML^3}{\hbar k}\frac{2\pi}{\hbar}|H'_{fi}|^2,$$

where now

$$H'_{fi} = \int d^3r\, \frac{e^{-i\vec{k}'\cdot\vec{r}}}{L^{3/2}} H'(\vec{r}) \frac{e^{i\vec{k}\cdot\vec{r}}}{L^{3/2}}.$$

Therefore we have in the first approximation, since $k' = k$,

$$\frac{d\sigma}{d\Omega} = \left(\frac{M}{2\pi\hbar^2}\right)^2 \left|\int d^3r\, V(\vec{r}) e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}\right|^2, \tag{8.12}$$

which is the Born approximation, as given by (4.137). Notice that the box size L, which served only an auxiliary purpose, has canceled out. The artificial assumption of a finite box, which makes the spectrum discrete, can be dropped, as we did earlier. In that case the proper delta-function normalization is achieved by chosing as the free wave functions $e^{i\vec{k}\cdot\vec{r}}/(2\pi)^{3/2}$, with one particle per volume $(2\pi)^3$, so that the flux is $v/(2\pi)^3$, and the density of states,

$$\rho(k) = \frac{Mk}{\hbar^2} d\Omega. \tag{8.13}$$

The same reasoning is applicable to *inelastic scattering.* In that case, the unperturbed Hamiltonian includes the interaction that leads to the inital and final bound states (such as the nuclear Coulomb potential responsible for the electronic states of an atom), and the initial and final states used for the matrix element are such bound states, combined with the free states of the scattered missile; the perturbation $\mathbf{H}'$ is the interaction that allows the system to make an energy nonconserving transition, the missing energy being accounted for by the difference between the initial and final energies of the missile. The Born-approximation result in that case is of the same form as given above, except that $k' \neq k$ and the formula (8.12) has to be multiplied by the ratio $k_{\text{final}}/k_{\text{initial}}$.

8.1.3 Second order perturbation theory

If the second-order term in the iterative solution of (8.4) is kept, the result is that (8.5) must be replaced by

$$\begin{aligned}\langle E_f|\,\rangle_t &\simeq \delta_{fi} + \frac{1}{i\hbar}\int_{-\infty}^{t} dt'\, H'_{fi}(t')e^{i\omega_{fi}t'} \\ &- \frac{1}{\hbar^2}\sum_k \int_{-\infty}^{t} dt''\, H'_{fk}(t'')e^{i\omega_{nk}t''}\int_{-\infty}^{t''} dt'\, H'_{ki}(t')e^{i\omega_{ki}t'}.\end{aligned}$$

For the case in which the interaction is simply constant until $t = T$, when it is switched off, the second-order term becomes

$$\begin{aligned}&- \frac{1}{\hbar^2}\sum_k H'_{fk}H'_{km}\int_{-\infty}^{t} dt''\, e^{i\omega_{fk}t''}\int_{-\infty}^{t''} dt'\, e^{i\omega_{ki}t'} \\ &= \frac{1}{\hbar^2}\sum_k \frac{H'_{fk}H'_{ki}}{\omega_{km}}\left[\frac{e^{i\omega_{fi}t}-1}{\omega_{fi}} - \frac{e^{i\omega_{fk}t}-1}{\omega_{fk}}\right],\end{aligned} \tag{8.14}$$

as a result of which we obtain

$$\begin{aligned}\langle E_f|\,\rangle_t \simeq \delta_{fi} &- \frac{1}{\hbar}\left[H'_{fi} - \sum_k \frac{H'_{fk}H'_{ki}}{E_k - E_i}\right]\frac{e^{i\omega_{fi}t}-1}{\omega_{fi}} \\ &- \sum_k \frac{H'_{fk}H'_{ki}}{\omega_{ki}}\,\frac{e^{i\omega_{fk}t}-1}{\omega_{fk}}.\end{aligned} \tag{8.15}$$

Hence, just as before, there will be a large contribution, increasing with time, when $\omega_{fi} \simeq 0$, but there is also such a contribution when $\omega_{fk} \simeq 0$, and the latter *need not conserve energy.* This effect, however, owes its existence entirely to the sudden switching on and off of the interaction; as we shall see shortly, when the switching is slow, no such energy non-conserving transitions will occur.

Paying attention only to the energy-conserving transitions, we thus find that the transition probability per unit time, to second order, is given by

$$\begin{aligned} W_{fi} &= \frac{2\pi}{\hbar}\,\rho(f)\left|H'_{fi}-\sum_k \frac{H'_{fk}H'_{ki}}{E_k-E_i}\right|^2 \\ &= \frac{2\pi}{\hbar}\rho(f)\left|\langle E_f|\mathbf{H}'-\mathbf{H}'\frac{1}{\mathbf{H}_0-E_i}\mathbf{H}'|E_i\rangle\right|^2. \end{aligned} \tag{8.16}$$

This form, however, works only when $H'_{fi}=0$, because otherwise the sum contains energy denominators that vanish; if $H'_{fi}\neq 0$, then the form given in (8.14), which remains finite when $k=i$, has to be retained.

Here is an intuitively appealing interpretation: In view of the physical meaning of the matrix element H'_{fi} as representing a transition from level i to level f, the terms $H'_{fk}H'_{ki}$ evidently represent two-step transitions via an intermediate state k. So the second-order term $\sum_k H'_{fk}H'_{ki}/(E_k-E_i)$ represents all the transitions via a "virtual" intermediate state. Since these intermediate states are of a transitory nature and the system is never actually experimentally caught there, energy need not be conserved. To say, then, that a certain transition is allowed only via n intermediate steps is equivalent to saying that the n^{th} order transition probability is the first one not to vanish. However, let us not attach too much reality to such physical interpretations of the mathematics; appealing and useful as a guide to our intuition as they may be, they are no more than ideas based on a particular method of approximation, namely, perturbation theory.

8.1.4 Interaction with electromagnetic fields

The interaction of charged particles with the electromagnetic field gives rise to the emission and absorption of radiation; if we shine light (or other electromagnetic radiation) on an atom, it will absorb photons and emit them, and the probability for such emissions and absorptions can be calculated by time-dependent perturbation theory. For a particle of charge e and no spin, subject to the scalar potential V, which may or may not be of electromagnetic origin, and to the external electromagnetic vector potential $\vec{A}$, the Schrödinger equation with minimal coupling reads

$$\left[\frac{1}{2M}(\vec{\mathbf{p}}-\frac{e}{c}\vec{A})^2+V\right]\psi = i\hbar\frac{\partial}{\partial t}\psi. \tag{8.17}$$

In the radiation gauge, where $\nabla\cdot\vec{A}=0$, the configuration-space Hamiltonian therefore is given by

$$\mathbf{H}=-\frac{\hbar^2}{2M}\nabla^2+V+\frac{e^2}{2Mc^2}\vec{A}^2+\frac{ie\hbar}{Mc}\vec{A}\cdot\nabla.$$

In order to study the emission and absorption of radiation, however, we have to treat the vector potential $\vec{A}$ not as an externally given function but as a quantum field like in Section 2.3.3. Of course, in the presence of charged particles, this field is not correctly described by the free Maxwell equations, but since we are going to confine ourselves to first-order perturbation theory, it will suffice to insert in (8.17) the free vector-potential operator, which is given by (2.102) for a fixed wave vector $\vec{k}$, apart from a multiplicative factor to be fixed later, in the form

$$\vec{\mathbf{A}}(\vec{r},t) = i[e^{i(\vec{k}\cdot\vec{r}-\omega t)}\vec{\mathsf{a}}(\vec{k}) - e^{-i(\vec{k}\cdot\vec{r}-\omega t)}\vec{\mathsf{a}}^\dagger(-\vec{k})], \tag{8.18}$$

in terms of the creation and annihilation operators $\vec{\mathsf{a}}^\dagger$ and $\vec{\mathsf{a}}$ of photons. The unperturbed, zero-order, state vectors of the system are then products of states of the free electromagnetic field, which we shall, for the time being, denote simply by the short-hand $|\mathrm{em}\rangle$, and, in the configuration representation, bound-state wave functions ψ_n of the particles, which are solutions of the Schrödinger equation (8.17) with $\vec{A} = 0$,

$$\mathbf{H}_0 = -\frac{\hbar^2}{2M}\nabla^2 + V.$$

The perturbation is given by

$$\mathbf{H}' = \frac{ie\hbar}{Mc}\vec{\mathbf{A}}\cdot\nabla,$$

while the $\mathbf{A}^2$-term will be neglected as too small to matter.

If the original state of the system is of the form $|i\rangle = |\mathrm{em}_i\rangle\psi_i(\vec{r})$ and the perturbation is turned on at the time $t = 0$, then at a later time t the inner product of the developing state $|\,\rangle_t$ with a state $|f\rangle = |\mathrm{em}_f\rangle\psi_f(\vec{r})$, according to (8.5), is given to first order by

$$\langle f|\,\rangle_t \simeq \delta_{fi} + \frac{1}{i\hbar}\int_0^t dt'\, H'_{fi}e^{i\omega_{fi}t}, \tag{8.19}$$

where

$$H'_{fi} \stackrel{\mathrm{def}}{=} \frac{ie\hbar}{Mc}\int d^3r\, \langle f|\vec{\mathbf{A}}\cdot\nabla|i\rangle. \tag{8.20}$$

This gives

$$\begin{aligned}\langle f|\,\rangle_t &\simeq \delta_{fi} + \frac{ie}{Mc}\left[H''\int_0^t dt'\, e^{i(\omega_{fi}-\omega)t} - H'''\int_0^t dt'\, e^{i(\omega_{fi}+\omega)t}\right] \\ &= \delta_{fi} + \frac{e}{Mc}\left[H''\frac{e^{i(\omega_{fi}-\omega)t}-1}{\omega_{fi}-\omega} - H'''\frac{e^{i(\omega_{fi}+\omega)t}-1}{\omega_{fi}+\omega}\right],\end{aligned} \tag{8.21}$$

where

$$\begin{aligned}H'' &\stackrel{\mathrm{def}}{=} \langle \mathrm{em}_f|\int d^3r\, \psi_f^*(\vec{r})e^{i\vec{k}\cdot\vec{r}}\vec{\mathsf{a}}(\vec{k})\cdot\nabla\psi_i(\vec{r})|\mathrm{em}_i\rangle \\ &= \langle \mathrm{em}_f|\vec{\mathsf{a}}|\mathrm{em}_i\rangle\cdot\int d^3r\, \psi_f^*(\vec{r})e^{i\vec{k}\cdot\vec{r}}\nabla\psi_i(\vec{r})\end{aligned}$$

and

$$\begin{aligned} H''' &\stackrel{\text{def}}{=} \langle \text{em}_f | \int d^3r\, \psi_f^*(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} \vec{\mathsf{a}}^\dagger(\vec{k}) \cdot \nabla \psi_m(\vec{r}) | \text{em}_i \rangle \\ &= \langle \text{em}_f | \vec{\mathsf{a}}^\dagger | \text{em}_i \rangle \cdot \int d^3r\, \psi_f^*(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} \nabla \psi_i(\vec{r}). \end{aligned}$$

As we discussed before, the expression (8.21) is of appreciable size only when the denominators in it are small, that is, when either $E_f = E_i + \hbar\omega$ or $E_f = E_i - \hbar\omega$. In the first instance, the factor H'' vanishes unless the final electromagnetic field contains one less photon of energy $\hbar\omega$, momentum $\hbar\vec{k}$, and polarization indicated by $\vec{\mathsf{a}}$, while in the second instance H''' vanishes unless the final field contains one additional such photon. Thus the first case is one in which the bound system, for example, an atom, absorbs a photon and jumps to a higher level whose additonal energy equals that of the photon, and the second case is one in which the bound system descends to a lower level, emitting a corresponding photon. Of course, since we have fixed the perturbation to have a given wave number, there may not be such energy levels of the bound system. Furthermore, we have already seen that a constant transition probability per unit time will emerge only if transitions to a group of states with closely spaced energies are considered.

Let us, then, assume that the radiation field consists, not of a single plane wave, but of an incoherent mixture of many plane waves with no phase relation between them, their intensities distributed according to a function $f(\omega)$. In that case we have to add the transition probabilities and we obtain for the case of absorption,

$$|\langle f | \rangle_t|^2 \simeq \frac{4e^2}{M^2c^2} \int d\omega\, f(\omega) |H''|^2 \frac{\sin^2[\frac{1}{2}(\omega_{fi} - \omega)t]}{(\omega_{fi} - \omega)^2}.$$

The direction of the matrix element $\langle \text{em}_f | \vec{\mathsf{a}} | \text{em}_i \rangle$ should be expressible in terms of a unit vector $\widehat{e}_{\text{pol}}$ in the direction of the polarization of the radiation. To express the magnitude of the matrix element in terms of an intensity, use the fact that the Poynting vector of the radiation described by the vector potential $\vec{A}(\vec{r},t) = \vec{A}_0 e^{i(\vec{k}\cdot\vec{r}-\omega t)} + \vec{A}_0^* e^{-i(\vec{k}\cdot\vec{r}-\omega t)}$ is given by $\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B} = \frac{\omega}{\pi}\vec{k}|\vec{A}_0|^2 \sin^2(\vec{k}\cdot\vec{r} - \omega t + \alpha)$, so that its *intensity*, the (time) average value of the magnitude of the Poynting vector, is

$$I = |\vec{S}|_{\text{av}} = \frac{\omega^2}{2\pi c} |f(\omega)\vec{A}_0|^2.$$

Therefore the square of its magnitude, weighted by $f(\omega)$, should be

$$\begin{aligned} f(\omega) |\langle \text{em}_f | \vec{\mathsf{a}} | \text{em}_i \rangle|^2 &= f(\omega) \sum_{n_f} \langle \text{em}_i | \vec{\mathsf{a}}^\dagger | n_f \rangle \langle n_f | \vec{\mathsf{a}} | \text{em}_i \rangle \\ &= f(\omega) \langle \text{em}_i | \vec{\mathsf{a}}^\dagger \vec{\mathsf{a}} | \text{em}_i \rangle = f(\omega) n_i(\omega) \\ &= \frac{2\pi c}{\omega^2} I(\omega), \end{aligned}$$

in which we have made use of the fact that all the additional terms in the sum over the number n_f of photons in the final state vanish because $\vec{\mathsf{a}}$ destroys just one photon of energy $\hbar\omega_i$. This allows us to write the transition probability as

$$|\langle f|\,\rangle_t|^2 \simeq \frac{8\pi e^2}{M^2 c}\int \frac{d\omega}{\omega^2}\, I(\omega)\frac{\sin^2[\frac{1}{2}(\omega_{fi}-\omega)t]}{(\omega_{fi}-\omega)^2}\left|\int d^3r\, e^{i\vec{k}\cdot\vec{r}}\psi_f^*\widehat{e}_{\rm pol}\cdot\nabla\psi_i\right|^2 .$$

From this point on we proceed as before: in the limit when $t\to\infty$ the probability becomes very sharply peaked at $\omega=\omega_{fi}$. Assuming that the intensity does not vary rapidly near $\omega=\omega_{fi}$, we take it and all the other functions of ω outside the integral, leaving us with the absorption probability per unit time,

$$\begin{aligned} P_{fi}^{\rm abs} &= \frac{1}{t}|\langle f|\,\rangle_t|^2 \simeq \frac{8\pi e^2}{M^2c}\frac{I(\omega_{fi})}{\omega_{fi}^2}\mathcal{M}_{fi}\int_{-\infty}^{\infty} d\omega\,\frac{\sin^2[\frac{1}{2}(\omega_{fi}-\omega)t]}{t(\omega_{fi}-\omega)^2} \\ &= \frac{4\pi^2e^2}{M^2c}\frac{I(\omega_{fi})}{\omega_{fi}^2}\mathcal{M}_{fi}, \end{aligned} \tag{8.22}$$

where

$$\mathcal{M}_{fi} \stackrel{\rm def}{=} \left|\int d^3r\, e^{i\vec{k}\cdot\vec{r}}\psi_f^*\widehat{e}_{\rm pol}\cdot\nabla\psi_i\right|^2, \tag{8.23}$$

with the understanding that $|\vec{k}|=|\omega_{fi}|/c$.

Similarly we proceed to calculate the probability per unit time for *induced emission* from the second term in (8.21), with the result

$$P_{fi}^{\rm emi} = \frac{4\pi^2e^2}{M^2c}\frac{I(\omega_{fi})}{\omega_{fi}^2}\mathcal{M}'_{fi},$$

where $I(\omega_{fi})=I(-\omega_{fi})=I(\omega_{if})$ is necessary for $I(t)$ to be real, and

$$\mathcal{M}'_{fi} \stackrel{\rm def}{=} \left|\int d^3r\, e^{-i\vec{k}\cdot\vec{r}}\psi_f^*\widehat{e}_{\rm pol}\cdot\nabla\psi_i\right|^2 = \left|\int d^3r\, e^{i\vec{k}\cdot\vec{r}}\psi_i^*\widehat{e}_{\rm pol}\cdot\nabla\psi_f\right|^2 = \mathcal{M}_{if},$$

which is proved by integrating by parts and using the fact that $\widehat{e}_{\rm pol}\cdot\nabla e^{i\vec{k}\cdot\vec{r}} = 0$, because $\widehat{e}_{\rm pol}\cdot\vec{k}=0$. We therefore find that

$$P_{fi}^{\rm emi} = P_{if}^{\rm abs}. \tag{8.24}$$

The radiation field induces as many transitions in one direction as in the other.

The dipole approximation

If the size (i.e., the spread of its bound-state wave function) of the bound system, the atom, is small compared to the wavelength of the emitted

or absorbed radiation, then in the region which contributes almost all of the value of the integral in $\mathcal{M}$ we have $\vec{k}\cdot\vec{r} \ll 1$ and it will be a good approximation to set $e^{i\vec{k}\cdot\vec{r}} \simeq 1$ in the integral, so that

$$\mathcal{M}_{fi} \simeq \left| \int d^3r\, \psi_f^* \widehat{e}\cdot\nabla\psi_i \right|^2 = \left| \frac{M}{\hbar}\omega_{if} \int d^3r\, \psi_f^* \widehat{e}\cdot\vec{r}\psi_i \right|^2 . \tag{8.25}$$

To prove this, use integrations by parts, assuming that the boundary terms vanish:

$$\begin{aligned}
&(E_i - E_f)\int d^3r\, \psi_f^* \widehat{e}\cdot\vec{r}\psi_i = \int d^3r\, \widehat{e}\cdot\vec{r}(\psi_f^*\mathbf{H}\psi_i - \psi_i\mathbf{H}\psi_f^*) \\
= \;& -\frac{\hbar^2}{2M}\int d^3r\, \widehat{e}\cdot\vec{r}(\psi_f^*\nabla^2\psi_i - \psi_i\nabla^2\psi_f^*) = -\frac{\hbar^2}{2M}\int d^3r\, \widehat{e}\cdot\vec{r}\nabla\cdot(\psi_f^*\nabla\psi_i - \psi_i\nabla\psi_f^*) \\
= \;& \frac{\hbar^2}{2M}\int d^3r\, \widehat{e}\cdot(\psi_f^*\nabla\psi_i - \psi_i\nabla\psi_f^*) = \frac{\hbar^2}{M}\int d^3r\, (\psi_f^*\widehat{e}\cdot\nabla\psi_i).
\end{aligned}$$

Therefore we obtain in this approximation,

$$P_{fi}^{\text{abs}} = \frac{4\pi^2 e^2}{c\hbar^2} I(\omega_{fi}) |\widehat{e}_{\text{pol}}\cdot(\vec{r})_{fi}|^2, \tag{8.26}$$

where $(\vec{r})_{fi}$ is shorthand for $\int d^3r\, \psi_f^*\vec{r}\psi_i$. Since ψ_f and ψ_i are orthogonal, the vector $\vec{r}$ may be taken with respect to any arbitrary origin, and $e\vec{r}$ is the electric dipole moment of a particle at $\vec{r}$ with respect to that origin. That is why the transitions for which (8.26) is the approximate probability per unit time are called *dipole transitions.* If θ is the angle between the direction of polarization $\widehat{e}_{\text{pol}}$ and $\vec{r}$, we have $|\widehat{e}_{\text{pol}}\cdot(\vec{r})_{fi}|^2 = |(\vec{r})_{fi}|^2\cos^2\theta$. Averaging this over all directions of polarization gives the absorption probability per unit time from an unpolarized radiation field[4]

$$P^{\text{abs}} = \mathfrak{B} I(\omega_{fi})/c, \tag{8.27}$$

where

$$\mathfrak{B} = \frac{4\pi^2 c\alpha}{3\hbar} |(\vec{r})_{fi}|^2 \tag{8.28}$$

is called the *Einstein B coefficient.* (Here α is the fine-structure constant $\alpha \stackrel{\text{def}}{=} e^2/\hbar c$.)

The fact that $\vec{r}$ is a vector immediately leads to certain *selection rules* for dipole transitions. As we have seen in Section 5.2, if the potential that binds the system is spherically symmetric, for (8.28) not to vanish we must have

$$\Delta l = \pm 1, \qquad \Delta m = 0, \pm 1,$$

[4] **Prove that for the three-dimensional avergage,** $\overline{\cos^2\theta} = 1/3$.

(if l is not a good quantum number, the first is replaced by $\Delta j = 0, \pm 1$), that is, dipole transitions can occur only between states for which $\Delta l = \pm 1$ and $\Delta m = 0, \pm 1$; if the dipole transition probability is zero, such transitions are called *forbidden.*

If the incident radiation is not incoherent but plane polarized along the z-axis, the relevant dipole matrix element is $(z)_{fi}$, for which the selection rule is $\Delta m = 0$; on the other hand, if the radiation is polarized along the x or y-axis (or in any direction perpendicular to the z-axis), then $\Delta m = \pm 1$. Of course, for these selection rules to make any observable difference, the m-degeneracy must be removed, say, by a strong external magnetic field (the Paschen-Back effect). In such a situation, we must have $\Delta m = 0$ for polarization along the magnetic field, and $\Delta m = \pm 1$ for polarizaton that is orthogonal to it; if the radiation is circularly polarized, the selection rule is $\Delta m = 1$ for right-circular polarization and $\Delta m = -1$ for left-circular polarization.[5]

In case a certain dipole transition is forbidden, more terms in the power-series expansion

$$e^{i\vec{k}\cdot\vec{r}} = 1 + i\vec{k}\cdot\vec{r} + \dots$$

or in the spherical-wave expansion (see Appendix D)

$$e^{i\vec{k}\cdot\vec{r}} = \sum_{l=0}^{\infty} i^l j_l(kr) P_l(\cos\theta)$$

have to be taken into account. Every term in the ensuing expansion of the transition probability is then smaller by a factor of ak relative to the previous one, if a is the spatial extent of the relevant bound-state wave function. Thus if $ak \ll 1$, the probability of a forbidden transition is suppressed at least by a factor of $(ak)^2$ relative to an allowed one.

In some instances, the exact matrix element (8.23) may vanish; in that case the transition is called *strictly forbidden* and one has to employ higher order perturbation theory to calculate the transition probability, which then is very small. An important example of such a situation is one in which both the initial and final states have $l = 0$. They are then both spherically symmetric, while $\widehat{e}_{\rm pol}\cdot\vec{k} = 0$; hence, if the z-axis is chosen along the polarization direction $\widehat{e}_{\rm pol}$, the integrand is odd in z and thus vanishes. Consequently, $l = 0 \to l = 0$ *transitions are strictly forbidden.*

Note that all of these statements hold both for absorption and for induced emission of radiation.

Spontaneous emission of radiation

The interaction of a bound electrically charged system with the radiation field gives rise not only to absorption and induced emission of photons but

[5] **Show this.**

also to their spontaneous emission, with no light or other radiation shining on the system. In order to calculate the probability for this to occur, we avail ourselves of Fermi's golden rule (8.8), which means we must first determine what the appropriate density ρ_f of final states is for photons. If the field is confined to a cubical box of sidelength L, we know that the allowed wavelengths are $\lambda = L/n$, so that the allowed wavenumbers for each of the three orthogonal directions are $2\pi n/L$, and the number of states in the wave-vector interval d^3k is $(L/2\pi)^3 d^3k$. Therefore,

$$\rho dE = \rho\hbar d\omega = \left(\frac{L}{2\pi}\right)^3 d^3k = d\Omega k^2 dk \left(\frac{L}{2\pi}\right)^3 = d\Omega \frac{\omega^2 d\omega}{c^3}\left(\frac{L}{2\pi}\right)^3,$$

or

$$\rho = d\Omega \frac{\omega^2}{\hbar c^3}\left(\frac{L}{2\pi}\right)^3. \tag{8.29}$$

Since no radiation shines on the atom, the initial state is the photon vacuum state and only the creation operator $\vec{\mathsf{a}}^\dagger$ in (8.18) will contribute, leading to a state with one photon of momentum $\vec{k}\hbar$, so that the appropriate normalization of $\vec{\mathbf{A}}$ for finding one photon per unit volume has to be such that

$$\vec{\mathbf{A}}|0\rangle = ic\sqrt{\frac{2\pi\hbar}{\omega}} L^{-3/2} e^{i(\vec{k}\cdot\vec{r}+\omega t)} \vec{\mathsf{a}}^\dagger(\vec{k})|0\rangle$$

and (8.20) gives

$$|H'_{fi}|^2 = \frac{2\pi\hbar^3 e^2}{M^2\omega L^3}\mathcal{M}_{fi},$$

where $\mathcal{M}_{fi}$ is defined by (8.23). Insertion of this and (8.29) in (8.8) gives the probability per unit time for the spontaneous emission of a single photon of plane polarization $\widehat{e}_{\rm pol}$ and momentum in the interval $d\Omega$ around $\vec{k}$ from the state ψ_i of the charged system to the state ψ_f, whose energy is lower by the amount $\hbar\omega$, the energy of the emitted photon,

$$W_{fi} = d\Omega \frac{e^2\hbar\omega}{2\pi M^2 c^3}\mathcal{M}_{fi}. \tag{8.30}$$

In the dipole approximation, in which $e^{i\vec{k}\cdot\vec{r}}$ is replaced by 1, this expression becomes, by (8.25),

$$W_{fi} = \frac{\alpha}{2\pi c^2}\omega^3 |\widehat{e}_{\rm pol}\cdot(\vec{r})_{fi}|^2 d\Omega.$$

If ϕ is the angle between the polarization of the emitted radiation and the real vector $\vec{R}_{fi}$ formed by the magnitudes of the three components of the complex vector $(\vec{r})_{fi}$, then $|\widehat{e}_{\rm pol}\cdot(\vec{r})_{fi}|^2 = |(\vec{r})_{fi}|^2\cos^2\phi$. Let θ be the angle between $\vec{R}_{fi}$ and $\vec{k}$, and φ' the angle between the plane formed by $\vec{k}$ and R_{fi} and the polarization vector $\widehat{e}_{\rm pol}$; then $\cos^2\phi = \sin^2\theta\cos^2\varphi'$, so that

$$W_{fi} = \frac{\alpha}{2\pi c^2}\omega^3 |(\vec{r})_{fi}|^2 d\Omega \sin^2\theta\cos^2\varphi',$$

whereas the probability of emitting a photon whose polarization is orthogonal to $\widehat{e}_{\rm pol}$ is given by the same expression with $\cos^2\varphi'$ replaced by $\sin^2\varphi'$. Therefore the probability of emitting a photon into $d\Omega$, irrespective of its polarization, is

$$W_{fi} = \frac{\alpha}{2\pi c^2}\omega^3 |(\vec{r})_{fi}|^2 d\Omega \sin^2\theta.$$

Finally we integrate over all emission directions and obtain the probability per unit time of emitting a photon in any direction while the system makes a transition with an energy difference $E_i - E_f = \hbar\omega_{if}$ (in the dipole approximation),

$$W_{fi} = \mathfrak{A} \stackrel{\rm def}{=} \frac{4}{3}\frac{\alpha\omega_{if}^3}{c^2}|(\vec{r})_{fi}|^2, \tag{8.31}$$

which is called the *Einstein A coefficient.* Since each photon carries away the energy $\hbar\omega$, the emitted power is given by

$$P = \frac{4}{3}\frac{e^2\omega^4}{c^3}|(\vec{r})_{fi}|^2,$$

an expression that should be compared to the classical result for the average power emitted by a moving electric dipole, which is

$$P = \frac{2}{3c^3}\overline{\ddot{\vec{\mathfrak{p}}}^2},$$

where $\vec{\mathfrak{p}}$ is the electric dipole moment. In comparing the two expressions, keep in mind that if $\vec{\mathfrak{p}} = \vec{\mathfrak{p}}_0 e^{-i\omega t} + c.c. = 2\vec{\mathfrak{p}}_0 \cos\omega t$, then $\overline{\ddot{\vec{\mathfrak{p}}}^2} = 4\omega^4 \vec{\mathfrak{p}}_0^2 \overline{\cos^2\omega t} = 2\omega^4 \vec{\mathfrak{p}}_0^2$, so that the quantum result can be simply obtained by the rule: replace $\vec{\mathfrak{p}}$ by the dipole matrix element $(e\vec{r})_{fi}$ corresponding to the transition.

The selection rules that hold for absorption are equally applicable to spontaneous emission. Suppose the atom is immersed in a strong static magnetic field, splitting the normal m-degeneracy. Adopt a z-axis in the direction of the magnetic field. Then $(z)_{fi}$ vanishes unless $\Delta m = 0$, and $(x)_{fi}$ and $(y)_{fi}$ vanish unless $\Delta m = \pm 1$. Therefore the light emitted in a transition with $\Delta m = 0$ must be polarized along the z-axis (i.e, parallel to the magnetic field), which implies that its intensity is maximal in the plane at right angles to $\vec{B}$; it is called π-light. On the other hand, the light emitted in a transition with $\Delta m = \pm 1$ can have no polarization component in the direction of the magnetic field; such transitions give rise to right and left circularly polarized light (σ-light), respectively, with maximal emission intensity parallel to the magnetic field.

Planck's formula

Suppose the walls of a cavity made of atoms at the temperature T contains radiation with the frequency distribution $I(\omega)$; in other words, the energy in

the frequency interval $d\omega$ around ω is $dE = I(\omega)d\omega$. The electric charges in the walls in a state of energy E_m are induced to emit radiation of frequency ω at the rate $I(\omega)\mathfrak{B}$, where $\mathfrak{B}$ is given by (8.28), and they spontaneously emit radiation at the rate $\mathfrak{A}$ given by (8.31), descending to the state E_n with $E_m - E_n = \hbar\omega$, while they absorb radiation at the rate $I(\omega)\mathfrak{B}$ making the opposite transition. When in thermodynamic equilibrium, the ratio of the number of atoms in the upper state to that in the lower one should be $\exp(-\hbar\omega/kT)$, where k is Boltzmann's constant, so that the following equation must hold in a state of eqilibrium:

$$e^{-\hbar\omega/kT}\left[\frac{4e^2\omega^3}{3\hbar c^3}|(\vec{r})_{nm}|^2 + \frac{4\pi^2e^2}{3\hbar^2 c}I(\omega)|(\vec{r})_{nm}|^2\right] = \frac{4\pi^2e^2}{3\hbar^2 c}I(\omega)|(\vec{r})_{nm}|^2.$$

As a result, the frequency distribution of the radiation in the cavity has to be given by the solution of this equation, which is known as the *Planck distribution formula:*

$$I(\omega) = \frac{\hbar\omega^3}{\pi^2c^2\left(e^{\hbar\omega/kT} - 1\right)}. \tag{8.32}$$

Note that this distribution is independent of all the quantities (such as the dipole moments $e\vec{r}_{nm}$ and the energies E_n and E_m) that refer to the properties of the walls and can therefore be expected to be universally applicable. As $\hbar \to 0$, it goes over into the classical Rayleigh-Jeans law,

$$I_{\text{class}}(\omega) = \frac{\omega^2}{\pi^2c^2}kT,$$

the modification of which to the form (8.32) by Max Planck, you will recall, was the beginning of the quantum theory. The experimental verification of Planck's formula serves as an indirect confirmation of Einstein's A and B coefficients.

Linebreadth

The fact that a bound system of electrically charged particles spontaneously emits radiation when at an energy level from which it is able to descend to one of lower energy implies that all such states, which would be bound states if it were not for the electric charge of the constituents, are unstable; only the ground state of a charged system is truly stable. Indeed, the total probability per unit time of a decay of a given level labeled i is given by

$$W_i = \sum_f W_{fi},$$

where the sum runs over all the states of lower energy than E_i; as a result, the lifetime of the state is $\tau = 1/W_i$. But, according to the uncertainty

principle, a state with the lifetime τ can be thought of as having a "width" $\Gamma = \hbar/\tau$; that is, its energy can be defined only with an error $\pm\frac{1}{2}\Gamma$. A classical analogue would be a simple, charged, harmonic oscillator, whose emission of radiation leads to a damping of its amplitude, so that its time dependence is given by $e^{-i\omega_0 t}e^{-\frac{1}{2}\gamma t}$ with $\gamma = 1/\tau$. If such an oscillator is set to start at the time $t = 0$, the Fourier transform of its motion is given by $1/[i(\omega_0 - \omega) - \frac{1}{2}\gamma]$, the square of whose absolute magnitude is $1/[(\omega_0 - \omega)^2 + \frac{1}{4}\gamma^2]$. Since $\frac{1}{2}\gamma$ is the fractional decrease in the oscillator's amplitude, γ is the fractional decrease of its energy and $\tau = 1/\gamma$ is the time after which the oscillator has lost approximately half of its energy; γ is also the width, at half maximum, of the "spectral line" whose shape is shown in Figure 4.2 on page 108. At the energies $E = E_0 \pm \frac{1}{2}\Gamma$, $\Gamma = \hbar\gamma = \hbar/\tau$, the photon emission has dropped to half its maximum at $E = E_0$. The resulting broadening effect on the atomic energy levels is called their *natural line breadth.* For atomic dipole transitions, the relative frequency width is about $\gamma/\omega \sim 10^{-7}$, which implies a lifetime[6] $\tau \sim \frac{1}{3} \times 10^{-8}$sec. Thus, the normal half-life of excited states of atomic systems is about $\frac{1}{3} \times 10^{-8}$sec.

8.2 Slowly Changing Hamiltonians

In order to solve the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = \mathbf{H}(t)\Psi \tag{8.33}$$

with a time-dependent Hamiltonian, it is useful to introduce *instantaneous eigenvectors* of $\mathbf{H}(t)$ (which, of course, vary with time),

$$\mathbf{H}(t)\Phi_n(t) = E_n(t)\Phi_n(t), \tag{8.34}$$

and we shall assume for simplicity that the spectrum is purely discrete, with no degeneracies. At the time t, the solution of (8.33) can be expanded on the basis of these states and we define the coefficients $c_n(t)$ by writing the expansion in the form

$$\Psi(t) = \sum_n c_n(t)\Phi_n(t)\, \exp\left[-\frac{i}{\hbar}\int_0^t dt'\, E_n(t')\right], \tag{8.35}$$

insertion of which in (8.33), with the use of (8.34), leads to

$$\sum_n (\dot{\Phi}_n c_n + \dot{c}_n\Phi_n) \exp\left[-\frac{i}{\hbar}\int_0^t dt'\, E_n(t')\right] = 0,$$

[6] **Check this.**

if the time-derivative is denoted by a dot. Using the orthogonality of the Φ_n (and assuming they are normalized), this leads to the equation

$$\dot{c}_k = -\sum_n c_n(\Phi_k, \dot{\Phi}_n)\exp\left[-\frac{i}{\hbar}\int_0^t dt'\,(E_n(t') - E_k(t'))\right]. \tag{8.36}$$

On the other hand, differentiating (8.34) with respect to the time and taking the inner product of the result with Φ_k, yields for $k \neq n$,

$$(\Phi_k, \dot{\Phi}_n) = \frac{\langle E_k|\dot{\mathbf{H}}|E_n\rangle}{E_n - E_k}. \tag{8.37}$$

Finally, defining the time-dependent phase of the state Φ_n by[7]

$$\dot{\gamma}_n \stackrel{\text{def}}{=} i(\Phi_n, \dot{\Phi}_n), \tag{8.38}$$

and defining

$$a_k \stackrel{\text{def}}{=} c_k \exp\left[-i\int_0^t dt'\,\dot{\gamma}_k(t')\right] = c_k e^{-i\gamma_k(t)},$$

we arrive at the following differential equation for the coefficients a_k

$$\dot{a}_k = \sum_{n\neq k} a_n \frac{\dot{H}_{kn}}{E_n - E_k}\exp\left[-i\int_0^t dt'\,\omega_{nk} - i(\gamma_k - \gamma_n)\right], \tag{8.39}$$

where we have written $\langle E_k|\dot{\mathbf{H}}|E_n\rangle \stackrel{\text{def}}{=} \dot{H}_{kn}$. Assuming that the time-change of the Hamiltonian began at the time $t = 0$ when the system was in the state $\Phi_m(0)$, the initial condition to go with the differential equations (8.39) is

$$a_k(0) = \delta_{km}.$$

Up to this point, we have made no approximations. However, suppose now the Hamiltonian changes very slowly; for example, let it have the form $\mathbf{H}(t) = f(\varepsilon t)\mathbf{H}$, where $f(t)$ is a given continuous function that describes the switching-on of the perturbation $\mathbf{H}$, such that $f(-\infty) = 0$ and $f(0) = 1$, and let $\varepsilon \ll 1$, making the rate of change very small. Then the right-hand side of (8.39) is negligible and so are all the a_k except for a_m, which remains close to 1, implying that

$$|(\Psi(t), \Phi_n(t)|^2 = |a_n(t)|^2 \simeq \delta_{nm}.$$

Thus the probability of finding the system in any state other than that of the slowly changing energy $E_m(t)$ is negligible. Instead of inducing transitions

[7] **Prove that $\dot{\gamma}_n$ is real.**

to states of other energies, as would be the case if the change were more rapid, the slowly varying Hamiltonian acts *adiabatically;* the system stays in the "same state" whose energy slowly adapts itself. This is called the *adiabatic approximation.*

There is another noteworthy effect associated with such an adiabatic change in the Hamiltonian, and that is the change in the phase $\gamma_m(t)$ of the state vector or the eigenfunction. Suppose that at the time T the Hamiltonian returns to its original value that it had at the time $t = 0$. The system will then have returned to its original energy E_m, but its eigenfunction will have acquired an additonal phase $\gamma_m(T)$, known as the *geometric phase* or *Berry's phase.*[8] It is experimentally observable when two identical systems, one of which has gone through a closed circuit, are brought together and the interference effects produced by their phase difference are measured.

For example, take a case in which the time-dependence of the Hamiltonian describes an externally produced motion of a part A of it in physical, three-dimensional space, parametrized by a vector $\vec{R}(t)$, such as the position of the center of mass of A, so that the time dependence of $\Phi_m(t)$ is the result of the fact that Φ_m depends on $\vec{R}$ and $\Phi_m(t) = \Phi_m(\vec{R}(t))$. If during the time from $t = 0$ to T, $\vec{R}$ slowly performs a closed loop, then

$$\begin{aligned}\gamma_m(T) &= -\Im \int_0^T dt\, (\Phi_m(\vec{R}(t)), \dot{\Phi}_m(\vec{R}(t)) \\ &= -\Im \oint d\vec{R} \cdot (\Phi_m(\vec{R}(t)), \nabla_{\vec{R}} \Phi_m(\vec{R}(t)),\end{aligned}$$

an integral that does not necessarily vanish. In fact, by Stokes's theorem,[9]

$$\begin{aligned}\gamma_m(T) &= -\Im \int_C d\vec{S} \cdot \nabla_{\vec{R}} \times \langle E_m | \nabla_{\vec{R}} | E_m \rangle \\ &= -\Im \int_C d\vec{S} \cdot (\nabla_{\vec{R}} \Phi_m, \times \nabla_{\vec{R}} \Phi_m) \\ &= -\Im \int_C d\vec{S} \cdot \sum_{n \neq m} (\nabla_{\vec{R}} \Phi_m, \Phi_n) \times (\Phi_n, \nabla_{\vec{R}} \Phi_m),\end{aligned}$$

where C denotes the area encircled by the loop described by A. Just as (8.37) was obtained, we now have for $n \neq m$

$$\langle E_m | \nabla_{\vec{R}} | E_n \rangle = (\nabla_{\vec{R}} H)_{mn} / (E_n - E_m),$$

where $(\nabla_{\vec{R}} H)_{mn} \stackrel{\text{def}}{=} \langle E_m | (\nabla_{\vec{R}} \mathbf{H}) | E_n \rangle$, and as a result, the geometric phase is given by

$$\gamma_m(T) = -\Im \int_C d\vec{S} \cdot \sum_{n \neq m} \frac{(\nabla_{\vec{R}} H)_{mn} \times (\nabla_{\vec{R}} H)_{nm}}{[E_n(\vec{R}) - E_m(\vec{R})]^2}. \tag{8.40}$$

[8] After M.V. Berry.

[9] **Why is $n = m$ excluded in the last line?**

8.3 Problems and Exercises

1. Integrate (8.32) and thereby derive the Stefan-Boltzmann law for radiation from a black body.

2. Using Fermi's golden rule, calculate the lifetime of a state of the hydrogen atom immersed in a weak, uniform, constant electric field.

3. An unpolarized ultraviolet light beam shines on a hydrogen atom in its ground state. (a) Calculate the absorption probability per unit time (in the dipole approximation) if the wavelength of the beam is tuned to 2.2157×10^{-5}cm. (b) Suppose the beam is plane polarized. Using the polarization direction as the axis of quantization, i.e., the z-axis, which m-sublevels of the excited atom will be populated? (c) If you wish to populate specifically a sublevel with $m = 1$ or $m = -1$ and no others, what will you have to do?

4. Taking the fine structure into account, calculate the absorption probability per unit time for linearly polarized radiation by a hydrogen atom at a frequency appropriate for the transition from the ground state (a) to the $2p\frac{1}{2}$ level, and (b) to the $2p\frac{3}{2}$ level. (Calculate this in lowest-order perturbation theory, without taking into account the change in the wave functions produced by the fine structure perturbation.)

5. Let the polar angles of the wave vector $\vec{k}$ be θ and φ, and let those of the polarization vector be θ' and φ'. Calculate the probability (in the dipole approximation) for the spontaneous emission of a photon by an atomic system if $\vec{r}_{fi}$ is (a) in the z-direction, (b) in the x-direction, and (c) in the y-direction. From these results, show that if the photon comes from a transition with $\Delta m = 0$, then the emission probability is maximal in the xy-plane, and if it comes from a transition with $\Delta m = \pm 1$, then the emission probability is maximal in the z-direction.

9
Multiparticle Systems

9.1 Two Particles Interacting with a Center

In Chapter 4 we studied systems consisting of two particles interacting with one another and saw that in such a case the Schrödinger equation can be reduced to that of a single-particle system in the center-of-mass coordinate frame. In this section we shall consider systems that consist of two particles in interaction not only with each other but also with an external center of force. Think, for example, of a helium atom with an infinitely massive nucleus. Whereas in Section 6.3.2 we applied perturbation theory and the variational method to the calculation of the ground-state energy of this particular system, here we want to study other, more general phenomena that can occur in such instances, such as inelastic scattering, resonance scattering, and auto-ionization.

The Hamiltonian of our system is assumed to be of the form

$$\mathbf{H} \stackrel{\text{def}}{=} \mathbf{T}_a + \mathbf{T}_b + V_a(\vec{\mathbf{r}}_a) + V_b(\vec{\mathbf{r}}_b) + V^{(ab)}(\vec{\mathbf{r}}_{ab}), \tag{9.1}$$

where $\mathbf{T}_a$ and $\mathbf{T}_b$ are the two kinetic-energy operators,

$$\mathbf{T}_a \stackrel{\text{def}}{=} \frac{\vec{\mathbf{p}}_a^{\,2}}{2M}, \qquad \mathbf{T}_b \stackrel{\text{def}}{=} \frac{\vec{\mathbf{p}}_b^{\,2}}{2M'},$$

and $\vec{\mathbf{r}}_{ab} \stackrel{\text{def}}{=} \vec{\mathbf{r}}_a - \vec{\mathbf{r}}_b$. In the absence of interaction between the two particles, that is, when $V^{(ab)} = 0$, the Schrödinger equation is separable and the solutions of (9.1) are linear combinations of products of solutions of the

one-particle Schrödinger equations,

$$[\mathbf{T}_a + V_a(\vec{\mathbf{r}}_a)]\psi^{(a)} = E^{(a)}\psi^{(a)}, \qquad [\mathbf{T}_b + V_b(\vec{\mathbf{r}}_b)]\psi^{(b)} = E^{(b)}\psi^{(b)}, \tag{9.2}$$

so that the eigenvalues and quasi-eigenvalues of (9.1) are the sums of the eigenvalues and quasi-eigenvalues of the two equations in (9.2). Thus the spectrum of the two-particle system looks schematically as indicated in Figure 9.1.

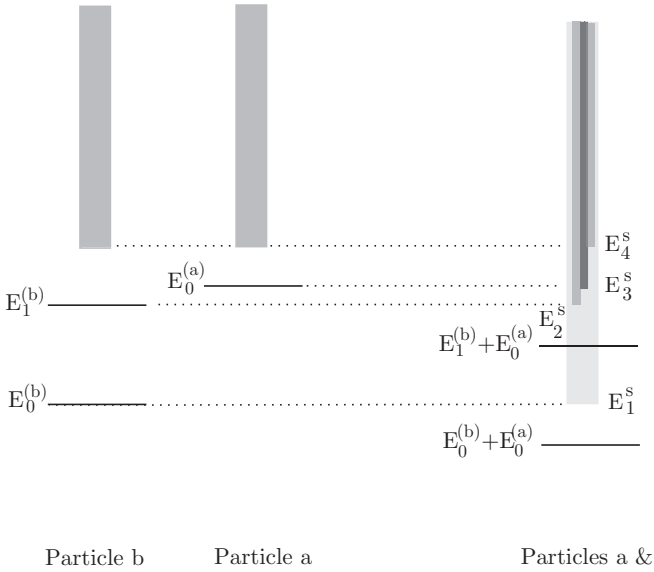

FIGURE 9.1. The left-hand figures show the separate spectra of the two particles, while the right-hand figure shows the two-particle spectrum for the non-interacting particles. The continuous spectrum is shown in gray; in the light gray energy region of the third figure, particle b is in its ground state and particle a is unbound. The energies E^s are the thresholds: in the region above E_2^s, b can be in its excited state $E_1^{(b)}$ and a unbound; above E_3^s, a can be bound and b free; above E_4^s both particles can be free.

You should look at this figure carefully and understand its implications. Whereas each of the one-particle systems has bound states of negative energy and a positive-energy continuous spectrum, which is separated from the bound states, the continuous spectrum of the two-particle system begins at E_1^s, the lowest energy of the two ground states, in this case the ground state of particle b. The physical reason for this is that above that energy particle a can escape to infinity: the atom can be ionized. The energy E_2^s is the threshold above which particle a can be free while b is in its excited state $E_1^{(b)}$; above this energy there could be inelastic scattering of a, with b initially in its ground state and finally in the excited state $E_1^{(b)}$, or vice versa, if there were an interaction that would make such a

transition go. Above E_3^s there could be exchange reactions, in which b collides with a bound and a emerges with b bound, or vice versa, if there were an interaction, and above E_4^s both a and b can be free, so that, with interaction, either particle could collide with the other one bound, ending up with both free (double ionizaton), or one of them could be captured and remain bound.

A particular feature to be noted is that an excited state of the combined system is embedded in the continuous spectrum (degenerate with a quasi-eigenvalue): the system can exist, at the same energy, both in a bound state and in a dissociated state, i.e., ionized, because the energy of a state in which particle b is in its second level and a is bound (so both are bound) is equal to the energy of a state in which b is in its lowest level and a has been expelled and is free.

Imagine now slowly cranking on the interaction $V^{(ab)}$ (say, multiplying $V^{(ab)}$ by a coupling parameter λ and allowing λ to increase from zero). Such a perturbation will, of course, shift the discrete energy levels, but it will also generally allow the bound state in which b is in its excited state and a in its ground state to "leak" and make a transition to the dissociated state in which b in its ground state and a free, thus destroying the stability of the original bound state and giving it a finite lifetime. For atoms, this is called the *Auger effect*, and for molecules, *predissociation.* A concrete example is that of a helium atom excited to its $(2s)^2$ state (both electrons in a 2s-state, with opposite spins, 81.6 eV above the ground state), which is degenerate with a state with one of the electrons in its 1s ground state and the other free (the ionization threshold is 54.4 eV above the 2s-state). The electron-electron interaction will therefore induce a *radiationless transition*, ionizing the atom and leaving it in its ground state.

So long as the interaction $V^{(ab)}$ is sufficiently weak, the Auger-transition probability W per unit time can be calculated by Fermi's golden rule, (8.8),

$$W = \frac{2\pi M k_1}{\hbar^3} \int d\Omega_f |\langle f|V^{(ab)}|i\rangle|^2, \tag{9.3}$$

where $k_1^2 \stackrel{\text{def}}{=} 2M(E - E_0^{(b)})/\hbar^2$, $|i\rangle$ is the initial state with both particles bound (the $(2s)^2$-state in the case of helium), and $\langle f|$ encompasses all the degenerate final states with b bound in its ground state (the 1s state for helium) and a free in the continuum but moving in various directions, and we must integrate over all these. [Equation (8.13) has been used here for the density of final states.] The lifetime of the unstable state is thus $\tau = 1/W$. (However, it is also possible for the bound state of the two particles to remain stably bound even when the interaction $V^{(ab)}$ is strong, remaining stable because the *exact* transition probability to the ionized state happens to be vanish; but this should be regarded as a more-or-less accidental situation.)

Since particle a is free at the energy in question, it can, of course, also come in and be scattered by the potential V_a. In the absence of the two-particle interaction $V^{(ab)}$, its scattering will be uninfluenced by the presence of particle b, but when $V^{(ab)}$ is turned on, the fact that the two-particle system can form an almost-bound state will manifest itself strongly as a resonance. There will be a sharp peak of width Γ in the squared magnitude of the scattering amplitude so that the scattering cross section has the characteristic shape

$$\sigma(E_2) \propto [(E_2 - R)^2 + \frac{1}{4}\Gamma^2]^{-1}, \tag{9.4}$$

shown in Figure 4.2, and Γ is related to the lifetime τ of the unstable state by the relation

$$\Gamma\tau = \hbar, \tag{9.5}$$

which we saw earlier. Moreover, the phase of the scattering amplitude at the resonance sharply *rises* by π, so as to produce a delay in the outgoing particles, just as it would classically if they emerged after *orbiting* the center for a while. Let's take a look at the mathematics of how this happens.

Suppose the two-particle wave function is expanded on the basis of products of one-particle wave functions, each of which solves the Schrödinger equation with its own potential, as in (9.2),

$$\psi(E, \vec{r}_a, \vec{r}_b) = \sum\!\!\!\!\!\!\int \psi^{(a)}(E^{(a)}, \vec{r}_a)\psi^{(b)}(E^{(b)}, \vec{r}_b), \tag{9.6}$$

where $E^{(a)} + E^{(b)} = E$ and the functions $\psi^{(b)}$ are assumed normalized, so that

$$\psi^{(a)}(E^{(a)}, \vec{r}_a) = \int d^3r_b\, \psi^{(b)*}(E^{(b)}, \vec{r}_b)\psi(E, \vec{r}_a, \vec{r}_b)$$

and $\psi^{(a)}(E^{(a)}, \vec{r}_a)$ satisfies the set of coupled Schrödinger equations[1]

$$\sum\!\!\!\!\!\!\int_\beta \mathbf{H}_{\alpha\beta}\psi^{(a)}_\beta = E_\alpha\psi^{(a)}_\alpha, \tag{9.7}$$

where

$$\mathbf{H}_{\alpha\beta} \stackrel{\text{def}}{=} \left(-\frac{\hbar^2}{2M}\nabla^2 + V_a(\vec{r})\right)\delta_{\alpha\beta} + V^{(ab)}_{\alpha\beta}(\vec{r}), \tag{9.8}$$

$$V^{(ab)}_{\alpha\beta}(\vec{r}) \stackrel{\text{def}}{=} \int d^3r_b\, \psi^{(b)}_\alpha(\vec{r}_b)V^{(ab)}(\vec{r}_b - \vec{r})\psi^{(b)*}_\beta(\vec{r}_b), \tag{9.9}$$

and the indices α and β run partly over discrete and partly over continuous values.

[1] **Show it.**

In order to deal with a manageable model system, let us truncate this infinite set of coupled equations by assuming that particle b has only two possible states (thereby removing the possiblity of exchange reactions in which initially particle b is bound and a free, while finally a is bound and b free), and write (9.7) in matrix notation,

$$(\mathbf{T} + \mathcal{V})\Psi = \mathcal{E}\Psi, \tag{9.10}$$

where Ψ has two components, the first being the wave function of particle a for the case in which b is in its ground state of energy $E_0^{(b)}$ and hence the energy of a is $E_1 \stackrel{\text{def}}{=} E - E_0^{(b)}$, and the second component being the wave fucntion of a when b is in its second state of energy $E_1^{(b)}$ and hence that of particle a is $E_2 \stackrel{\text{def}}{=} E - E_1^{(b)}$; $\mathcal{E}$ is the diagonal matrix with E_1 and E_2 on its diagonal. The two states are referred to as the two *channels* of the system. Since the bound-state wave functions $\psi_1^{(b)}$ and $\psi_2^{(b)}$ can be chosen real, the matrix $\mathcal{V}$ is real and symmetric.

To calculate the scattering of particle a we proceed just as in the single-channel case considered in Chapter 4: we have to solve the Lippmann-Schwinger equation

$$\Psi^+ = \Psi_0 + \mathbf{G}_0^+ \mathcal{V} \Psi^+, \tag{9.11}$$

in which $\mathbf{G}_0^+$ is the diagonal (in channel space) matrix that solves the equation

$$(\mathcal{E} - \mathbf{T})\mathbf{G}_0^+ = \mathbb{1} \tag{9.12}$$

with the usual outgoing-wave boundary condition. (Here $\mathbb{1}$ is both the unit matrix in channel space and the unit operator.) The two components of Ψ_0 depend on the initial channel: if initially particle b is in its ground state while particle a comes in, the first component of Ψ_0 is the plane wave $(2\pi)^{-3/2} e^{i\vec{k}_1 \cdot \vec{r}}$, where $k_1^2 = |\vec{k}_1|^2 \stackrel{\text{def}}{=} 2ME_1/\hbar^2$, and the second component equals zero; if initially b is in its second state, the first component of Ψ_0 vanishes and its second component equals $(2\pi)^{-3/2} e^{i\vec{k}_2 \cdot \vec{r}}$, where $k_2^2 \stackrel{\text{def}}{=} |\vec{k}_2|^2 = 2ME_2/\hbar^2$. These two column matrices can be arranged in a 2×2 matrix, with the free solution written as $(2\pi)^{-3/2} e^{i\vec{K} \cdot \vec{r}}$, where $\vec{K}$ is a diagonal matrix with the diagonal elements $\vec{k}_1$ and $\vec{k}_2$. In terms of the matrix K, which has k_1 and k_2 on the diagonal, $K \stackrel{\text{def}}{=} \sqrt{2M\mathcal{E}}/\hbar$, the free Green's function is expressible in the form

$$G_0^+(E, \vec{r}, \vec{r}\,') = -\frac{M}{2\pi\hbar^2} \frac{e^{iK|\vec{r} - \vec{r}\,'|}}{|\vec{r} - \vec{r}\,'|},$$

from which we obtain the asymptotic form of Ψ^+

$$\Psi^+(E, \vec{r}) \sim (2\pi)^{-3/2} [e^{i\vec{K} \cdot \vec{r}} + \frac{1}{r} e^{iKr} A(E; \widehat{k}_f, \widehat{k}_i)],$$

where $\widehat{k}_i$ and $\widehat{k}_f$ are the initial and final momentum directions, and A is the 2×2 matrix

$$A(E; \widehat{k}_f, \widehat{k}_i) = -\frac{M}{2\pi\hbar^2}\Theta(f,i), \tag{9.13}$$

$$\Theta(f,i) \stackrel{\text{def}}{=} (\Psi_0^{(f)}, \mathcal{V}\Psi^{(i)+}) \stackrel{\text{def}}{=} \int d^3r\, \Psi_0^*(\vec{K}_f, \vec{r})\mathcal{V}(\vec{r})\Psi^+(\vec{K}_i, \vec{r}). \tag{9.14}$$

The scattering cross section from the initial channel i and direction $\widehat{k}_i$, to the final channel f and direction $\widehat{k}_f$ is given by[2]

$$\frac{d\sigma(f, \widehat{k}_f, i, \widehat{k}_i)}{d\Omega} = \frac{k_f}{k_i}|A_{fi}(E; \widehat{k}_f, \widehat{k}_i)|^2. \tag{9.15}$$

Specification of a channel amounts to specifying in which of its two states particle b is and at the same time determining the values of k_f and k_i. If the two states are the same, the scattering is elastic; otherwise it is inelastic, with different inital and final energies of particle a as well as of particle b but with the total energy conserved.

Just as in Section 4.4, it is convenient to define a complete Green's function by the equation

$$(\mathcal{E} - \mathbf{T} - \mathcal{V})\mathcal{G}^+ = \mathbb{1} \tag{9.16}$$

and the outgoing-wave boundary condition, and solve (9.11) by writing

$$\Psi^+ = \Psi_0 + \mathcal{G}^+\mathcal{V}\Psi_0,$$

so that $\Theta(fi)$ has the form

$$\Theta(f,i) = (\Psi_0^{(f)}, \mathcal{V}\Psi_0^{(i)}) + (\Psi_0^{(f)}, \mathcal{V}\mathcal{G}^+\mathcal{V}\Psi_0^{(i)}). \tag{9.17}$$

Let us also define the diagonal matrix Green's function G^+ whose elements g_1^+ and g_2^+ satisfy

$$[E_1 - \mathbf{T} - v_1]g_1^+ = \mathbb{1}, \qquad [E_2 - \mathbf{T} - v_2]g_2^+ = \mathbb{1}, \tag{9.18}$$

and the outgoing-wave boundary condition, where

$$v_1(\vec{r}) \stackrel{\text{def}}{=} V_a(\vec{r}) + V_{11}^{(ab)}(\vec{r}), \qquad v_2(\vec{r}) \stackrel{\text{def}}{=} V_a(\vec{r}) + V_{22}^{(ab)}(\vec{r}). \tag{9.19}$$

We are particularly interested in the case in which channel 1 is open, i.e., particle a has enough energy to be in its continuous spectrum, while channel 2 is closed so that it is below its continuum (the energy region between E_1^s and E_2^s of Figure 9.1).

Multiplying (9.16) by G^+ on the left, we obtain the equation for the full Green's function, $(\mathbb{1} - G^+\mathcal{V}^{\text{int}})\mathcal{G}^+ = G^+$, or,

$$\mathcal{G}^+ = G^+ + G^+\mathcal{V}^{\text{int}}\mathcal{G}^+, \tag{9.20}$$

---

[2] **Where does the factor k_f/k_i come from?**

where

$$\mathcal{V}^{\rm int} \stackrel{\rm def}{=} V_{12}^{(ab)} Q$$

in terms of the matrix

$$Q \stackrel{\rm def}{=} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Also define

$$P_1 \stackrel{\rm def}{=} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad P_2 \stackrel{\rm def}{=} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Since P_1 commutes with G^+ and $P_1 Q = Q P_2$, multiplication of (9.20) by P_1 and by P_2 leads to

$$P_1 \mathcal{G}^+ = g_1^+ P_1 + g_1^+ \mathcal{V}^{\rm int} P_2 \mathcal{G}^+,$$

$$P_2 \mathcal{G}^+ = g_2^+ P_2 + g_2^+ \mathcal{V}^{\rm int} P_1 \mathcal{G}^+,$$

and substitution of the first equation into the second, to

$$P_2 \mathcal{G}^+ = g_2^+ P_2 + g_2^+ \mathcal{V}^{\rm int} g_1^+ P_1 + g_2^+ V_{12}^{(ab)} g_1^+ V_{12}^{(ab)} P_2 \mathcal{G}^+,$$

or,

$$(E_2 - \mathbf{T} - v_2 - V_{12}^{(ab)} g_1 V_{12}^{(ab)}) P_2 \mathcal{G}^+ = P_2 + \mathcal{V}^{\rm int} g_1^+ P_1, \tag{9.21}$$

implying that $\mathcal{G}_2^+ \stackrel{\rm def}{=} P_2 \mathcal{G}^+ P_2$ is a Green's function that satisfies the equation

$$[E_2 - \mathfrak{H}(E_1)] \mathcal{G}_2^+ = P_2, \tag{9.22}$$

with the nonlocal, energy-dependent, non-Hermitian, effective Hamiltonian

$$\begin{aligned} \mathfrak{H}(E_1) &\stackrel{\rm def}{=} \mathbf{T} + v_2 + V_{12}^{(ab)} g_1(E_1) V_{12}^{(ab)} \\ &= \mathbf{T} + v_2 + V_{12}^{(ab)} \mathcal{P}(E_1 - \mathbf{H}_1)^{-1} V_{12}^{(ab)} \\ &\quad - i\pi V_{12}^{(ab)} \delta(E_1 - \mathbf{H}_1) V_{12}^{(ab)}, \end{aligned} \tag{9.23}$$

in which use has been made of (A.2) as well as of the symbolic expression in (4.127), $\mathcal{P}$ stands for Cauchy's principal value, and $\mathbf{H}_1 \stackrel{\rm def}{=} \mathbf{T} + v_1$. The intuitive physical interpretation of this effective Hamiltonian seen by particle a when it is in channel 2 is that in addition to seeing the ordinary potential v_2, it sees the effect of the interaction $V^{(ab)}$ causing a transition to channel 1, where it can propagate freely (subject only to the potential v_1), followed by a transition back to channel 2.

The operator $\mathcal{G}_2 \stackrel{\rm def}{=} [E_2 - \mathfrak{H}(E_1)]^{-1} P_2$ will have a pole at an eigenvalue $A(E_1^{(B)})$ of $\mathfrak{H}$,

$$\mathfrak{H}(E_1^{(B)}) u = A(E_1^{(B)}) u,$$

which in the absence of $V^{(ab)}$ would simply be the eigenvalue $E_0^{(a)}$ of $\mathbf{T}+V_2$. But (9.23) and (9.19) show that the presence of $V^{(ab)}$ shifts it not only along the real axis but generally into the complex plane,

$$A(E_1^{(B)}) \stackrel{\text{def}}{=} R - \frac{1}{2}i\Gamma,$$

which is the reason why the state becomes unstable, as we have seen. When $V^{(ab)}$ is small, therefore, $u \simeq \psi^{(a)}$ as well as $\Gamma \ll R$. For real energies in the vicinity of $E_2 = R$ (9.17) shows that the elastic scattering amplitude of particle a in channel 1 contains a rapidly varying term that, as a function of the energy, behaves like

$$\frac{(\Psi_0(\vec{k}_1'), V^{(ab)}u)(u, V^{(ab)}\Psi_0(\vec{k}_1))}{[E_2 - R + \frac{1}{2}i\Gamma]}, \tag{9.24}$$

producing a sharp peak of width Γ in the magnitude of the amplitude and in the elastic scattering cross section that has the characteristic shape (9.4) shown in Figure 4.2. From (9.23) and a comparison of (4.127) with (4.128) we calculate the value of Γ for weak $V^{(ab)}$ to be

$$\begin{aligned}
\Gamma &= 2\pi(\psi^{(a)}(E_0^{(a)}), V_{12}^{(ab)}\delta(E_1 - \mathbf{H}_1)V_{12}^{(ab)}\psi(E_0^{(a)})) \\
&= \frac{2\pi M k_1}{\hbar^2}\int d\Omega_1 |(\psi^{(a)}(E_0^{(a)}), V_{12}^{(ab)}\psi^{(a)}(\vec{k}_1))|^2 \\
&= \frac{2\pi M k_1}{\hbar^2}\int d\Omega_1 \left| \int d^3r_a d^3r_b\, \psi^{(a)*}(E_0^{(a)}, \vec{r}_a)\psi^{(b)*}(E_1^{(b)}, \vec{r}_b) \right. \\
&\qquad \left. \times V^{(ab)}(\vec{r}_a - \vec{r}_b)\psi^{(a)}(\vec{k}_1, \vec{r}_a)\psi^{(b)}(E_0^{(b)}, \vec{r}_b) \right|^2 .
\end{aligned}$$

Comparison with (9.3) shows that the width Γ of the resonance peak is related to the lifetime of the unstable state by (9.5). It follows from (9.24) that the phase of the scattering amplitude at the resonance sharply *rises* by π.[3]

9.2 Identical particles

Even in classical physics certain phenomena involving more then one particle depend on whether the particles are identical or distinguishable by their masses or other characteristics. For example, if two idential particles are scattered by one another and their emergence is counted by detectors, the numbers of particles #1 registering in detector A and of #2 in detector B has to be added to the number of particles #2 registering in A and of #1 in

[3] **Show this.**

B: the cross sections in which the final particles are and are not exchanged have to be added to arrive at the correct count.

In quantum mechanics, however, the situation is very much more serious. Classically, two particles with identical properties can still be distinguished by their history; they have a certain individuality: we can keep track of them by following them along their trajectories. In quantum mechanics, this cannot be done; any attempt to identify one along its track would alter the experiment. Since neither individuality nor history can be used for identification, the only possible marker of a particle is its *state,* specified by the measured values of a complete set of commuting observables. Two identical particles in the same state are fundamentally indistinguishable, in keeping with the assumption that the state of a system, as given by its state vector (in the ideal situation of a pure state), completely specifies all the properties of the system at one time. Moreover, in order to arrive at the correct scattering cross section in a situation such as described in the first paragraph above, the *scattering amplitudes* have to be added rather than the cross sections, which results in additional interference effects that depend on the relative phases of the two amplitudes.

Consider, then, a wave function of N identical particles $\psi(1, 2, \ldots, N; t)$ (where the numbers indicate the particles in some specified states), and let $\mathcal{E}_{12}$ be the operator that exchanges particles 1 and 2, so that

$$\mathcal{E}_{12}\psi(1, 2, \ldots, N; t) \stackrel{\text{def}}{=} \psi(2, 1, \ldots, N; t).$$

Since, in order to ensure that experimental results cannot distinguish between particles, all expectation values have to be invariant under the exchange, we must have $|\mathcal{E}_{12}\psi|^2 = |\psi|^2$, and hence $\mathcal{E}_{12}\psi = e^{i\varphi}\psi$ with φ real: ψ must be an eigenfunction of $\mathcal{E}_{12}$ with an eigenvalue of modulus 1. What is more, if we want single-valuedness of the wave function, we ought to have $\mathcal{E}_{12}^2 = \mathbf{1}$. The eigenvalues of $\mathcal{E}_{12}$ must therefore be ± 1, and all physically acceptable wave functions ψ of these particles have to be eigenfunctions of $\mathcal{E}_{12}$.

Next, take exchanges between identical particles 2 and 3 and suppose that

$$\mathcal{E}_{12}\psi = \psi \qquad \text{and} \qquad \mathcal{E}_{23}\psi = -\psi.$$

But[4] $\mathcal{E}_{13} = \mathcal{E}_{12}\mathcal{E}_{23}\mathcal{E}_{12}$ and therefore $\mathcal{E}_{13}\psi = -\psi$, while also $\mathcal{E}_{13} = \mathcal{E}_{23}\mathcal{E}_{12}\mathcal{E}_{23}$ and therefore $\mathcal{E}_{13}\psi = \psi$. Consequently we have to conclude that there have to be two classes of identical particles: those whose wave functions are symmetric under exchange of any two of them, and those whose wave functions are anti-symmetric; no wave function of N identical particles can be symmetric under the exchange of some of them and anti-symmetric under the exchange of others. Note that the exchange operators $\mathcal{E}_{12}$ and $\mathcal{E}_{23}$ don't

[4] **Check this.**

commute, so there is no complete set of common eigenfunctions, but there are eigenfunctions that are symmetric under all exchanges, and others that are anti-symmetric, and these two classes are the only ones physically acceptable. Therefore, *the Hilbert space of several identical particles contains vectors that do not correspond to physically realizable states.*

There is, however, the following **caveat:** The argument that the eigenvalues $p = e^{i\varphi}$ of the exchange operator $\mathcal{E}_{12}$ have to be ± 1 is based on the assumption that when two particles are physically exchanged by transporting them along some path, and the process is repeated along another path, the wave function has to return to its original value, else it would not be single-valued. For purely topological reasons, this reasoning is correct in three dimensions, where two exchanges are topologically equivalent to none, even if the exclusion principle (or a hard core repulsion) is taken into account and the particles are prevented from occupying the same point. However, in two dimensions (or in a multiply connected three-dimensional manifold), the argument breaks down: no number of repeated exchanges is necessarily topologically equivalent to zero exchanges if coincidence is excluded. The reason is the same as that discussed in Section 7.2 in the context of the Aharanov-Bohm effect. As a result, identical particles confined to a two-dimensional configuration space need not be members of one of the only two classes allowed in three dimensions; their wave functions can have more complicated properties under exchange than simple symmetry or anti-symmetry: under exchange their wave functions may change by arbitrary phase factors. Such particles, which have come to be known as **anyons,** play significant roles in the physics of the solid state, where two-dimensional films and surface structures come into play. We shall not pursue this topic further in this book.[5]

If a wave function is symmetric or anti-symmetric at one time, it will retain this property for all time. This is because $\psi(t) = \exp(-i\mathbf{H}t/\hbar)\psi(0)$, and we must have $[\mathcal{E}_{ij}, \mathbf{H}] = 0$, or else the particles would be dynamically distinguishable. Therefore if $\mathcal{E}_{ij}\psi(0) = \pm\psi(0)$, then

$$\mathcal{E}_{ij}\psi(t) = \exp(-i\mathbf{H}t/\hbar)\mathcal{E}_{ij}\psi(0) = \pm\exp(-i\mathbf{H}t/\hbar)\psi(0) = \pm\psi(t).$$

The next question is how to construct such totally symmetric or anti-symmetric wave functions. Suppose that $\psi(1, 2, \ldots, N)$ is an arbitrary wave function of N identical particles. In order to construct a symmetric function, we simply add all the wave functions obtained by exchanging pairs of particles:

$$\psi_{\text{sy}}(1, 2, \ldots, N) \stackrel{\text{def}}{=} \sum_{\text{perm}} \psi(i, j, \ldots, N),$$

[5]For a good survey, see [Canright]; see also [Shapere].

where the sum runs over all permutations. Similarly, an anti-symmetric function is constructed by

$$\psi_{\text{antisy}}(1,2,\ldots,N) \stackrel{\text{def}}{=} \sum_{\text{perm}} \varepsilon_{ij\ldots}\psi(i,j,\ldots,N),$$

where $\varepsilon_{ij\ldots}$ is the sign of the permutation; that is, $\varepsilon_{ij\ldots} = 1$ if $i,j,\ldots$ is an even permutation of $1,2,\ldots$ (it takes an even number of exchanges to get from one to the other), and $\varepsilon_{ij\ldots} = -1$ if $i,j,\ldots$ is an odd permutation of $1,2,\ldots$ (it takes an odd number of exchanges).[6]

As an example take the case of N non-interacting identical particles. Any eigenfunction of the Hamiltonian $\mathbf{H} = \sum_{j=1}^{N} \mathbf{H}_j$ of such a system can be written as a linear combination of product wave functions, each of which are eigenfunctions of identical individual Hamiltonians $\mathbf{H}_j$:

$$\psi = \psi_{n_1}(1)\psi_{n_2}(2)\cdots\psi_{n_N}(N),$$

where

$$\mathbf{H}_j\psi_{n_j} = E_{n_j}\psi_{n_j}, \quad \text{and} \quad E = E_{n_1} + E_{n_2} + \ldots + E_{n_N}.$$

Such functions, however, are generally neither symmetric nor anti-symmetric. Since there clearly is exchange degeneracy (unless all the E_{n_i} are equal)—that is, the total energy of the system is invariant under exchange—we can take linear combinations of such product wave functions with the same total energy and still have an energy eigenfunction. So to form a symmetric energy eigenfunction we simply take

$$\psi_{\text{sy}}(1,2,\ldots,N) \stackrel{\text{def}}{=} \frac{1}{\sqrt{N!}} \sum_{\text{perm}} \psi_{n_1}(i_1)\psi_{n_2}(i_2)\cdots\psi_{n_N}(i_N), \tag{9.25}$$

and the exchange degeneracy is thereby generally removed.

A simple way to form an anti-symmetric wave function is to take the *Slater determinant*,[7]

$$\psi_{\text{antisy}}(1,2,\ldots,N) \stackrel{\text{def}}{=} \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{n_1}(1) & \psi_{n_1}(2) & \ldots & \psi_{n_1}(N) \\ \psi_{n_2}(1) & \psi_{n_2}(2) & \ldots & \psi_{n_2}(N) \\ \vdots & \vdots & & \vdots \\ \psi_{n_N}(1) & \psi_{n_N}(2) & \ldots & \psi_{n_N}(N) \end{vmatrix}. \tag{9.26}$$

If the one-particle wave functions are all ortho-normal and each state is occupied by one particle, the wave functions defined by (9.25) and (9.26)

[6] **You should check that any exhange of two particles in ψ_{antisy} as defined above changes the sign of that wave function.**

[7] **Why is this wave function anti-symmetric?**

are properly normalized. To see this, take the square of the norm of the symmetric wave function,

$$(\psi_{\rm sy},\psi_{\rm sy}) = \frac{1}{N!}\sum_{\rm perm}(\psi_{n_1}(1),\psi_{n_1}(1))\,(\psi_{n_2}(2),\psi_{n_2}(2))\cdots = \frac{1}{N!}\sum_{\rm perm}1 = 1.$$

(As an **exercise,** I will ask you to calculate the proper normalization factor if some levels are occupied by more than one particle.) The argument for the anti-symmetric case is the same as for the symmetric case.

If two particles described by an anti-symmetric wave function are in the same state, the wave function, of course, has to vanish. Therefore such particles obey the *Pauli exclusion principle:* no state can accommodate more than one of them. The distinction between the two kinds of particles that exist in nature can therefore be put most simply by saying: One kind is subject to the Pauli principle and the other kind is not. This distinction has important consequences in statistical mechanics. When in thermodynamic equilibrium at the temperature T, the energies E_n of fundamentally indistinguishable particles not subject to the Pauli principle are distributed according to *Bose-Einstein* statistics:

$$P(E_n) \propto \frac{1}{e^{(E_n-\mu)/kT}-1}, \tag{9.27}$$

where k is Boltzmann's constant and μ is the so-called chemical potential. [For a derivation, see (9.47) below.] On the other hand, if they are subject to the exclusion principle, they obey *Fermi-Dirac* statistics,

$$P(E_n) \propto \frac{1}{e^{(E_n-\mu)/kT}+1}. \tag{9.28}$$

[For a derivation, see (9.48).] The first are therefore called *bosons* and the second, *fermion.* (Now you may appreciate why the other kinds of particles possible in two-dimensional or multiply connected three-dimensional structures are called *anyons.*) Inspection of the two distributions shows that both approach Maxwell-Boltzmann statistics for high energies, but for low energies (compared to kT) they deviate markedly from $e^{-E_n/kT}$ as well as from one another.

Of the known particles, electrons, protons, neutrons, muons, and neutrinos are examples of fermions; photons and pions are bosons. It is one of the most fundamental facts of physics that *all fermions have half (odd) integral spin, while all bosons have integral spin.* The ultimate explanation of this *spin-statistics connection* rests on the theory of relativity and requires tools that go beyond the scope of this book;[8] in Chapter 10 we shall see a partial explanation. The spin-statistics connection holds not only for

[8]For a good review, see [Duck].

elementary particles but also for compound systems such as atoms and nuclei: those of integral spin act like bosons, and those of half-integral spin like fermions. For example, the fact that C^{12}-nuclei are bosons manifests itself in their mutual scattering by a constructive interference peak at 90^0 in their center-of-mass system because the two scattering amplitudes for direct and exchange scattering have to be *added*; were they fermions, the two amplitudes would have to be subtracted, which would result in destructive interference and a minimum.

Perhaps the most dramatic manifestation of the fact that even bound systems of integral spin behave like bosons is their observed behavior at extremely low temperatures. As we shall see in the next section, bosons are, in fact, actually attracted into states that are already heavily occupied, the result of which is called **Bose-Einstein condensation**. The atoms of liquid ^{4}He all congregate in the state of lowest energy when cooled to near 0^0 Kelvin, whereas those of ^{3}He behave quite differently; as is appropriate for them as fermion they must all occupy different states and thus fill the lowest energy levels as high as necessary, the top level being called the *Fermi surface.*

A gas that is a Bose-Einstein condensate, whose existence was predicted by Satyendranath Bose and Albert Einstein in 1925, is not to be confused with an ordinary collection of atoms or molecules all of which are in their lowest energy state. In the condensate, the N-particle state is *pure*, and the entire system consisting of many atoms is describable by a wave function in its $3N$-dimensional configuration space rather than by a density matrix: the states of all the atoms are phase-correlated. Such a state of a gas at ultra-low temperature was first observed experimentally only a few years ago.

9.2.1 Fock space

In view of the essential indistinguishability of identical particles, a way of labeling states of more than one of them that is more convenient than trying to specify which particle is in which state, is to use the occupation-number representation already introduced in Section 1.4.3, which is a natural expression of the view that particles are no more than manifestations of an underlying quantum field. Here one defines a standard order of the states (which requires infinitely many tags), for example, by increasing energy and some fixed conventional order for degenerate states, and labels the states of the system by identifying which of them are occupied and by how many particles, as in

$$\Psi(0,1,0,0,2,0,\ldots) \stackrel{\text{def}}{=} \Psi([n]) \stackrel{\text{def}}{=} |[n]\rangle,$$

where $[n]$ denotes the set of occupation numbers in the standard order. This is clearly a more appropriate manner of dealing with the fact that

the particles have no identity than attempting to specify which of them is where. It is called the *occupation-number representation* and the Hilbert space whose basis vectors are so labeled is also called *Fock space*, after the Russian physicist V. Fock. Whereas for bosons the occupation numbers are unrestricted, for fermions, of course, none of them can exceed 1.

As we saw in Section 1.4.3, the proper mathematical tools for the description of these states are the annihilation and creation operators, which are conventionally written as $\mathbf{a}_i$ and $\mathbf{a}_i^\dagger$, respectively (these are the same as the $\widetilde{\Psi}_i$ and $\widetilde{\Psi}_i^\dagger$ in Section 1.4.3). For bosons they satisfy the commutation relations

$$[\mathbf{a}_i, \mathbf{a}_j^\dagger] = \delta_{ij}, \qquad [\mathbf{a}_i, \mathbf{a}_j] = 0. \tag{9.29}$$

In terms of the occupation-number operator for state i, $\mathsf{N}_i \stackrel{\text{def}}{=} \mathbf{a}_i^\dagger \mathbf{a}_i$, we then have

$$\mathsf{N_i}|[n]\rangle = n_i|[n]\rangle. \tag{9.30}$$

The total number operator,

$$\mathsf{N} \stackrel{\text{def}}{=} \sum_i \mathsf{N}_i = \sum_i \mathbf{a}_i^\dagger \mathbf{a}_i, \tag{9.31}$$

has as its eigenstates all the states of a system containing a fixed, given number of particles, and its eigenvalues are these total numbers of particles. (The *system* is now a field rather than a collection of particles.)

As we saw,

$$\mathbf{a}_i^\dagger|\ldots, n_i, \ldots\rangle = \sqrt{n_i + 1}|\ldots, n_i + 1, \ldots\rangle,$$

which can also be written as

$$\mathbf{a}_i^\dagger|[n]\rangle = \sqrt{n_i + 1}|[n] + 1_i\rangle, \tag{9.32}$$

so that

$$|[n]\rangle = \frac{\cdots (\mathbf{a}_2^\dagger)^{n_2} (\mathbf{a}_1^\dagger)^{n_1}}{\sqrt{n_1! n_2! \cdots}}|0\rangle = \frac{(\mathbf{a}_1^{n_1} \mathbf{a}_2^{n_2} \cdots)^\dagger}{\sqrt{n_1! n_2! \cdots}}|0\rangle, \tag{9.33}$$

as in (1.43) and (1.44).

For fermions, we have to be a little more careful about the order of these operators because they anti-commute:

$$\{\mathbf{a}_i, \mathbf{a}_j^\dagger\} = \delta_{ij}, \qquad \{\mathbf{a}_i, \mathbf{a}_j\} = 0. \tag{9.34}$$

As we saw in Section 1.4.3 these anti-commutation relations lead to the Pauli exclusion principle, so that only occupation numbers 0 and 1 are allowed. In order to end up with consistent signs, the action of $\mathbf{a}_i^\dagger$ has to include a sign factor such as

$$\mathbf{a}_i^\dagger|[n]\rangle = (-1)^p(1 - n_i)|[n] + 1_i\rangle, \tag{9.35}$$

where p is the number of occupied states in $[n]$ higher than i. (This sign then of course depends on the order of states adopted by convention, but it will remain consistent once this order is fixed.) Let us check that this works correctly:

$$\begin{aligned} \mathbf{a}_1^\dagger|0\rangle &= |1,0,0,\ldots\rangle \\ \mathbf{a}_2^\dagger|1,0,0,\ldots\rangle &= |1,1,0,0,\ldots\rangle = -\mathbf{a}_1^\dagger|0,1,0,0,\ldots\rangle = -\mathbf{a}_1^\dagger\mathbf{a}_2^\dagger|0\rangle \\ &= \mathbf{a}_2^\dagger\mathbf{a}_1^\dagger|0\rangle = \mathbf{a}_2^\dagger|1,0,0,\ldots\rangle. \end{aligned}$$

Had the factor $(-1)^p$ (or some other conventional sign factor) not been present in (9.35), this would not have come out consistent.[9]

The state vectors of one-particle systems in which the particle occupies the state i are $\Psi_i = |i\rangle_1 = \mathbf{a}_i^\dagger|0\rangle$ and the completeness of the states used as a basis implies that $\sum_i |i\rangle_{1\,1}\langle|i| = \mathbb{1}_1$, or

$$\sum_i \mathbf{a}_i^\dagger|0\rangle\langle 0|\mathbf{a}_i = \mathsf{P}_1, \tag{9.36}$$

where P_1 is the projection on the subspace of all one-particle states in the Fock space. Both P_1 and N should, of course, be independent of the basis of states used in the definition, and they are.[10] This includes even quasi-bases such as the configuration basis, which was employed in Section 1.4.3. Thus,

$$\rho(\vec{q}) \stackrel{\text{def}}{=} \Psi^\dagger(\vec{q})\Psi(\vec{q})$$

is the density of particles in configuration space and

$$\mathsf{N} = \int d^3q\, \Psi^\dagger(\vec{q})\Psi(\vec{q})$$

their total number; the projection on the one-particle subspace is given by

$$\mathsf{P}_1 = \int d^3q\, \Psi^\dagger(\vec{q})|0\rangle\langle 0|\Psi(\vec{q}).$$

In this language a general one-particle state is given by $|\,\rangle_1 = \Psi_1 = \sum_j c_j\mathbf{a}_j^\dagger|0\rangle$, so that the probability of finding the particle in the state i is given by

$$|\langle 1_i|\,\rangle_1|^2 = |\langle 1_i|\sum_j c_j\mathbf{a}_j^\dagger|0\rangle|^2 = |\langle 0|\sum_j c_j\mathbf{a}_i\mathbf{a}_j^\dagger|0\rangle|^2 = |c_i|^2,$$

where the commutation relations (9.29) have been used and the fact that $\mathbf{a}_i|0\rangle = 0$. To obtain a one-particle wave function, we proceed as in Section

[9] **Verify this.**
[10] **Prove this as an exercise.**

1.4.3. Using either (1.36) or (1.51), depending on whether the particles are bosons or fermions, and (1.48), with the result (1.49), and

$$\psi(\vec{q}) = \langle 0|\Psi(\vec{q})|\,\rangle_1 = \sum_j c_j \langle 0|\Psi(\vec{q})\mathbf{a}_j^\dagger|0\rangle = \sum_j c_j \psi_j(\vec{q}), \tag{9.37}$$

where $\psi_j(\vec{q})$ is the wave function of the state j.

If a new set of annihilation operators is introduced by a unitary transformation U in the sense that

$$\mathbf{a}'_i = \sum_j U_{ij}\mathbf{a}_j,$$

then for bosons

$$[\mathbf{a}'_i, \mathbf{a}_j^\dagger] = \sum_k U_{ik}[\mathbf{a}_k, \mathbf{a}_j^\dagger] = U_{ij}, \tag{9.38}$$

and therefore, since $\mathbf{a}_i'^\dagger = \sum_j U_{ij}^* \mathbf{a}_j^\dagger$,

$$|i\rangle'_1 = \mathbf{a}_i'^\dagger|0\rangle = \sum_j U_{ij}^* \mathbf{a}_j^\dagger|0\rangle = \sum_j U_{ij}^*|j\rangle_1,$$

which means that the matrix U_{ij} is the transformation function connecting the two one-particle bases $|i\rangle'_1$ and $|j\rangle_1$, $U_{ij} = (\Psi'_i, \Psi_j)$. If the new "basis" is a quasi-basis, for example the quasi-basis for the configuration representation, then $U_j(\vec{q}) = \langle\vec{q}|j\rangle = \psi_j(\vec{q})$, the normalized one-particle wave function of the j^{th} state, and we have by (9.38),

$$[\Psi(\vec{q}), \mathbf{a}_j^\dagger] = \psi_j(\vec{q}). \tag{9.39}$$

For fermions the same equations hold, with the commutators in (9.38) and (9.39) replaced by anti-commutators. The commutators $[\Psi(\vec{q}), \mathbf{a}_j]$, resp. anti-commutators $\{\Psi(\vec{q}), \mathbf{a}_j\}$, of course, all vanish.

The general two-boson wave function is obtained by repeated application of (9.39), remembering that whenever $\Psi(\vec{q})$ abuts $|0\rangle$ the result vanishes,[11]

$$\begin{aligned}
\psi(\vec{q}_1, \vec{q}_2) &= \langle 0|\Psi(\vec{q}_2)\Psi(\vec{q}_1)|\,\rangle_2 = \sum_{i\neq j} c_{ij}\langle 0|\Psi(\vec{q}_2)\Psi(\vec{q}_1)\mathbf{a}_i^\dagger\mathbf{a}_j^\dagger|0\rangle \\
&\qquad + \sum_j \frac{c_{jj}}{\sqrt{2}}\langle 0|\Psi(\vec{q}_2)\Psi(\vec{q}_1)\mathbf{a}_j^{\dagger 2}|0\rangle \\
&= \sum_{i\neq j} c_{ij}[\psi_i(\vec{q}_2)\psi_j(\vec{q}_1) + \psi_i(\vec{q}_1)\psi_j(\vec{q}_2)] + \sum_j c_{jj}\sqrt{2}\psi_j(\vec{q}_2)\psi_j(\vec{q}_1) \\
&= \sum_{ij} d_{ij}\psi_i(\vec{q}_1)\psi_j(\vec{q}_2),
\end{aligned} \tag{9.40}$$

[11] **Fill in the missing steps.**

where $d_{ij} \stackrel{\text{def}}{=} c_{ij} + c_{ji} = d_{ji}$ for $i \neq j$, and $d_{jj} \stackrel{\text{def}}{=} \sqrt{2}c_{jj}$. Thus $\psi(\vec{q}_1, \vec{q}_2)$ is automatically symmetric. The two-fermion wave function, on the other hand, is given by the anti-symmetric combination[12]

$$\psi(\vec{q}_1, \vec{q}_2) = \sum_{ij} c_{ij}[\psi_i(\vec{q}_2)\psi_j(\vec{q}_1) - \psi_i(\vec{q}_1)\psi_j(\vec{q}_2)] = \sum_{ij} d_{ij}\psi_i(\vec{q}_2)\psi_j(\vec{q}_1), \tag{9.41}$$

where $d_{ij} = c_{ij} - c_{ji} = -d_{ji}$.

For N particles, we have similarly, $|\,\rangle_N = \sum_{[n]} c_{[n]}|[n]\rangle$, and the n-particle wave function for bosons,

$$\begin{aligned} & \psi(\vec{q}_1, \vec{q}_2, \ldots) = \langle 0|\cdots\Psi(\vec{q}_2)\Psi(\vec{q}_1)|\,\rangle_N \\ = & \sum_{[n]} \frac{c_{[n]}}{\sqrt{n_1!n_2!\cdots}}\langle 0|\cdots\Psi(\vec{q}_2)\Psi(\vec{q}_1)\cdots\mathbf{a}_2^{\dagger n_2}\mathbf{a}_1^{\dagger n_1}|0\rangle, \end{aligned} \tag{9.42}$$

where the sum runs over all $[n]$ such that $n_1 + n_2 + \ldots = N$, and the right-hand side can be expressed as a linear combination of products of N one-particle wave functions by successively commuting, or anti-commuting, all the $\Psi(\vec{q})$ to the right. For bosons, the result is automatically symmetric, and for fermions anti-symmetric.

The Hamiltonian in Fock space is expressed in terms of the operators $\mathbf{a}_j$ and $\mathbf{a}_j^\dagger$. In the simplest instance,

$$\mathbf{H} = \sum_{kl} H_{kl}^{(1)}\mathbf{a}_k^\dagger\mathbf{a}_l. \tag{9.43}$$

Acting on a one-particle state, in which the particle is in the state Ψ_n, it results in

$$\mathbf{H}|1_n\rangle = \sum_{kl} H_{kl}^{(1)}\mathbf{a}_k^\dagger\mathbf{a}_l\mathbf{a}_n^\dagger|0\rangle = \sum_k H_{kn}^{(1)}\mathbf{a}_k^\dagger|0\rangle = \sum_k H_{kn}^{(1)}|1_k\rangle,$$

which identifies $H_{kl}^{(1)}$ as

$$H_{kl}^{(1)} = \langle 1_k|\mathbf{H}|1_l\rangle = (\Psi_k|\mathbf{H}|\Psi_l).$$

Acting on a two-particle state, the result is

$$\mathbf{H}|1_n, 1_m\rangle = \sum_{kl} H_{kl}^{(1)}\mathbf{a}_k^\dagger\mathbf{a}_l\mathbf{a}_n^\dagger\mathbf{a}_m^\dagger|0\rangle = \sum_k [H_{km}^{(1)}|1_k, 1_m\rangle \pm H_{kn}^{(1)}|1_k, 1_n\rangle],$$

the sign being + for bosons and − for fermions. Similarly for N-particle states. As you can see, this Hamiltonian acts on one particle at a time,

[12] **Verify this.**

taking it out of its original state and inserting it another, conserving the total number of particles,[13] $[\mathbf{H}, \mathsf{N}] = 0$.

For a system of identical non-interacting particles it is most convenient to use a representation that diagonalizes the one-particle Hamiltonian, so that the basis for the occupation number representation diagonalizes $H^{(1)}$,

$$\mathbf{H} = \sum_k \mathbf{a}_k^\dagger \mathbf{a}_k H_{kk}^{(1)} = \sum_k \mathsf{N}_k E_k,$$

E_k being the energy of a particle in the state k. The Fock states are eigenstates of this Hamiltonian:

$$\mathbf{H}|[n]\rangle = \sum_k E_k \mathsf{N}_k |[n]\rangle = \sum_k n_k E_k |[n]\rangle.$$

Let us now add a one-particle perturbation $\mathbf{H}'$, so that

$$\mathbf{H} = \sum_k \mathsf{N}_k E_k + \sum_{kl} \mathbf{a}_k^\dagger H'_{kl} \mathbf{a}_l.$$

This perturbation may, for example, be an electromagnetic wave producing a transition from $|[n]\rangle$ to $|[n']\rangle = |[n] + 1_k - 1_l\rangle$. In first-order perturbation theory, the transition probability is then proportional to $|\langle[n']|\mathbf{H}'|[n]\rangle|^2$, and for $k \neq l$ we have for fermion

$$\begin{aligned} \langle[n']|\mathbf{H}'|[n]\rangle &= \langle[n] + 1_k - 1_l|\mathbf{H}'|[n]\rangle = H'_{kl}\langle[n] + 1_k - 1_l|\mathbf{a}_k^\dagger \mathbf{a}_l|[n]\rangle \\ &= \pm(-1)^{p_l + p_k}(1 - n_k) n_l H'_{kl} \end{aligned}$$

so that

$$|\langle[n] + 1_k - 1_l|\mathbf{H}'|[n]\rangle|^2 = |H'_{kl}|^2 n_l (1 - n_k). \tag{9.44}$$

Suppose, on the other hand, that the particles are bosons. In that case we have $\langle[n] + 1_k - 1_l|\mathbf{a}_k^\dagger \mathbf{a}_l|[n]\rangle = \sqrt{n_l(n_k + 1)}$ for $k \neq l$ and as a result,

$$|\langle[n] + 1_k - 1_l|\mathbf{H}'|[n]\rangle|^2 = |H'_{kl}|^2 n_l (1 + n_k). \tag{9.45}$$

Thus both for fermions and for bosons, the transition probability is proportional to the number of particles in the initial state l before, times the number of particles in the final state k after, the transition. In the fermion case, if the final state is occupied, there can be no transition, a fact that may be interpreted as an effective repulsion. In the boson case, on the other hand, the transition probability is *enhanced* if the final state is occupied: there is, in effect, an *attraction,* as though bosons were gregarious and prefered to be in the same state.

---

[13] **Prove this.**

The next more complicated kind of Hamiltonian is of the form

$$\mathbf{H} = \frac{1}{2}\sum_{klrs} H^{(2)}_{kl,sr}\mathbf{a}^\dagger_k\mathbf{a}^\dagger_l\mathbf{a}_r\mathbf{a}_s, \tag{9.46}$$

which annihilates one-particle states and which, on others, describes the interaction of two particles by extracting them from their orginal states and inserting them into new ones. It also conserves the total number of particles.[14] (All operators that are products of the same number of creation as annihilation operators conserve the number of particles.)

For a system with one-particle and pairwise interactions (and again a representation is used that diagonalizes the one-particle Hamiltonian) the total Hamiltonian will be

$$\mathbf{H} = \sum_k \mathsf{N}_k E_k + \frac{1}{2}\sum_{klrs} H^{(2)}_{kl,sr}\mathbf{a}^\dagger_k\mathbf{a}^\dagger_l\mathbf{a}_r\mathbf{a}_s$$

and if there is no degeneracy, the first-order energy shift resulting from the interaction will be given by

$$\Delta E_{[n]} = \frac{1}{2}\sum_{klsr} H^{(2)}_{kl,sr}\langle[n]|\mathbf{a}^\dagger_k\mathbf{a}^\dagger_l\mathbf{a}_r\mathbf{a}_s|[n]\rangle.$$

If the particles are fermions, we find that[15]

$$\langle[n]|\mathbf{a}^\dagger_k\mathbf{a}^\dagger_l\mathbf{a}_r\mathbf{a}_s|[n]\rangle = n_k n_l(\delta_{rl}\delta_{sk} - \delta_{kr}\delta_{ls}).$$

Therefore,

$$\Delta E_{[n]} = \frac{1}{2}\sum_{k\neq l} n_k n_l[H^{(2)}_{kl,kl} - H^{(2)}_{kl,lk}],$$

where the second term on the right is called the *exchange-energy* term, which arises because the particles may be exchanged in the course of the interaction. For example, if the states k and l of two spin-1/2 particles simply differ by their spins, it is convenient to use the spin-exchange operator defined in (5.69), which leads to

$$\begin{aligned}\Delta E_{[n]} &= \frac{1}{2}\sum_{k\neq l} n_k n_l(\Psi_{kl}, [\mathbf{H}^{(2)} - \mathbf{H}^{(2)}\mathsf{P}_{\text{exch}}]\Psi_{kl}) \\ &= \frac{1}{2}\sum_{k\neq l} n_k n_l(\Psi_{kl}, \mathbf{H}^{(2)}\frac{1}{2}(\mathbb{1} - \vec{\sigma}^{(k)}\cdot\vec{\sigma}^{(l)})\Psi_{kl}).\end{aligned}$$

The exchange interaction of two spin-1/2 fermions can therefore in some circumstances be regarded as a spin-dependent force.

---

[14] **Prove this.**
[15] **Show this.**

For bosons, on the other hand, the corresponding result for the first-order energy shift is[16]

$$\Delta E_{[n]} = \frac{1}{2}\sum_k n_k(n_k-1)H^{(2)}_{kk,kk} + \frac{1}{2}\sum_{k\neq l} n_k n_l [H^{(2)}_{kl,kl} + H^{(2)}_{kl,lk}].$$

The first term represents the interaction of two particles in the same state k, $\frac{1}{2}n_k(n_k-1)$ being the number of pairs in the state, which in the fermion case, of course, vanishes.

9.2.2 Bose-Einstein and Fermi-Dirac distributions

As an application, consider an assembly of bosons in equilibrium by collisions with an atomic system (say, making up the walls of the container) at temperature T. The number of transitions from N_1 atoms at energy E_1 and n_m bosons of energy ε_m to N_2 atoms at energy E_2 and n_l bosons of energy ε_l will be, according to (9.45), $AN_1n_m(n_l+1)$ with some constant A, wheras the number of transitions going in the reverse direction is $AN_2n_l(n_m+1)$, so that at equilibrium

$$n_m(n_l+1)N_1 = n_l(n_m+1)N_2,$$

and the conservation of energy requires $E_1+\varepsilon_m = E_2+\varepsilon_l$. At equilibrium, the ratio N_1/N_2 has to follow the Boltzmann law,

$$\frac{N_1}{N_2} = \exp[-\frac{E_1-E_2}{kT}] = \exp[-\frac{\varepsilon_l-\varepsilon_m}{kT}].$$

The result is the equation

$$\frac{n_m}{n_m+1}\frac{n_l+1}{n_l} = \exp\frac{\varepsilon_l-\varepsilon_m}{kt},$$

or

$$\left(1+\frac{1}{n_l}\right)\exp[-\frac{\varepsilon_l}{kT}] = \left(1+\frac{1}{n_m}\right)\exp[-\frac{\varepsilon_m}{kT}] = C,$$

where C is a constant, independent of the energies of the bosons. Therefore

$$n_m = \frac{1}{C\exp(\varepsilon_m/kT)-1}, \tag{9.47}$$

and C is determined by the constraint

$$N = \sum_m n_m = \sum_m \frac{1}{C\exp(\varepsilon_m/kT)-1}.$$

[16] **Show this.**

Eq. (9.47) is the Planck distribution law for bosons, or *Bose-Einstein statistics.*

For fermions (9.44) has to be used instead of (9.45) and the result is[17] the *Fermi-Dirac* statistical distribution

$$n_m = \frac{1}{C \exp(\varepsilon_m / kT) + 1}, \tag{9.48}$$

where, again, C has to be calculated from the constraint that the total number of particles is fixed.

[17] **Check this.**

9.3 Problems and Exercises

1. Use Fermi's golden rule (8.8) to calculate the Auger transition probability per unit time for a helium atom from its $(2s)^2$ state. (Enclose the atom in a large cubic box to discretize the continuous spectrum.)

2. Suppose that state n_1 of a given system is occupied by m_1 identical, non-interacting particles, state n_2 by m_2, etc. Express the properly normalized multiparticle wave function in terms of the single-particle functions.

3. Suppose two spin-1/2 particles are in a state in which the total angular momentum is a good quantum number; if they are in a singlet state, what is their lowest possible orbital angular momentum? What if they are in a triplet state?

4. Suppose that a new ortho-normal basis is used for the definition of occupation numbers in Fock space, $\Psi'_i = \sum_j U_{ij}\Psi_j$. How are the new annihilation operators $\mathbf{a}'_i$ related to the old ones? Show that the total number operator N and the projection operator on one-particle states P_1 are invariant under such a change in basis.

5. Let two identical spin-1/2 particles interact via a spin-independent attractive central potential strong enough to bind.

   (a) What are the values of the total angular momentum j and of the parity in the ground state?

   (b) For a given value of j and a given parity, how many possible states are there? Disregard the different orientations (i.e., m values) and make a list of the possible states (i.e., their l values and their total spin values) in each of the four relevant kinds of cases.

6. Consider a helium atom in the simplest approximation, in which the electrons don't interact with one another. If the electrons had spin 1, and hence obeyed Bose-Einstein statistics, what would be their total spin in the ground state?

7. Two identical spin-0 particles are free to move along a line; they do not interact with one another, but each is subject to a harmonic force from the origin that corresponds to the classical frequency ω. The two-particle system is in an eigenstate of the total energy with the eigenvalue $2\hbar\omega$.

   (a) Write down the density matrix of the system.

   (b) Calculate both the expectation value and the root-mean-square deviation from that expectation of the energy of one of the particles alone, with the other one ignored.

10
Relativistic Electron Theory

10.1 Preliminaries

The Lagrangian (2.68) for the nonrelativistic matter field had been adopted because it leads to the equation of motion (2.69) and the Hamiltonian (2.70), which is equivalent to (2.34) for a quantum, and the latter equation is the quantum version of the Newtonian energy-momentum relation for a particle. In order to arrive at the appropriate relativistic field equation, it suggests itself that one proceed analogously with the relativistic energy-momentum relation, which for a free particle of mass M reads

$$E^2 = \vec{p}^{\,2}c^2 + M^2c^4,$$

where c is the velocity of light. The result of thus defining the Hamiltonian operator to be $\mathbf{H} \stackrel{\text{def}}{=} \sqrt{\vec{\mathbf{p}}^{\,2}c^2 + M^2c^4}$ is a Schrödinger equation that reads either

$$i\hbar\frac{d\Psi}{dt} = \sqrt{\vec{\mathbf{p}}^{\,2}c^2 + M^2c^4}\,\Psi \tag{10.1}$$

or else its once iterated version

$$-\hbar^2\frac{d^2\Psi}{dt^2} = (\vec{\mathbf{p}}^{\,2}c^2 + M^2c^4)\Psi. \tag{10.2}$$

In the configuration representation, (10.1) would then lead to the equation

$$i\hbar\frac{\partial\psi}{\partial t} = \sqrt{-\hbar^2c^2\nabla^2 + M^2c^4}\,\psi,$$

which would be hard to handle because of the square root on the right, while (10.2) would lead to the more manageable

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = (-\hbar^2 c^2 \nabla^2 + M^2 c^4)\psi,$$

or

$$\left[-\hbar^2 \left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) + M^2 c^2\right] \psi = 0. \tag{10.3}$$

The field producing quanta whose one-particle wave function satisfies this equation can be generated from the Lagrangian density

$$\mathcal{L} = \hbar^2 \dot{\Psi}^\dagger \dot{\Psi} - \hbar^2 \nabla \Psi^\dagger \cdot \nabla \Psi - M^2 c^4 \Psi^\dagger \Psi,$$

which leads to the canonically conjugate momentum operator

$$\Pi = \hbar^2 \dot{\Psi}^\dagger$$

and the Hamiltonian[1]

$$\mathbf{H} = \int d^3 r \, [\hbar^{-2} \Pi \Pi^\dagger + \hbar^2 c^2 \nabla \Psi^\dagger \cdot \nabla \Psi + M^2 c^4 \Psi^\dagger \Psi]. \tag{10.4}$$

In the presence of an electromagnetic field described by the scalar potential ϕ and a vector potential $\vec{A}$, the concomitant relativistic Schrödinger equation in the configuration representation for a particle of charge e then is the following:

$$\left(i\hbar \frac{\partial}{\partial t} - e\phi\right)^2 \psi = \left[(-ic\hbar\nabla - e\vec{A})^2 + M^2 c^4\right] \psi, \tag{10.5}$$

which is, in fact, the equation proposed by Schrödinger as the appropriate relativistic generalization of his original, nonrelativistic equation. Eq. (10.5) is known as the *Klein-Gordon equation.*[2]

Eqs.(10.3) and (10.5), while relativistically invariant, have two serious flaws as proposed wave equations for electrons: first, while the field Hamiltonian (10.4) is positive definite, its eigenvalues are the squares of the energies, and there is no way of forcing these energies to be positive; they automatically come out both positive and negative. Second, if (10.5) is used for an electron in a hydrogen atom, it, by itself, does not account correctly for the fine structure of the spectrum, as we saw in Section 6.2.6; the Pauli term would have to be added ad hoc, in a Lorentz-invariant manner. Therefore, to describe the behavior of the wave function (or the underlying field) of an electrically charged particle of spin-1/2, such as the electron, a totally new equation is needed. (However, the Klein-Gordon equation does have its use as a relativistic wave equation for particles of spin 0.)

---

[1] **Show it.**

[2] After O. Klein and W. Gordon.

10.1.1 Relativistic spin

In order to be able to exploit the extensive analogies that exist between Lorentz transformations, which leave $x_1^2 + x_2^2 + x_3^2 - (ct)^2$ invariant, and rotations, which leave the Euclidean distance invariant, let us temporarily introduce a four-dimensional Euclidean space by setting $x_4 \stackrel{\text{def}}{=} ix_0 \stackrel{\text{def}}{=} ict$. This means that, at the expense of having to deal with a purely imaginary x_4, the Lorentz group is replaced by the rotation group O(4), which leaves $x_1^2 + x_2^2 + x_3^2 + x_4^2$ invariant. Furthermore, for reasons that will become apparent, it will be most convenient to generalize such objects as rotation and reflection operators from 3 to 5 dimensions, thus dealing with O(5), and subsequently to descend to 4, the dimensionality of space-time.

Recall what we know about the spin-1/2 rotation generators $\frac{1}{2}\vec{\sigma}$ in three dimensions. The unitary rotation operator for the spin degrees of freedom is given in terms of them by

$$\mathsf{S} = \exp[i\frac{\omega}{2}\widehat{n}\cdot\vec{\sigma}], \tag{10.6}$$

if the rotation axis and angle are $\widehat{n}$ and ω, respectively. The three components of $\vec{\sigma}$ form an axial vector and are, more properly, the components of an anti-symmetric tensor of rank two, as in (5.55), so that $\frac{1}{2}\sigma_{lk}$ is the generator of rotations that leave the (lk)-plane invariant. The appropriate commutation relations for such rotation generators are (5.56), and the anti-commutation relations for spin-1/2 are (5.57).

Next, define R_k as the unitary operator for a reflection *through the k-axis* for spinors, so that, for example, R_1 corresponds to the transformation $x_1' = x_1$, $x_2' = -x_2$, $x_3' = -x_3$. Since the σ_{kl} are the components of a tensor of rank 2, they must transform like

$$\begin{aligned}\sigma_1' &= \mathsf{R}_1^{-1}\sigma_1\mathsf{R}_1 = \mathsf{R}_1^{-1}\sigma_{23}\mathsf{R}_1 = \sigma_{23} = \sigma_1,\\ \sigma_2' &= \mathsf{R}_1^{-1}\sigma_{31}\mathsf{R}_1 = -\sigma_{31} = -\sigma_2,\\ \sigma_3' &= \mathsf{R}_1^{-1}\sigma_{12}\mathsf{R}_1 = -\sigma_{12} = -\sigma_3,\end{aligned}$$

etc., which implies

$$\mathsf{R}_k^{-1}\sigma_l\mathsf{R}_k = \begin{cases} \sigma_l & \text{for } k = l \\ -\sigma_l & \text{for } k \neq l, \end{cases}$$

(In other words, R_k commutes with σ_k and anti-commutes with the other σs.) or

$$\mathsf{R}_k^{-1}\sigma_{rs}\mathsf{R}_k = \begin{cases} -\sigma_{rs} & \text{for } k = r \text{ or } k = s \\ \sigma_{rs} & \text{for } k \neq r, k \neq s. \end{cases}$$

Because R_1 is a reflection operator through the 1-axis, it is equivalent to a rotation by $\pm\pi$ in the (23)-plane and must therefore be

$$\mathsf{R}_1 = \exp\left[\pm i\frac{\pi}{2}\sigma_{23}\right] = \pm i\sigma_{23} = \pm i\sigma_1,$$

according to (10.6) and (5.49), and similarly for the others:

$$\mathsf{R}_k = \pm i\sigma_k. \tag{10.7}$$

As reflection operators, the R_k must be the components of an axial vector, just like σ_k, and hence the signs in (10.7) must be the same for all three of them. They are thus skew-Hermitian and, since we are dealing with spin-1/2, it follows from (10.7) that they satisfy the anti-commutation relations

$$\frac{1}{2}\{\mathsf{R}_k, \mathsf{R}_l\} = -\delta_{kl}. \tag{10.8}$$

Furthermore, we find that

$$\mathsf{R}_1\mathsf{R}_2\mathsf{R}_3 = (\pm i)^3\sigma_1\sigma_2\sigma_3 = \pm\mathbb{1}$$

and if the plus sign in (10.7) is adopted by convention, i.e., $\mathsf{R}_k = +i\sigma_k$, then

$$\mathsf{R}_1\mathsf{R}_2\mathsf{R}_3 = \mathbb{1}. \tag{10.9}$$

From (5.56) we obtain the commutation relations

$$\frac{1}{2i}[\mathsf{R}_k, \sigma_{lm}] = \delta_{km}\mathsf{R}_l - \delta_{kl}\mathsf{R}_m. \tag{10.10}$$

Finally, notice that

$$\mathsf{R}_1\mathsf{R}_2 = -\sigma_1\sigma_2 = -i\sigma_3 = -i\sigma_{12},$$

so that more generally,

$$\mathsf{R}_k\mathsf{R}_l = -i\sigma_{kl} = \exp\left[-i\frac{\pi}{2}\sigma_{kl}\right], \tag{10.11}$$

which expresses a rotation by π in the (kl)-plane as two reflections, through the k- and l-axes.

In any dimension, the number $n(n-1)/2$ of real parameters in O(n) equals the number of components of anti-symmetric tensors of rank 2, so that rotations can be generated by $\frac{1}{2}\sum_{\mu,\nu=1}^{n}\varepsilon_{\mu\nu}\sigma_{\mu\nu}$ in terms of the rotation generators $\sigma_{\mu\nu}$. Alternatively, they may be viewed as products of $n(n-1)/2$ individual rotations $\frac{1}{2}\varphi\sigma_{\mu\nu}$, each of which keeps a plane invariant and requires only a single parameter. The rotation generators $\sigma_{\mu\nu}$ are therefore the most basic objects needed.

10.1.2 The γ-matrices

We now turn to O(5). Using Greek indices running from 1 to 5, there are 10 rotation generators $\Sigma_{\mu\nu} = -\Sigma_{\nu\mu}$, and 5 reflection operators, which we

shall call γ_μ for a reflection through the μ-axis. The rotation operators have to satisfy commutation relations analogous to (5.56),

$$\frac{1}{2i}[\Sigma_{\mu\nu}, \Sigma_{\lambda\kappa}] = \delta_{\mu\lambda}\Sigma_{\nu\kappa} - \delta_{\mu\kappa}\Sigma_{\nu\lambda} + \delta_{\nu\kappa}\Sigma_{\mu\lambda} - \delta_{\nu\lambda}\Sigma_{\mu\kappa}, \tag{10.12}$$

and because they are supposed to be a representation for spin-1/2, the anti-commutation relations

$$\frac{1}{2}\{\Sigma_{\mu\nu}, \Sigma_{\lambda\kappa}\} = \delta_{\mu\lambda}\delta_{\nu\kappa} - \delta_{\mu\kappa}\delta_{\nu\lambda} \tag{10.13}$$

analogous to (5.57), provided that one member of the pair $(\mu\nu)$ agrees with one of the $(\lambda\kappa)$. [Equation (10.12) implies that $[\Sigma_{\mu\nu}, \Sigma_{\lambda\kappa}] = 0$ if all four of the indices are different.]

In order to obtain the relation between the reflection operators and the rotation generators, we have to realize that whereas in three dimensions a reflection through the 1-axis is equivalent to a rotation by π in the (23)-plane, in five dimensions it is accomplished by a rotation in the (23)-plane combined with a rotation in the (45)-plane. Therefore, in place of (10.7), we must have, except possibly for an arbitrary sign, $\gamma_1 = (i\Sigma_{23})(i\Sigma_{45})$, and take conventionally

$$\gamma_1 = \Sigma_{23}\Sigma_{45} \tag{10.14}$$

and its cyclic permutations of the indices. As a result, we find that $\gamma_\mu^2 = \mathbb{1}$, and the reflection operators are Hermitian, whereas in three dimensions $\mathsf{R}_k^2 = -\mathbb{1}$ and they are skew-Hermitian. For $\mu \neq \nu$ we find from (10.12) and(10.14),[3]

$$\gamma_\mu\gamma_\nu = i\Sigma_{\mu\nu} = \exp[i\frac{\pi}{2}\Sigma_{\mu\nu}] \tag{10.15}$$

and hence

$$\frac{1}{2i}[\gamma_\mu, \gamma_\nu] = \Sigma_{\mu\nu}. \tag{10.16}$$

We also obtain from (10.14) and the commutation relations of the Σs,

$$\gamma_1\gamma_2\gamma_3\gamma_4\gamma_5 = -\mathbb{1}. \tag{10.17}$$

In fact, as you should verify for yourself, all the properties of the rotation operators imply those of the reflection operators and vice versa.

It is instructive to write down the entire list:

$$\begin{aligned}
\gamma_1 &= \Sigma_{23}\Sigma_{45} = \Sigma_{25}\Sigma_{34} = \Sigma_{24}\Sigma_{53} \\
\gamma_2 &= \Sigma_{34}\Sigma_{51} = \Sigma_{31}\Sigma_{45} = \Sigma_{35}\Sigma_{14} \\
\gamma_3 &= \Sigma_{45}\Sigma_{12} = \Sigma_{42}\Sigma_{51} = \Sigma_{41}\Sigma_{25} \\
\gamma_4 &= \Sigma_{51}\Sigma_{23} = \Sigma_{53}\Sigma_{12} = \Sigma_{52}\Sigma_{31} \\
\gamma_5 &= \Sigma_{12}\Sigma_{34} = \Sigma_{14}\Sigma_{23} = \Sigma_{13}\Sigma_{42}.
\end{aligned}$$

[3] **Check this.**

Note particularly that the expressions for γ_1, γ_2, and γ_3 are just like those for the Rs in three dimensions, except that they are all multiplied by Σ_{45}.

So here we have 16 linearly independent Hermitian operators,

$$\mathbb{1},\ \Sigma_{\mu\nu},\ \gamma_\mu, \qquad \mu,\nu = 1,\ldots,5,$$

which form the basis of a linear algebra. [The product of a γ with a Σ can be expressed as a multiple of a Σ by means of (10.15) and (10.17), the product of two γs in terms of a Σ by (10.15), and the product of two Σs in terms of γs by (10.15) and (10.17). This is the reason for going to five dimensions rather than directly to four, where the rotation and reflection operators, together with the unit operator, would not form the basis of an algebra.] Therefore, for a representation of these 16 operators by means of 16 linearly independent matrices, these matrices can be no smaller than 4×4.

How many of these 4×4 matrices can be chosen to be real and how many purely imaginary? Since the operators are Hermitian, so must be the matrices; hence a real matrix has to be symmetric and an imaginary one anti-symmetric. Among the 16 independent 4×4 matrices, exactly ten can be chosen symmetric and six anti-symmetric; therefore ten of the Hermitian matrices can be taken to be real and six imaginary. The only consistent choice, as you should check for yourself (for example, all six γs real would lead to all ten Σs imaginary, which is too many), is to have three real and two imaginary γs, which yields four imaginary Σs and six real ones; since $\mathbb{1}$ is real, this works out correctly. So we choose γ_1, γ_2, and γ_3 to be real (because eventually, the first three coordinates play a different physical role from the others) and γ_4 and γ_5 purely imaginary.

When descending to four dimensions, from now on unless otherwise noted, *Greek subscripts will be understood to run from 1 to 4,* with 5 explicitly written out. Since physics demands a Lorentzian rather then Euclidean metric, the requirement is that the fourth component of any "real" vector be imaginary, while the first three have to be real. Similarly, all the components of a tensor of rank two have to be real, except those with *one* subscript equal to 4, which are imaginary. Applied to the tensor $\langle\Sigma_{\mu\nu}\rangle$, this implies that, even though $\Sigma_{\mu\nu}$ is Hermitian, we must have

$$\langle\Sigma_{\mu\nu}\rangle^* = \begin{cases} -\langle\Sigma_{\mu\nu}\rangle & \text{for } \mu \text{ or } \nu = 4 \\ \langle\Sigma_{\mu\nu}\rangle & \text{for } \mu,\nu \neq 4 \end{cases},$$

which implies

$$\langle\Sigma_{\mu\nu}\rangle^* = \langle\gamma_4\Sigma_{\mu\nu}\gamma_4\rangle. \tag{10.18}$$

Because the carrier space corresponding to the 4×4 rotation matrices is four-dimensional, the vectors u in this space, i.e., the relativistic spinors,

have to have four components:[4] u_α, $\alpha = 1, \ldots, 4$, and we shall regard them as column matrices, writing $(u, v) \stackrel{\text{def}}{=} \sum_\alpha u^*_\alpha v_\alpha \stackrel{\text{def}}{=} \widetilde{u}^* v = u^\dagger v$. However, if we were to define $\langle \Sigma_{\mu\nu} \rangle \stackrel{\text{def}}{=} \widetilde{u}^* \Sigma_{\mu\nu} u$ for normalized spinors u, as we did nonrelativistically, the result would be $\langle \Sigma_{\mu\nu} \rangle^* = \langle \Sigma_{\mu\nu} \rangle$, rather than (10.18). The expectation value therefore has to be defined relativistically as[5]

$$\langle \Sigma_{\mu\nu} \rangle \stackrel{\text{def}}{=} u^\dagger \gamma_4 \Sigma_{\mu\nu} u.$$

It is consequently convenient to define an *adjoint spinor* by

$$\bar{u} \stackrel{\text{def}}{=} \widetilde{u}^* \gamma_4 = u^\dagger \gamma_4, \tag{10.19}$$

so that

$$\langle \Sigma_{\mu\nu} \rangle \stackrel{\text{def}}{=} u^\dagger \gamma_4 \Sigma_{\mu\nu} u = \bar{u} \Sigma_{\mu\nu} u, \tag{10.20}$$

as well as generally for any operator $\mathbf{A}$ acting a relativistic spinor wave functions,

$$\langle \mathbf{A} \rangle \stackrel{\text{def}}{=} \bar{u} \mathbf{A} u, \tag{10.21}$$

(assuming u is normalized, see below) from which it follows that

$$\langle \mathbf{A} \rangle^* = \langle \gamma_4 \mathbf{A}^\dagger \gamma_4 \rangle,$$

implying that $\langle i\gamma_5 \rangle$ is real and $\langle i\gamma_\mu \rangle$ and $\langle i\gamma_5\gamma_\mu \rangle$ have the correct reality properties of four-vectors. For example, let us test the reality property of $\langle i\gamma_\mu \rangle$ with this new definition of the expectation value. We find that for $\mu = 1, 2, 3$,

$$\langle i\gamma_\mu \rangle^* = (\widetilde{u}^* \gamma_4 i \gamma_\mu u)^* = -\widetilde{u}^* i \gamma_\mu \gamma_4 u = \bar{u} i \gamma_\mu u = \langle i\gamma_\mu \rangle,$$

whereas $\langle i\gamma_4 \rangle^* = -\langle i\gamma_4 \rangle$, which is just the way a "real" four-vector ought to behave. Similarly for $\langle i\gamma_5\gamma_\mu \rangle$.

There is, however, another difficulty. The interpretation of $\bar{u}\mathbf{A}u$ as the expectation value of $\mathbf{A}$ is based on the assumption that $\bar{u}u = 1$. But, although $\bar{u}u$ is real, there is no assurance that it is positive for all u, so that $\bar{u}u$ cannot necessarily be made equal to 1 by rescaling of u.

For infinitesimal rotations in the $(\mu\nu)$-plane by an angle $\varepsilon \ll \pi$ the rotation operator that transforms spinors according to $u' = \mathsf{S}u$ is given by

$$\mathsf{S} = \mathbb{1} + \frac{i}{2} \varepsilon \Sigma_{\mu\nu},$$

[4]These four spinor components are not to be confused with four space-time components!

[5]**Exercise: Show that this definition satisfies (10.18).**

which leads to the rotation operator for finite rotations by the angle φ in the $(\mu\nu)$-plane,

$$\mathsf{S} = \lim_{n\to\infty}\left(\mathbb{1} + \frac{i}{2}\frac{\varphi}{n}\Sigma_{\mu\nu}\right)^n = \exp[\frac{i}{2}\varphi\Sigma_{\mu\nu}] = \mathbb{1}\cos\frac{\varphi}{2} + i\Sigma_{\mu\nu}\sin\frac{\varphi}{2}. \quad (10.22)$$

The most general rotation in 5 dimensions is the product of 10 such operators, analogous to the way in which a general rotation in three dimensions is the product of three rotations by Euler angles. However, the general rotation matrix for infinitesimal rotations can also be written in the form

$$\mathsf{S} = \mathbb{1} + \frac{i}{2}\sum_{\mu>\nu=1}^{4}\varepsilon_{\mu\nu}\Sigma_{\mu\nu} = \mathbb{1} + \frac{i}{4}\sum_{\mu,\nu=1}^{4}\varepsilon_{\mu\nu}\Sigma_{\mu\nu},$$

where $\varepsilon_{\mu\nu} = -\varepsilon_{\nu\mu}$, $|\varepsilon_{\mu\nu}| \ll 1$, but this cannot be simply exponentiated as in (10.22).

The inverse of the rotation matrix (10.22) is of course given by

$$\mathsf{S}^{-1} = \exp[-\frac{i}{2}\varphi\Sigma_{\mu\nu}] = \mathbb{1}\cos\frac{\varphi}{2} - i\Sigma_{\mu\nu}\sin\frac{\varphi}{2},$$

which for real φ is equal to $\mathsf{S}^\dagger$, because $\Sigma_{\mu\nu}$ is Hermitian; but for actual Lorentz transformations, including boosts, for which one of the indices μ or ν equals 4, the angle φ has to be taken imaginary (see below), so that in that case $\mathsf{S}^\dagger = \mathsf{S}$. These cases can all be combined in the formula

$$\mathsf{S}^\dagger = \gamma_4\mathsf{S}^{-1}\gamma_4; \quad (10.23)$$

thus S is generally not unitary. As a result of (10.23), if under a Lorentz transformation a spinor u goes over into

$$u' = \mathsf{S}u,$$

then its adjoint becomes

$$\bar{u}' = u^{\dagger\prime}\gamma_4 = u^\dagger\mathsf{S}^\dagger\gamma_4 = \bar{u}\mathsf{S}^{-1},$$

so that

$$\langle\mathbf{A}\rangle' = \bar{u}'\mathbf{A}u' = \bar{u}\mathsf{S}^{-1}\mathbf{A}\mathsf{S}u = \langle\mathsf{S}^{-1}\mathbf{A}\mathsf{S}\rangle. \quad (10.24)$$

The matrices S form a four-dimensional representation of the Lorentz group which is different from the defining representation given by the 4×4 matrices that transform space-time four-vectors, and the carrier space of this representation is the space of relativistic spinors which transform like

$$u' = \mathsf{S}u.$$

We now abandon imaginary fourth components of venctors and use $\mu, \nu = 0, 1, 2, 3$, defining $\gamma_0 \stackrel{\text{def}}{=} -i\gamma_4$, and the Lorentz metric $-g_{00} = g_{11} = g_{22} = g_{33} = 1$, $g_{\mu\nu} = 0$ for $\mu \neq \nu$, and $g_{\mu\nu} = g^{\mu\nu}$; then

$$\gamma_\mu \gamma_\nu = g_{\mu\nu} + i\Sigma_{\mu\nu}. \tag{10.25}$$

According to (10.17), the Hermitian matrix γ_5 is related to the other γ matrices by

$$\gamma_5 = -\gamma_1\gamma_2\gamma_3\gamma_4 = i\gamma_0\gamma_1\gamma_2\gamma_3. \tag{10.26}$$

and the adjoint spinor, in terms of which expectation values are defined as in (10.21), is given by

$$\bar{u} \stackrel{\text{def}}{=} iu^\dagger\gamma_0.$$

The γ matrices play a central role in the relativistic formulation of quantum mechanics and it is important to study their properties.

Let us, then, verify that $\langle i\gamma_\mu \rangle$ behaves like a vector under rotations. (Since we found $\langle i\gamma_4 \rangle$ to be imaginary, $\langle i\gamma_0 \rangle$ is real, so that all four components of $\langle i\gamma_\mu \rangle$, $\mu = 0, \ldots, 3$ are real.) You will easily check,[6] using (10.16), that for (10.22), that is, for a rotation in the $(\mu\nu)$-plane,

$$\mathsf{S}^{-1} i\gamma_\mu \mathsf{S} = i\gamma_\mu \cos\varphi + i\gamma_\nu \sin\varphi,$$

$$\mathsf{S}^{-1} i\gamma_\nu \mathsf{S} = i\gamma_\nu \cos\varphi - i\gamma_\mu \sin\varphi,$$

whereas S leaves the other three γs unchanged. It therefore follows that under the rotation (10.22) in the $(\mu\nu)$-plane, for $\mu, \nu = 1, 2, 3$,

$$\langle i\gamma_\mu \rangle' = \cos\varphi \, \langle i\gamma_\mu \rangle + \sin\varphi \, \langle i\gamma_\nu \rangle,$$

$$\langle i\gamma_\nu \rangle' = -\sin\varphi \, \langle i\gamma_\mu \rangle + \cos\varphi \, \langle i\gamma_\nu \rangle,$$

and for $\kappa \neq \mu, \nu$,

$$\langle i\gamma_\kappa \rangle' = \langle i\gamma_\kappa \rangle.$$

For an actual Lorentz transformation describing boosts, say, of velocity v along the x-axis, we have to take $\mu = 1$, $\nu = 4$, and set $\varphi = i\psi$, where

$$\sinh\psi \stackrel{\text{def}}{=} \frac{v/c}{\sqrt{1-(v/c)^2}}, \qquad \cosh\psi = \frac{1}{\sqrt{1-(v/c)^2}},$$

which makes

$$\mathsf{S} = \exp[-\frac{1}{2}\psi\,\Sigma_{14}] = \mathbf{1}\cosh\psi - \Sigma_{14}\sinh\psi.$$

The result is

$$\langle i\gamma_1 \rangle' = \cosh\psi \, \langle i\gamma_1 \rangle + i\sinh\psi \, \langle i\gamma_4 \rangle,$$

[6] **Do it.**

$$\langle i\gamma_4\rangle' = -i\sinh\psi\,\langle i\gamma_1\rangle + \cosh\psi\,\langle i\gamma_4\rangle,$$

or

$$\langle i\gamma_1\rangle' = \frac{\langle i\gamma_1\rangle - (v/c)\langle i\gamma_0\rangle}{\sqrt{1-(v/c)^2}},$$

$$\langle i\gamma_0\rangle' = \frac{\langle i\gamma_0\rangle - (v/c)\langle i\gamma_1\rangle}{\sqrt{1-(v/c)^2}}.$$

Thus for general proper, orthochronous Lorentz transformations given by

$$x'_\mu = \sum_{\nu=0}^{3} A_\mu{}^\nu x_\nu, \quad \mu = 0,1,2,3,$$

(where now only real coordinates have been employed, with $x_0 = ct$), we have

$$\mathsf{S}^{-1}\gamma_\mu\mathsf{S} = \sum_{\nu=0}^{3} A_\mu{}^\nu\gamma_\nu, \quad \mu = 0,1,2,3, \tag{10.27}$$

and therefore

$$\langle i\gamma_\mu\rangle' = \sum_{\nu=0}^{3} A_\mu{}^\nu\langle i\gamma_\nu\rangle, \quad \mu = 0,1,2,3,$$

which means that the four quantities $\langle i\gamma_\mu\rangle$, $\mu = 0,1,2,3$, transform like the coordinates x_μ and therefore form a Lorentz four-vector.

We similarly find[7] that under proper rotations and Lorentz transformations, $\langle i\gamma_5\rangle$ is invariant and $\langle i\gamma_5\gamma_\mu\rangle$ is a four-vector, while under *space reflection,* for which $\mathsf{S} = i\gamma_4$, $\langle i\gamma_\mu\rangle$ is a vector, $\langle i\gamma_5\gamma_\mu\rangle$ a pseudo-vector, $\langle\,|\,\rangle$ a scalar, and $\langle i\gamma_5\rangle$ a pseudo-scalar, and under *time reversal,* with $\mathsf{S} = i\gamma_4\gamma_5$, $\langle\,|\,\rangle$ changes sign, $\langle i\gamma_5\rangle$ is invariant, $\langle i\gamma_\mu\rangle$ is a pseudo-vector, and $\langle i\gamma_5\gamma_\mu\rangle$ is a vector. The operator $\mathsf{S} = \gamma_5$ does a complete space-time reflection, and under this transformation $\langle\,|\,\rangle$ and $\langle i\gamma_5\rangle$ are pseudo-scalars, while $\langle i\gamma_\mu\rangle$ and $\langle i\gamma_5\gamma_\mu\rangle$ are pseudo-vectors.

Just as it was useful nonrelativistically to define a spin exchange matrix by (5.69), so we define here the matrix

$$\mathfrak{I} \stackrel{\text{def}}{=} \frac{1}{4}\sum_{A=1}^{16} \mathbf{S}_A^{(1)} \otimes \mathbf{S}_A^{(2)}, \tag{10.28}$$

where $\mathbf{S}_A$ is one of the 16 Hermitian matrices (all of which are such that $\mathbf{S}_A^2 = \mathbb{1}$) $\mathbb{1}, \gamma_\mu, \Sigma_{\mu\nu}$, $\mu,\nu = 1,\ldots,5$, and the superscript distinguishes between the two systems, $\mathbf{S}_A^{(1)}$ acting on system #1 and $\mathbf{S}_A^{(2)}$ acting on #2. It has the property that

$$\mathbf{S}_C^{(1)}\mathfrak{I} = \mathfrak{I}\mathbf{S}_C^{(2)}, \quad C = 1,\ldots,16, \tag{10.29}$$

[7] **Verify all these statements.**

as well as

$$\mathbf{S}_C^{(2)}\mathfrak{I} = \mathfrak{I}\mathbf{S}_C^{(1)}, \quad C = 1, \ldots, 16. \tag{10.30}$$

Moreover, as you will **prove in a homework problem,**

$$\mathfrak{I}^2 = \mathbb{1}, \tag{10.31}$$

where $\mathbb{1}$ is the unit matrix both in the space #1 and in #2.

Since the $\mathbf{S}_A$ form the basis of an algebra, we have for $A \neq B$,

$$\mathbf{S}_A\mathbf{S}_B = \sum_D a_{ABD}\mathbf{S}_D. \tag{10.32}$$

The coefficients a_{ABD} can be calculated by taking traces, and we find (as I will ask you to **prove as a homework problem**)

$$a_{ABC} = \frac{1}{4}\mathrm{tr}\mathbf{S}_A\mathbf{S}_B\mathbf{S}_C, \tag{10.33}$$

which implies that

$$a_{ABC} = a_{CAB} = a_{BCA}.$$

Consequently we obtain

$$\begin{aligned}\mathbf{S}_C^{(1)}\mathfrak{I} &= \sum_A \mathbf{S}_C^{(1)}\mathbf{S}_A^{(1)} \otimes \mathbf{S}_A^{(2)} = \sum_{AD} a_{CAD}\mathbf{S}_D^{(1)} \otimes \mathbf{S}_A^{(2)} \\ &= \sum_{AD} \mathbf{S}_D^{(1)} \otimes a_{DCA}\mathbf{S}_A^{(2)} = \sum_D \mathbf{S}_D^{(1)} \otimes \mathbf{S}_D^{(2)}\mathbf{S}_C^{(2)} = \mathfrak{I}\mathbf{S}_C^{(2)}.\end{aligned}$$

Equation (10.30) is proved similarly.

Therefore $\mathfrak{I}$ is the exchange operator between #1 and #2, and because of (10.31), its eigenvalues are ± 1. If two particles are in a state that is symmetric under spin exchange, then $\mathfrak{I} = \mathbb{1}$, and if they are in an anti-symmetric spin state, $\mathfrak{I} = -\mathbb{1}$. The exchange operator can also be expressed in terms of the γ-matrices,

$$\mathfrak{I} = \mathbb{1} - \frac{1}{8}(\mathbb{1} - \sum_1^5 \gamma_\mu^{(1)} \otimes \gamma_\mu^{(2)})^2, \tag{10.34}$$

as I will ask you to **prove as a homework problem.** It therefore follows that for anti-symmetric states,

$$\sum_1^5 \gamma_\mu^{(1)} \otimes \gamma_\mu^{(2)} = \begin{cases} -3\mathbb{1} \\ 5\mathbb{1} \end{cases},$$

while for states symmetric under spin exchange

$$\sum_1^5 \gamma_\mu^{(1)} \otimes \gamma_\mu^{(2)} = \mathbb{1},$$

i.e.,

$$\sum_{1}^{4} \gamma_\mu^{(1)} \otimes \gamma_\mu^{(2)} = \mathbb{1} - \gamma_5^{(1)} \otimes \gamma_5^{(2)}, \tag{10.35}$$

and

$$\sum_{\mu<\nu=1}^{5} \Sigma_{\mu\nu}^{(1)} \otimes \Sigma_{\mu\nu}^{(2)} = 2\mathbb{1},$$

which implies that for symmetric states

$$\sum_{\mu<\nu=1}^{4} \Sigma_{\mu\nu}^{(1)} \otimes \Sigma_{\mu\nu}^{(2)} = \mathbb{1} + \gamma_5^{(1)} \otimes \gamma_5^{(2)}. \tag{10.36}$$

As an application, suppose that $u^{(1)}$ and $u^{(2)}$ are the same states, so that $u^{(1)} \otimes u^{(2)} = u \otimes u$ is symmetric under exchange. It then follows from (10.35) that

$$-\sum_{\mu,\nu=0}^{3} (\bar{u} i\gamma_\mu u) g^{\mu\nu} (\bar{u} i\gamma_\nu u) = (\bar{u}u)^2 + (\bar{u} i\gamma_5 u)^2 > 0, \tag{10.37}$$

since both $\bar{u}u = \widetilde{u}^* \gamma_4 u$ and $\bar{u} i\gamma_5 u$ are real. Therefore, $\langle i\gamma_\mu \rangle$ is a *time-like* four-vector. If (10.35) is multiplied by $i\gamma_5^{(1)} \otimes i\gamma_5^{(2)}$, the result is

$$\sum_{1}^{4} (i\gamma_5^{(1)} \gamma_\mu^{(1)}) \otimes (i\gamma_5^{(2)} \gamma_\mu^{(2)}) = \mathbb{1} + i\gamma_5^{(1)} \otimes i\gamma_5^{(2)},$$

which, when expectation values are taken, leads to

$$\sum_{1}^{4} (\bar{u} i\gamma_5 \gamma_\mu u)^2 = \sum_{\mu,\nu=0}^{3} (\bar{u} i\gamma_5 \gamma_\mu u) g^{\mu\nu} (\bar{u} i\gamma_5 \gamma_\nu u) = (\bar{u}u)^2 + (\bar{u} i\gamma_5 u)^2 > 0, \tag{10.38}$$

implying that $\langle i\gamma_5 \gamma_\mu \rangle$ is a *space-like* four-vector. Not only that, but since multiplication of (10.35) by $i\gamma_5^{(1)}$ yields

$$\sum_{\mu=1}^{4} i\gamma_5^{(1)} \gamma_\mu^{(1)} \otimes \gamma_\mu^{(2)} = i\gamma_5^{(1)} - i\gamma_5^{(2)},$$

it follows that

$$\sum_{1}^{4} (\bar{u} i\gamma_5 \gamma_\mu u)(\bar{u} i\gamma_\mu u) = \sum_{\mu,\nu=0}^{3} (\bar{u} i\gamma_5 \gamma_\mu u) g^{\mu\nu} (\bar{u} i\gamma_\nu u) = 0, \tag{10.39}$$

which says that $\langle i\gamma_5 \gamma_\mu \rangle$ and $\langle i\gamma_\mu \rangle$ are orthogonal to one another.

In a similar manner we find from (10.36) that

$$\sum_{\mu<\nu=1}^{4}(\bar{u}\Sigma_{\mu\nu}u)^2=\sum_{\mu>\nu=0,\lambda>\kappa=0}^{3}(\bar{u}\Sigma_{\mu\nu}u)g^{\mu\lambda}g^{\nu\kappa}(\bar{u}\Sigma_{\lambda\kappa}u)=(\bar{u}u)^2-(\bar{u}i\gamma_5u)^2 \tag{10.40}$$

as well as

$$\begin{aligned}\sum_{\mu>\nu=1}^{4}(\bar{u}i\gamma_5\Sigma_{\mu\nu}u)^2 &= \sum_{\mu>\nu=0,\lambda>\kappa=0}^{3}(\bar{u}i\gamma_5\Sigma_{\mu\nu}u)g^{\mu\lambda}g^{\nu\kappa}(\bar{u}\gamma_5\Sigma_{\lambda\kappa}u)\\ &= (\bar{u}u)^2+(\bar{u}i\gamma_5u)^2\end{aligned} \tag{10.41}$$

and

$$\begin{aligned}\sum_{\mu>\nu=1}^{4}(\bar{u}i\gamma_5\Sigma_{\mu\nu}u)(\bar{u}\Sigma_{\mu\nu}u) &= \sum_{\mu>\nu=0,\lambda>\kappa=0}^{3}(\bar{u}i\gamma_5\Sigma_{\mu\nu}u)g^{\mu\lambda}g^{\nu\kappa}(\bar{u}\Sigma_{\lambda\kappa}u)\\ &= 2(\bar{u}u)(\bar{u}i\gamma_5u).\end{aligned} \tag{10.42}$$

These equations imply that under proper Lorentz transformations, which leave $\bar{u}u$ and $\bar{u}i\gamma_5u$ invariant, $\langle\Sigma_{\mu\nu}\rangle$ and $\langle i\gamma_5\Sigma_{\mu\nu}\rangle$ behave like tensors of rank two.

Charge conjugation

There is another important transformation of the γ matrices, which, for reasons that will become clear in Section 10.2.1, is called *charge conjugation.* It is defined by a unitary matrix C, $\mathsf{C}^\dagger\mathsf{C}=\mathbb{1}$, that takes a spinor u into

$$u^c\stackrel{\text{def}}{=}\mathsf{C}\widetilde{\bar{u}}, \tag{10.43}$$

and which is such that for $\mu=0,\ldots,3$,

$$\mathsf{C}^{-1}\gamma_\mu\mathsf{C}=-\widetilde{\gamma}_\mu. \tag{10.44}$$

Such a transformation leaves the anti-commutation relations implied by (10.25) unchanged, and it may be chosen to be $\mathsf{C}=\gamma_0$ if the four γ-matrices $\gamma_0,\ldots,\gamma_3$ are taken to be real. The four-component objects u^c are spinors under proper Lorentz transformations,[8] that is, they are such that $u^{c\prime}=\mathsf{S}u^c$ if $u'=\mathsf{S}u$.

It is easy to check[9] that the definition (10.43) implies

$$\bar{u^c}=-\widetilde{u}\mathsf{C}^{-1},$$

[8] **It will be a homework exercise to show this.**
[9] **Do it.**

and therefore

$$\bar{u^c}u^c = -\bar{u}u.$$

Similarly, we find that

$$\bar{u^c}i\gamma_\mu u^c = \bar{u}i\gamma_\mu u,$$

as well as

$$\bar{u^c}i\gamma_5 u^c = -\bar{u}i\gamma_5 u,$$

and

$$\bar{u^c}i\gamma_5\gamma_\mu u^c = -\bar{u}i\gamma_5\gamma_\mu u.$$

So, under charge conjugation, $\langle i\gamma_\mu\rangle$ remains invariant, while $\langle i\gamma_5\rangle$, $\langle i\gamma_5\gamma_\mu\rangle$, and $\langle\,|\,\rangle$ change sign.

10.2 The Dirac Equation

We now have the tools to construct a suitable differential operator that can be used in a relativistically invariant field equation. Since the set γ_μ transforms under Lorentz transformations as in (10.27), it follows that

$$\gamma_\mu \frac{\partial}{\partial x_\mu} \stackrel{\text{def}}{=} \gamma_\mu \partial^\mu$$

is such that in a new coordinate frame

$$(\gamma_\mu \partial^\mu)' \stackrel{\text{def}}{=} \gamma_\mu \frac{\partial}{\partial x'_\mu} = \mathsf{S}\gamma_\mu \partial^\mu \mathsf{S}^{-1}.$$

(From now on we shall use the *summation convention,* so that repeated indices are automatically summed over from 0 to 3.) This operator will be denoted by $\not\partial$:

$$\not\partial \stackrel{\text{def}}{=} \gamma_\mu \partial^\mu = \vec{\gamma}\cdot\nabla + \gamma_0 \frac{1}{c}\frac{\partial}{\partial t},$$

(here $\vec{\gamma}$ denotes the three-vector with the components $\gamma_1, \gamma_2, \gamma_3$) so that under Lorentz transformations

$$\not\partial' = \mathsf{S}\not\partial\mathsf{S}^{-1}, \tag{10.45}$$

where $\not\partial' \stackrel{\text{def}}{=} \gamma_\mu \partial'^\mu$. (Note particularly that in every frame the same set of γ matrices may be used.) According to (10.25)

$$\not\partial^2 = \frac{1}{2}\{\gamma_\mu, \gamma_\nu\}\partial^\mu\partial^\nu = g_{\mu\nu}\partial^\mu\partial^\nu = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}. \tag{10.46}$$

Thus $\not\partial$ may be regarded as a square root of the relativistically invariant d'Alembert operator $\nabla^2 - \partial^2/\partial(ct)^2$.

Let us, then, write down the simplest candidate for a Lorentz invariant Lagrangian density for a free electron field:

$$\begin{aligned}\mathcal{L} &= \bar{\Psi}\left(\frac{1}{2}\hbar c\not{\partial} - \frac{1}{2}\hbar c\not{\partial}_{\leftarrow} + Mc^2\right)\Psi \\ &= \Psi^\dagger\left(\frac{1}{2}\hbar c\gamma_4\vec{\gamma}\cdot\nabla - \frac{1}{2}\hbar c\gamma_4\vec{\gamma}\cdot\nabla_{\leftarrow}\right. \\ &\qquad \left. - i\,\frac{\hbar}{2}\frac{\partial}{\partial t} + i\frac{\hbar}{2}\frac{\partial}{\partial t_{\leftarrow}} + \gamma_4 Mc^2\right)\Psi, \end{aligned} \tag{10.47}$$

(where the subscript $_{\leftarrow}$ indicates that the differentiation is to operate toward the left) if we assume the field Ψ to have four components that transform like those of a relativistic spinor, i.e., $\Psi' = \mathsf{S}\Psi$ under Lorentz transformations. [We will see shortly what the physical meaning of the constant term Mc^2 in (10.47) is.] The Lagrangian equation of motion obtained from (2.65) for the free field then reads[10]

$$(\hbar\not{\partial} + Mc)\Psi = 0, \tag{10.48}$$

which is the *Dirac equation.* The result of multiplying it by $(\hbar\not{\partial} - Mc)$ is, according to (10.46),

$$\left(\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} - \hbar^2\nabla^2 + M^2c^2\right)\Psi = 0, \tag{10.49}$$

so that every solution of (10.48) must also solve (10.49), which is identical to (10.3). Thus M in the Lagrangian (10.47) will be the mass of the quanta produced by the field. The Hamiltonian is obtained just like (2.71),

$$\mathbf{H} = -\int d^3r\,\bar{\Psi}(\hbar c\vec{\gamma}\cdot\nabla + Mc^2)\Psi. \tag{10.50}$$

That (10.48) is a covariant equation follows immediately from the fact that the Lagrangian from which it is generated is invariant, but we also find directly that after a Lorentz transformation (10.48) reads

$$(\hbar\not{\partial}' + Mc)\Psi' = 0.$$

It thus has the same form in the moving or rotated laboratory as before, with the same γ-matrices.

Before we can find out what kinds of particles will be produced by the field that satisfies (10.48), we have to look at the solutions of the same equation as applied, not to a field but to a numerical spinor function:

$$(\hbar\not{\partial} + Mc)u(\vec{r},t) = 0. \tag{10.51}$$

[10] **Show this as an exercise.**

This equation will obviously have plane wave solutions of the form

$$u(\vec{r},t) = u\exp[i(\vec{k}\cdot\vec{r}-\omega t)] = u\exp(ix_\mu k^\mu),$$

where $k^0 \stackrel{\text{def}}{=} -\omega/c$ and the constant spinor u must satisfy the equation

$$(i\hbar\gamma_\mu k^\mu + Mc)u = 0. \tag{10.52}$$

Multiplication of this equation by $(i\hbar\gamma_\mu k^\mu - Mc)$ yields $(M^2c^2 + \hbar^2\vec{k}^2 - \hbar^2\omega^2/c^2)u = 0$, which implies that ω is related to $\vec{k}$ by

$$\omega = \pm\sqrt{\frac{M^2c^4}{\hbar^2} + c^2\vec{k}^2}.$$

For a given vector $\vec{k}$ and a given value of M, the Hermitian 4×4 matrix $\Gamma(\vec{k}) \stackrel{\text{def}}{=} \gamma_0(-\hbar\vec{\gamma}\cdot\vec{k}+iMc)$ must have four linearly independent eigenvectors; two of them belong to the eigenvalue $\hbar\omega/c = +\sqrt{M^2c^2+\hbar^2\vec{k}^2}$ and two to $\hbar\omega/c = -\sqrt{M^2c^2+\hbar^2\vec{k}^2}$. To label the degenerate states, we introduce the Hermitian matrix $\vec{\Sigma}\cdot\widehat{k}$ [where $\Sigma_1 \stackrel{\text{def}}{=} \Sigma_{23}$, and its cyclic permutations, $\Sigma_{\mu\nu}$ is given by (10.25), and $\widehat{k} \stackrel{\text{def}}{=} \vec{k}/|\vec{k}|$]; this matrix, which has the eigenvalues ± 1 called positive and negative *helicity*, commutes with $\Gamma(\vec{k})$ and can be employed to label the states by $u_\pm(\vec{k},\omega)$. Moreover, (10.44) together with the facts that the components of $\vec{\gamma}$ are Hermitian while γ_0 is skew-Hermitian, leads to

$$\mathsf{C}\Gamma^*(\vec{k})\mathsf{C}^{-1} = -\Gamma(-\vec{k}),$$

from which it follows by (10.43) that if $\Gamma(\vec{k})u(\vec{k}) = (\hbar\omega/c)u(\vec{k})$, then

$$\Gamma(\vec{k})u^c(-\vec{k}) = -(\hbar\omega/c)\,u^c(-\vec{k}).$$

Therefore *every negative-frequency solution equals the charge conjugate of a positive-frequency solution, and vice versa.* We also find[11] that if $\vec{\Sigma}\cdot\widehat{k}u = u$, then $\vec{\Sigma}\cdot\widehat{k}u^c = -u^c$. Thus there are four mutually orthogonal eigenvectors of $\Gamma(\vec{k})$, which can be labeled $u_{\pm,\pm}$, the first $\pm$ indicating positive or negative helicity and the second the eigenvalues $\omega = \pm\sqrt{c^2k^2 + M^2c^4/\hbar^2}$. [However, since $u_{\pm,-}(\vec{k}) = -(u_{\mp,+})^c(-\vec{k})$, we can equally well take the four to be $u_{+,+}$, $u_{-,+}$, and their charge conjugates.] We will assume that the $u_{\pm,\pm}(\vec{k})$ are normalized, $u^\dagger_{\pm,\pm}(\vec{k})u_{\pm,\pm}(\vec{k}) = 1$. This implies[12] that

$$\bar{u}_{\pm,\pm}(\vec{k})u_{\pm,\pm}(\vec{k}) = \frac{Mc^2}{\hbar\omega}u^\dagger_{\pm,\pm}(\vec{k})u_{\pm,\pm}(\vec{k}) = \frac{Mc^2}{\hbar\omega}.$$

[11] **Show this.**
[12] **Show this.**

The spinor functions $u_{\pm,+}(\vec{k})\exp(i\vec{k}\cdot\vec{r}-i\omega t)$ and $u_{\pm,-}(\vec{k})\exp(i\vec{k}\cdot\vec{r}+i\omega t)$, with $\omega = +\sqrt{c^2k^2 + M^2c^4/\hbar^2}$, form a complete set of solutions of (10.51) and the spinor field $\Psi(\vec{r},t)$ subject to (10.48) can be expanded in the form

$$\Psi(\vec{r},t) = (2\pi)^{-3/2}\int d^3k \sum_{h=\pm}\Big[e^{i(\vec{k}\cdot\vec{r}-\omega t)}u_{h,+}(\vec{k})\mathbf{a}_h(\vec{k}) + e^{i(\vec{k}\cdot\vec{r}+\omega t)}u_{h,-}(\vec{k})\mathbf{b}_h^\dagger(\vec{k})\Big], \quad (10.53)$$

where the coefficients $\mathbf{a}_h(\vec{k})$ and $\mathbf{b}_h(\vec{k})$ are operators given by

$$\mathbf{a}_h(\vec{k}) = (2\pi)^{-3/2}\int d^3r\, u_{h,+}^\dagger(\vec{k})\Psi(\vec{r},t)e^{-i(\vec{k}\cdot\vec{r}-\omega t)}$$

$$\mathbf{b}_h^\dagger(\vec{k}) = (2\pi)^{-3/2}\int d^3r\, u_{h,-}^\dagger(\vec{k})\Psi(\vec{r},t)e^{-i(\vec{k}\cdot\vec{r}+\omega t)}.$$

Since Ψ is a spinor field, it has to satisfy anti-commutation relations (the spin-statistics connection) as in (2.80), (2.81), and (2.82), whose use leads to the anti-commutators[13]

$$\{\mathbf{a}_h(\vec{k}),\mathbf{a}_h^\dagger(\vec{k}')\} = \delta_{hh'}\delta(\vec{k}-\vec{k}'), \quad \{\mathbf{b}_h(\vec{k}),\mathbf{b}_h^\dagger(\vec{k}')\} = \delta_{hh'}\delta(\vec{k}-\vec{k}'), \quad (10.54)$$

and

$$\{\mathbf{a}_h(\vec{k}),\mathbf{a}_h(\vec{k}')\} = 0, \quad \{\mathbf{b}_h(\vec{k}),\mathbf{b}_h(\vec{k}')\} = 0,$$
$$\{\mathbf{a}_h(\vec{k}),\mathbf{b}_h(\vec{k}')\} = 0, \quad \{\mathbf{a}_h(\vec{k}),\mathbf{b}_h^\dagger(\vec{k}')\} = 0.$$

These anti-commutation relations are now employed to construct the number representation as in Section 1.4.3: the vacuum state $|0\rangle$ is defined by

$$\mathbf{a}_h(\vec{k})|0\rangle = \mathbf{b}_h(\vec{k})|0\rangle = 0$$

for all $\vec{k}$ and $h = \pm$, and the state of one particle of helicity h in the state $\vec{k}$ with $\omega > 0$ is generated from it by

$$|1\rangle = \mathbf{a}_h^\dagger(\vec{k})|0\rangle,$$

while the one-particle state with helicity h and $\omega < 0$ is generated by

$$|1\rangle = \mathbf{b}_h^\dagger(\vec{k})|0\rangle.$$

There are no states with more than one particle of the same $\vec{k}$, the same helicity, and the same sign of ω. The Hamiltonian (10.50) turns out to be[14]

$$\mathbf{H} = \int d^3k\,\hbar\omega\sum_h\Big[\mathbf{a}_h^\dagger(\vec{k})\mathbf{a}_h(\vec{k}) - \mathbf{b}_h(\vec{k})\mathbf{b}_h^\dagger(\vec{k})\Big].$$

[13] **Show this.**
[14] **Check this.**

At this point it is best to discretize the $\vec{k}$-values, as we did before. Using the discretized version of (10.54), we then obtain, apart from an infinite constant, after returning to the continuous form,

$$\begin{aligned} \mathbf{H} &= \int d^3k\,\hbar\omega \sum_h \left[\mathbf{a}_h^\dagger(\vec{k})\mathbf{a}_h(\vec{k}) + \mathbf{b}_h^\dagger(\vec{k})\mathbf{b}_h(\vec{k})\right] \\ &= \int d^3k\,\hbar\omega \sum_h \left[\mathsf{N}_{h,+}(\vec{k}) + \mathsf{N}_{h,-}(\vec{k})\right], \end{aligned} \tag{10.55}$$

where $\mathsf{N}_{h,+}(\vec{k}) \stackrel{\text{def}}{=} \mathbf{a}_h^\dagger(\vec{k})\mathbf{a}_h(\vec{k})$ and $\mathsf{N}_{h,-}(\vec{k}) \stackrel{\text{def}}{=} \mathbf{b}_h^\dagger(\vec{k})\mathbf{b}_h(\vec{k})$ are the number operators for a given $\vec{k}$, helicity h, and positive and negative frequencies, respectively. Each of the particles with wave vector $\vec{k}$ produced by the field thus has the positive energy $\hbar\omega$. You should particularly note *that this is the case only because of the employment of anti-commutation relations.* Had we instead used commutation relations, the negative-frequency particles would have had negative energy. By contrast, recall that for the electromagnetic field we had to use commutation relations: these are specific instances of the spin-statistics connection.

The total momentum operator is also easily calculated by means of (2.67),[15] with the result

$$\vec{\mathbf{P}} = \int d^3k\,\hbar\vec{k} \sum_h \left[\mathsf{N}_{h,+}(\vec{k}) + \mathsf{N}_{h,-}(\vec{k})\right]. \tag{10.56}$$

Thus we have found that, for each given $\vec{k}$, the particles produced by the field Ψ have the energy $E = \hbar\omega > 0$ and the momentum $\vec{p} = \hbar\vec{k}$, and since $\omega^2 = \vec{k}^2c^2 + M^2c^4/\hbar^2$, their energy and momentum are connected by the equation $E^2 = \vec{p}^2c^2 + M^2c^4$, which implies that they have the mass M and satisfy the correct relativistic energy-momentum relation. Furthermore, the physical meaning of the helicity $\vec{\Sigma}\cdot\widehat{k} = \vec{\Sigma}\cdot\widehat{p}$ is the projection of the spin on the particle's momentum direction, a coordinate-invariant use of the spin projection.

The general one-particle configuration-space wave function is defined by

$$\psi(\vec{r},t) \stackrel{\text{def}}{=} \langle\, |\widetilde{\Psi}^\dagger(\vec{r},t)|0\rangle,$$

(the transpose has to be taken in order to have the spinor ψ come out as a column vector) and as a result of (10.48), the equation satisfied by the one-particle wave function is[16]

$$(\hbar\partial\!\!\!/^* + Mc)\psi = 0.$$

[15] **Do it.**
[16] **Show this.**

If the four γ-matrices $\gamma_0, \ldots, \gamma_3$ are taken to be real, as we chose them, this becomes

$$(\hbar \not{\partial} + Mc)\psi = 0; \tag{10.57}$$

for other choices, the γ-matrices in (10.57) and (10.48) differ by a canonical transformation without physical effect.

Equation (10.57) can be rewritten in the form of a Schrödinger equation after multiplying it by γ_0 and defining $\vec{\alpha} \stackrel{\text{def}}{=} \gamma_0 \vec{\gamma} = \vec{\alpha}^\dagger$ and $\beta \stackrel{\text{def}}{=} -i\gamma_0 = \beta^\dagger$,

$$i\hbar \frac{\partial}{\partial t}\psi = \left(i\hbar c \vec{\alpha} \cdot \nabla - \beta M c^2\right) \psi = \mathsf{H}\psi \tag{10.58}$$

if the Hamiltonian is defined as

$$\mathsf{H} \stackrel{\text{def}}{=} i\hbar c \vec{\alpha} \cdot \nabla - \beta M c^2 = -c\vec{\alpha} \cdot \vec{\mathbf{p}} - \beta M c^2, \tag{10.59}$$

which is not to be confused with the field Hamiltonian $\mathbf{H}$ of (10.55), and where $\vec{\mathbf{p}}$ is the momentum operator of the particle. The vector $\vec{\alpha}$ has the components $\alpha_i = -\Sigma_{4i}$, so that it is the generator of a relativistic boost in the direction i, and $\{\alpha_i, \alpha_j\} = 2\delta_{ij}$. Since $\alpha_i^2 = \mathbb{1}$, the eigenvalues of α_i are ± 1.

Equation (10.58) may be treated just like the nonrelativistic Schrödinger equation, and we may calculate the velocity of the particle by

$$\frac{d\vec{\mathbf{q}}}{dt} = \frac{1}{i\hbar}[\vec{\mathbf{q}}, \mathsf{H}] = -\frac{c}{i\hbar}[\vec{\mathbf{q}}, \vec{\alpha} \cdot \vec{\mathbf{p}}] = -c\vec{\alpha},$$

which implies that, remarkably enough, the eigenvalues of any component of the particle's velocity are $\pm c$. Whenever an instantaneous measurement of the speed of an electron is made, it can therefore come out only as the velocity of light, in one direction or the other. You have to remember that relativistically, the velocity is not simply a multiple of the momentum, and in order to measure it, two closely spaced, precise position measurements have to be made, which renders the particle's corresponding momentum quite uncertain. (If all momenta are equally likely, the expectation or average value of the momentum is infinite.) The expectation value of any component of its velocity, which of course is always less than that of light, comes about as a result of a rapidly oscillatory motion (which has come to be known by the German word *Zitterbewegung*) over a distance equal to the electron's Compton wavelength $\hbar/Mc$.

The Heisenberg equation of motion for the velocity of the electron is given by

$$i\hbar\dot{\vec{\alpha}} = [\vec{\alpha}, \mathsf{H}] = 2\vec{\alpha}\mathsf{H} + 2c\vec{\mathbf{p}},$$

so that $i\hbar\ddot{\vec{\alpha}} = 2\dot{\vec{\alpha}}\mathsf{H}$; therefore

$$\dot{\vec{\alpha}} = \dot{\vec{\alpha}}_0 \, e^{2\mathsf{H}t/i\hbar},$$

and as a result,

$$\vec{\alpha} = \frac{i\hbar}{2}\dot{\vec{\alpha}}\mathsf{H}^{-1} - c\vec{\mathbf{p}}\mathsf{H}^{-1} = \frac{i\hbar}{2}\dot{\vec{\alpha}}_0\, e^{2\mathsf{H}t/i\hbar}\mathsf{H}^{-1} - c\vec{\mathbf{p}}\mathsf{H}^{-1}.$$

Consequently,

$$\dot{\vec{\mathbf{q}}} = -c\vec{\alpha} = -\frac{i\hbar c}{2}\dot{\vec{\alpha}}_0\, e^{2\mathsf{H}t/i\hbar}\mathsf{H}^{-1} + c^2\vec{\mathbf{p}}\mathsf{H}^{-1}.$$

The first term on the right-hand side oscillates rapidly with an amplitude equal to half the Compton wavelength. Any practical velocity measurement takes much longer than the period $\hbar/2Mc^2 \sim 10^{-21}$sec of this oscillation, so that the first term averages to zero and the result of the measurement is only the second term, which has the classical value $\dot{\vec{\mathbf{q}}} = c^2\vec{\mathbf{p}}\mathsf{H}^{-1}$. It is the rapidly oscillating first part that gives each component of the velocity the eigenvalues $\pm c$.

The adjoint equation of (10.57) reads

$$\bar{\psi}(\hbar\partial\!\!\!/_{\leftarrow} - Mc) = 0,$$

and if this equation, multiplied by ψ, is added to (10.57), multiplied on the left by $\bar{\psi}$, the result is

$$\partial^\mu\bar{\psi}\gamma_\mu\psi = 0, \tag{10.60}$$

which has the form of a current conservation equation $\nabla\cdot\vec{j} + \partial\rho/\partial t = 0$ if the relativistic probability-current density four-vector is defined by $j_\mu \stackrel{\text{def}}{=} \langle ic\gamma_\mu\rangle$, so that $\vec{j} \stackrel{\text{def}}{=} \langle ic\vec{\gamma}\rangle$ and $\rho \stackrel{\text{def}}{=} \langle i\gamma_0\rangle = \psi^\dagger\psi$ for normalized ψ.

Solving the free Dirac equation

Up to this point we have had no need to write down the γ-matrices explicitly. All that was ever required were the commutation and anti-commutation relations implied by (10.25) and those of the rotation generators $\Sigma_{\mu\nu}$ for spin-1/2, though it was sometimes convenient to assume the four matrices $\gamma_0,\ldots,\gamma_3$ to be real. Let us now, instead, choose the matrices $\vec{\alpha}$ and β to have the form

$$\vec{\alpha} = i\begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \tag{10.61}$$

which implies that[17]

$$\vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \qquad \vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}. \tag{10.62}$$

(Here $\mathbb{1}$ denotes the 2×2 unit matrix, and the components of $\vec{\sigma}$ are the Pauli spin matrices.)

---

[17] **Show this; also show that these matrices satisfy the correct commutation and anti-commutation relations.**

After setting $\psi(\vec{r},t) = u\exp[i(\vec{p}\cdot\vec{r} - p_0 t)/\hbar]$, (10.58) turns into the set of algebraic equations for u

$$(\vec{\alpha}\cdot\vec{p} + \beta Mc)u = -\frac{p_0}{c}u.$$

Write the four-component spinor u in terms of two two-component spinors v and w,

$$u \stackrel{\text{def}}{=} \begin{pmatrix} v \\ w \end{pmatrix},$$

and you obtain the two equations

$$i\vec{\sigma}\cdot\vec{p}w + Mcv = -\frac{p_0}{c}v, \qquad i\vec{\sigma}\cdot\vec{p}v + Mcw = \frac{p_0}{c}w,$$

which imply that $p_0^2 = M^2c^4 + \vec{p}^2c^2$; using the subscripts $\pm$ for the helicity, $\vec{\Sigma}\cdot\vec{p}u_\pm = \pm|\vec{p}|u_\pm$, we obtain the equations[18]

$$v_\pm = \mp\frac{ic|\vec{p}|}{|p_0| + Mc^2}w_\pm, \qquad \text{for } \omega = p_0/\hbar > 0,$$

$$w_\pm = \mp\frac{ic|\vec{p}|}{|p_0| + Mc^2}v_\pm, \qquad \text{for } \omega = p_0/\hbar < 0.$$

In the nonrelativistic limit, the momentum is small compared to the rest mass (times c), $|\vec{p}| \ll Mc$. Therefore for $\omega > 0$, $\| v \| \ll \| w \|$, while for $\omega < 0$ it's the other way around, $\| w \| \ll \| v \|$; the upper two components of a spinor u (in this representation) are thus referred to as the "small components" when $\omega > 0$, whereas for $\omega < 0$, the lower two are the "small components." The large components are then the eigenvectors of the ordinary, nonrelativistic, two-component spin in the momentum direction, $\vec{\sigma}\cdot\hat{p}$.

10.2.1 Electrons in an electromagnetic field

In the presence of an electromagnetic field generated by the potentials $\vec{A}$ and ϕ, these potentials have to enter the matter-field equations in the usual manner together with the derivatives, so that $i\hbar\nabla$ is replaced by $i\hbar\nabla + \frac{e}{c}\vec{A}$ and $i\hbar\frac{\partial}{\partial t}$ by $i\hbar\frac{\partial}{\partial t} - e\phi$. Since the four quantities ϕ and $\vec{A}$ behave under Lorentz transformations like the components of a four-vector, this means that if we define $A^0 \stackrel{\text{def}}{=} -\phi$, the Dirac equation for the field Ψ, in the presence of the electromagnetic potentials A^μ reads

$$[\gamma_\mu(\hbar\partial^\mu - i\frac{e}{c}A^\mu) + Mc]\Psi = 0, \tag{10.63}$$

[18] **Check this.**

and the Dirac equation for the one-particle wave function of the quanta produced by the field reads similarly,

$$[\gamma_\mu(\hbar\partial^\mu - i\frac{e}{c}A^\mu) + Mc]\psi = 0, \tag{10.64}$$

which implies that the quanta of the matter field are coupled to the electromagnetic field and have the electric charge e.

The new version of (10.49) in the presence of the electromagnetic potentials is obtained by multiplying (10.64) by $\gamma_\mu(\hbar\partial^\mu - i\frac{e}{c}A^\mu) - Mc$, so that if ψ solves (10.64) it also satisfies the equation

$$\left[\gamma_\mu(\hbar\partial^\mu - i\frac{e}{c}A^\mu)\gamma_\nu(\hbar\partial^\nu - i\frac{e}{c}A^\nu) - M^2c^2\right]\psi = 0.$$

Use of (10.25) and the equations (2.92) connecting the potentials to the electric and magnetic fields leads to[19]

$$\begin{aligned}
&\gamma_\mu(\hbar\partial^\mu - i\frac{e}{c}A^\mu)\gamma_\nu(\hbar\partial^\nu - i\frac{e}{c}A^\nu)\\
=\;& g_{\mu\nu}(\hbar\partial^\mu - i\frac{e}{c}A^\mu)(\hbar\partial^\nu - i\frac{e}{c}A^\nu) + \frac{i}{2}\Sigma_{\mu\nu}[(\hbar\partial^\mu - i\frac{e}{c}A^\mu),(\hbar\partial^\nu - i\frac{e}{c}A^\nu)]\\
=\;& -g_{\mu\nu}(-i\hbar\partial^\mu - \frac{e}{c}A^\mu)(-i\hbar\partial^\nu - \frac{e}{c}A^\nu) + \frac{e\hbar}{c}\vec{\Sigma}\cdot\vec{B} + \frac{ie\hbar}{c}\vec{\alpha}\cdot\vec{E}.
\end{aligned}$$

Therefore ψ has to solve the equation

$$[-M^2c^4 - g_{\mu\nu}(c\mathbf{p}^\mu - eA^\mu)(c\mathbf{p}^\nu - eA^\nu) + ec\hbar\vec{\Sigma}\cdot\vec{B} + iec\hbar\vec{\alpha}\cdot\vec{E}]\psi = 0$$

if we define $\mathbf{p}^0 \stackrel{\text{def}}{=} -i\hbar\partial/(c\partial t)$. In the nonrelativistic approximation the last term in the bracket is dropped because $\vec{\alpha}\sim\vec{v}/c$; moreover,

$$\begin{aligned}
&(c\mathbf{p}^0 - e\phi)^2 - M^2c^4 - (c\vec{\mathbf{p}} - e\vec{A})^2 + ec\hbar\vec{\Sigma}\cdot\vec{B}\\
=\;& \left[c\mathbf{p}^0 - e\phi - \sqrt{M^2c^4 + (c\vec{\mathbf{p}} - e\vec{A})^2 - ec\hbar\vec{\Sigma}\cdot\vec{B}}\right]\\
&\qquad\times\left[c\mathbf{p}^0 - e\phi + \sqrt{M^2c^4 + (c\vec{\mathbf{p}} - e\vec{A})^2 - ec\hbar\vec{\Sigma}\cdot\vec{B}}\right]\\
&\sim (Mc^2 + \ldots)\left[c\mathbf{p}^0 - e\phi - Mc^2 - \frac{(c\vec{\mathbf{p}} - e\vec{A})^2}{2Mc^2} + \frac{e\hbar}{2Mc}\vec{\Sigma}\cdot\vec{B} + \ldots\right]\\
&\to Mc^2\left[E_{\text{nr}} - e\phi - \frac{(c\vec{\mathbf{p}} - e\vec{A})^2}{2Mc^2} + \frac{e\hbar}{2Mc}\vec{\Sigma}\cdot\vec{B}\right]
\end{aligned}$$

where $E_{\text{nr}} \stackrel{\text{def}}{=} cp^0 - Mc^2$ is the nonrelativistic energy. The nonrelativistic time-independent Schrödinger equation obtained from (10.64) thus reads

$$\left[\frac{1}{2M}(\vec{\mathbf{p}} - \frac{e}{c}\vec{A})^2 - \frac{e\hbar}{2Mc}\vec{\sigma}\cdot\vec{B} + e\phi\right]\psi = E\psi,$$

[19] **Do the details of this calculation.**

which means that the Hamiltonian agrees with (7.14). The Dirac equation thus automatically leads to the correct Schrödinger equation for spin-1/2 particles in the nonrelativistic limit, including the Pauli term with the Bohr-magneton strength and the correct gyromagnetic ratio.

Next, let us check what charge conjugation accomplishes, when applied to (10.63). Using (10.43) and (10.44), we find[20] that if Ψ satisfies (10.63), then its charge conjugate Ψ^c satisfies the equation

$$[\gamma_\mu(\hbar\partial^\mu + i\frac{e}{c}A^\mu) + Mc]\Psi^c = 0,$$

which implies that the field Ψ^c gives rise to particles whose wave functions satisfy (10.64) with the same mass but with the opposite sign of the charge. (That is why the transformation (10.43) is called charge conjugation.) Thus, the Dirac equation automatically contains the prediction of the existence of positrons (though this was recognized as a prediction only after the experimental discovery of positrons).

Here, then, is a summary of the remarkable features of Dirac's equation:

- It is form-invariant under Lorentz transformations.
- It describes a field that gives rise to spin-1/2 particles with the correct gyromagnetic ratio.
- Even though it has solutions of positive and negative frequency, the energy of its solutions is always positive, provided its solutions are assumed to obey anti-commutation relations and thus give rise to fermions; therefore it incorporates the spin-statistics connection.
- It contains within it the prediction of the existence of positrons.

Historical note: When Dirac invented his equation, it was in the form of the wave equation (10.64) rather than for a field. He therefore had to struggle with the problem of how to deal with the negative-frequency solutions, which appeared to have negative energies because $E = \hbar\omega$. His solution to this quandary was the ingenious *hole theory:* all the negative-energy states are normally occupied by an infinite sea of electrons, which as fermions obey Pauli's exclusion principle, so that under ordinary circumstances these negative-energy states are unavailable to the electrons we observe. When one of these states happens to be empty, creating a hole in the sea of negative-energy states, this hole, as an *absence* of a particle of negative energy and negative charge, acts like a particle of positive energy and positive charge. So the hole theory was a device to solve the negative-energy problem without using a field, but it made sense only if electrons were assumed to be fermions. Therefore, it too used the spin-statistics connection in an essential fashion. Using the Dirac equation for

[20] **Show this.**

the quantum field as a starting point, rather than as the result of "second quantization," avoids the artificiality of the hole theory and solves the negative-energy problem by means of the spin-statistics connection in the form of anti-commutation rules for the spinor field.

The Dirac equation as a field equation, together with the Maxwell equations, with the electric four-current operator[21]

$$\mathbf{J}_\mu \stackrel{\text{def}}{=} ec\bar{\Psi} i\gamma_\mu \Psi$$

as a source term (the space part of $\mathbf{J}_\mu$ is the ordinary electric current density and $\mathbf{J}_0$ is the electric charge density), regarded as equations for the electromagnetic quantum field [so that the electromagnetic potentials in (10.63) are operators] make up the enormously successful theory of quantum electrodynamics (QED). However, it took another 30 years before the difficulties arising from the infinities this theory leads to were overcome by Feynman, Schwinger, and Tomonaga, and it lies beyond the scope of this book.

10.3 The Dirac Equation for a Central Potential

In the presence of a static electric field generated by a central potential $\phi(r)$, the time-independent Dirac equation for the wave function of an electron reads

$$\mathsf{H}_\phi \psi \stackrel{\text{def}}{=} [-c\vec{\alpha}\cdot\vec{\mathbf{p}} - \beta Mc^2 + e\phi(r)]\psi = E\psi. \tag{10.65}$$

The orbital angular momentum $\hbar\vec{\mathcal{L}}$ now is not a constant of the motion, but the total angular momentum $\hbar\vec{\mathcal{J}} = \hbar(\vec{\mathcal{L}} + \frac{1}{2}\vec{\Sigma})$ is: since ϕ and $\vec{\alpha}\cdot\vec{\mathbf{p}}$ are invariant under rotations, so is the entire Hamiltonian,

$$[\vec{\mathcal{J}}, \mathsf{H}_\phi] = 0,$$

and the total angular momentum is conserved. As usual, the eigenvalues of $\vec{\mathcal{J}}^2$ are $j(j+1)$ with $j = 1/2, 3/2, \ldots$. Now, it follows from Eq. (5.70) that $\vec{\mathcal{J}}^2 + \frac{1}{4} = (\mathbb{1} + \vec{\Sigma}\cdot\vec{\mathcal{L}})^2$, which implies that $(\mathbb{1} + \vec{\Sigma}\cdot\vec{\mathcal{L}})^2$, but not necessarily $\mathbb{1} + \vec{\Sigma}\cdot\vec{\mathcal{L}}$ itself, is a constant of the motion. However, it turns out that the Hermitian operator

$$\mathbf{K} \stackrel{\text{def}}{=} \beta(\mathbb{1} + \vec{\mathcal{L}}\cdot\vec{\Sigma})$$

does commute with the Hamiltonian H_ϕ.

Since $\vec{\mathcal{L}}\cdot\vec{\Sigma}$ commutes with β and with $\phi(r)$, it is only necessary to show that $[\beta(\vec{\mathcal{L}}\cdot\vec{\Sigma}+\mathbf{1}), \vec{\alpha}\cdot\vec{\mathbf{p}}] = 0$ in order to prove that $[\mathbf{K}, \mathsf{H}_\phi] = 0$. Furthermore, since β anti-commutes

[21] **Show that this current operator changes sign under charge conjugation.**

with $\vec{\alpha}$, what needs to be shown is that $\{(\vec{\mathcal{L}}\cdot\vec{\Sigma}+\mathbf{1}),\vec{\alpha}\cdot\vec{\mathbf{p}}\}=0$. So let us calculate $R\stackrel{\text{def}}{=}\{\vec{\mathcal{L}}\cdot\vec{\Sigma},\vec{\alpha}\cdot\vec{\mathbf{p}}\}$.

Since $[\vec{\mathcal{J}},\vec{\alpha}\cdot\vec{\mathbf{p}}]=0$ and $\vec{\mathcal{J}}\cdot\vec{\Sigma}=\vec{\mathcal{L}}\cdot\vec{\Sigma}+\frac{3}{2}$, we have $R=\{\vec{\mathcal{J}}\cdot\vec{\Sigma}-\frac{3}{2},\vec{\alpha}\cdot\vec{\mathbf{p}}\}=\vec{\mathcal{J}}\cdot\{\vec{\Sigma},\vec{\alpha}\cdot\vec{\mathbf{p}}\}-3\vec{\alpha}\cdot\vec{\mathbf{p}}$. Next, use $\vec{\alpha}=-\gamma_5\vec{\Sigma}$ and obtain $\{\vec{\Sigma},\vec{\alpha}\cdot\vec{\mathbf{p}}\}=-\gamma_5\{\vec{\Sigma},\vec{\Sigma}\cdot\vec{\mathbf{p}}\}=-2\gamma_5\vec{\mathbf{p}}$, as a result of which $\vec{\mathcal{J}}\cdot\{\vec{\Sigma},\vec{\alpha}\cdot\vec{\mathbf{p}}\}=-2\gamma_5(\vec{\mathcal{L}}+\frac{1}{2}\vec{\Sigma})\cdot\vec{\mathbf{p}}=\vec{\alpha}\cdot\vec{\mathbf{p}}$. Therefore, $R=-2\vec{\alpha}\cdot\vec{\mathbf{p}}$, so that $\{\vec{\mathcal{L}}\cdot\vec{\Sigma}+\mathbf{1},\vec{\alpha}\cdot\vec{\mathbf{p}}\}=0$ and consequently, $[\beta(\vec{\mathcal{L}}\cdot\vec{\Sigma}+\mathbf{1}),\vec{\alpha}\cdot\vec{\mathbf{p}}]=0$, which proves that $[\mathbf{K},\mathsf{H}_\phi]=0$.

According to (5.70), $\mathbf{K}^2=(\mathbb{1}+\vec{\mathcal{L}}\cdot\vec{\Sigma})^2=\vec{\mathcal{J}}^2+\frac{1}{4}$, so that the eigenvalues κ^2 of $\mathbf{K}^2$ are $j(j+1)+\frac{1}{4}=(j+\frac{1}{2})^2$, which implies that the eigenvalues of $\mathbf{K}$ are $\kappa=\pm(j+\frac{1}{2})$, and the values of $|\kappa|$ are the positive integers. Moreover, since[22] $\mathbf{K}^2=\vec{\mathcal{L}}^2+\mathbb{1}+\vec{\mathcal{L}}\cdot\vec{\Sigma}$, we have

$$\vec{\mathcal{L}}^2=\mathbf{K}^2-\beta\mathbf{K}.$$

Therefore, in the representation (10.61), when $\beta=1$ (the upper components of the four-spinor), then for $\kappa=j+\frac{1}{2}$ we must have $l(l+1)=(j+\frac{1}{2})^2-(j+\frac{1}{2})=(j+\frac{1}{2})(j-\frac{1}{2})$, so $l=j-\frac{1}{2}=\kappa-1$, and for $\kappa=-(j+\frac{1}{2})$ we have $l(l+1)=(j+\frac{1}{2})^2+(j+\frac{1}{2})=(j+\frac{1}{2})(j+\frac{3}{2})$, so $l=j+\frac{1}{2}=-\kappa$; when $\beta=-1$ (the lower components) we find similarly that for $\kappa=j+\frac{1}{2}$, $l=j+\frac{1}{2}=\kappa$ and for $\kappa=-(j+\frac{1}{2})$, $l=j-\frac{1}{2}=-(\kappa+1)$. So here is the table of values.

β	κ	j	$\vec{\Sigma}\cdot\vec{\mathcal{L}}$
1	$l+1$	$l+\frac{1}{2}$	$\kappa-1$
1	$-l$	$l-\frac{1}{2}$	$\kappa-1$
-1	l	$l-\frac{1}{2}$	$-(\kappa+1)$
-1	$-(l+1)$	$l+\frac{1}{2}$	$-(\kappa+1)$

Note that β is not a constant of the motion, but in the nonrelativistic limit it is: In the representation (10.61), $\beta=1$ for the upper components and $\beta=-1$ for the lower ones. So if the small components are set equal to zero, as they are in the nonrelativistic approximation, the spinors u are eigenfunctions of β. For positive frequencies, the lower two components are the large ones and therefore $\beta=-1$ in the nonrelativistic limit; for negative frequencies, the upper components are large, so that $\beta=+1$. As a result, in the nonrelativistic limit the sign of the eigenvalue of $\mathbf{K}$ is in a one-to-one relation with the two cases $l=j\pm\frac{1}{2}$, but in general there is no such correspondence.

To expand the angle dependence of solutions of (10.65) we use the two-component spinor angle functions $\mathfrak{Z}^m_{lj}$ defined by (5.73) and (5.74). How-

[22] **Show this.**

ever, they will be relabeled, using κ instead of l,

$$\widehat{\mathfrak{Z}}^m_{\kappa j} \stackrel{\text{def}}{=} \begin{cases} \mathfrak{Z}^m_{j-\frac{1}{2}\,j} & \text{when} \quad \kappa = j + \frac{1}{2} \\ \mathfrak{Z}^m_{j+\frac{1}{2}\,j} & \text{when} \quad \kappa = -(j + \frac{1}{2}) \end{cases},$$

so that

$$(\vec{\sigma} \cdot \vec{\mathcal{L}} + \mathbb{1})\widehat{\mathfrak{Z}}^m_{\kappa j} = \kappa \widehat{\mathfrak{Z}}^m_{\kappa j}.$$

Since, furthermore,[23] with $\widehat{r} \stackrel{\text{def}}{=} \vec{r}/r$, $\{\vec{\Sigma} \cdot \widehat{r}, \vec{\sigma} \cdot \vec{\mathcal{L}} + \mathbb{1}\} = 0$, it follows that

$$\vec{\sigma} \cdot \widehat{r}\widehat{\mathfrak{Z}}^m_{\kappa j} = \widehat{\mathfrak{Z}}^m_{-\kappa j}.$$

We can therefore write a four-component spinor eigenfunction of $\mathbf{K}$ that is also a function of r in the form

$$\Psi_\kappa(r) = \frac{1}{r} \begin{pmatrix} F(r)\widehat{\mathfrak{Z}}^m_{\kappa j} \\ -G(r)\vec{\sigma} \cdot \widehat{\mathbf{r}}\widehat{\mathfrak{Z}}^m_{\kappa j} \end{pmatrix},$$

insertion of which in (10.65), together with the fact that

$$\vec{\sigma} \cdot \vec{\mathbf{p}}\, \vec{\sigma} \cdot \vec{r} = \vec{\mathbf{p}} \cdot \vec{r} - i\hbar\vec{\sigma} \cdot \vec{\mathcal{L}} = -3i\hbar\mathbb{1} - i\hbar r\frac{\partial}{\partial r} - i\hbar\vec{\sigma} \cdot \vec{\mathcal{L}},$$

$$\vec{\sigma} \cdot \vec{r}\, \vec{\sigma} \cdot \vec{\mathbf{p}} = \vec{r} \cdot \vec{\mathbf{p}} + i\hbar\vec{\sigma} \cdot \vec{\mathcal{L}} = -i\hbar r\frac{\partial}{\partial r} + i\hbar\vec{\sigma} \cdot \vec{\mathcal{L}},$$

yields the two coupled equations[24]

$$\left(\frac{d}{dr} + \frac{\kappa}{r}\right) G = \left(a_1 - \frac{e\phi}{\hbar c}\right) F, \qquad \left(\frac{d}{dr} - \frac{\kappa}{r}\right) F = \left(a_2 + \frac{e\phi}{\hbar c}\right) G,$$

where

$$a_1 \stackrel{\text{def}}{=} \frac{Mc^2 + E}{\hbar c}, \qquad a_2 \stackrel{\text{def}}{=} \frac{Mc^2 - E}{\hbar c}.$$

Defining

$$a \stackrel{\text{def}}{=} \sqrt{a_1 a_2} = \sqrt{M^2c^4 - E^2}/\hbar c$$

and the dimensionless radial distance $\rho \stackrel{\text{def}}{=} ar$, we obtain

$$\left(\frac{d}{d\rho} + \frac{\kappa}{\rho}\right) G = \left(\frac{a_1}{a} - \frac{e\phi}{a\hbar c}\right) F, \tag{10.66}$$

$$\left(\frac{d}{d\rho} - \frac{\kappa}{\rho}\right) F = \left(\frac{a_2}{a} + \frac{e\phi}{a\hbar c}\right) G. \tag{10.67}$$

[23] **Prove this.**

[24] **Show this as an exercise.**

In the free case, when $\phi = 0$, these two equations can be uncoupled[25] to read

$$-G'' + \frac{\kappa(\kappa+1)}{\rho^2}G = aG, \qquad -F'' + \frac{\kappa(\kappa-1)}{\rho^2}F = aF,$$

and $a = |\vec{p}|/\hbar = k$, so that $G = \rho j_\kappa(\rho)$ and $F = \rho j_{\kappa-1}(\rho)$. For general potentials, however, Eqs. (10.66) and (10.67) cannot be uncoupled.

10.3.1 The hydrogen atom

For a nuclear point charge of strength Ze, the Coulomb potential is given by $\phi = Ze/r$, and the potential energy seen by an electron is $e\phi = -Ze^2/r$ or

$$\frac{e\phi}{a\hbar c} = -\frac{Ze^2}{\hbar car} = -\frac{\gamma}{\rho}, \qquad \gamma \stackrel{\text{def}}{=} \frac{Ze^2}{\hbar c} = \alpha Z,$$

where $\alpha \stackrel{\text{def}}{=} e^2/\hbar c$ is the fine structure constant.

We proceed as in the nonrelativistic case, defining $F(\rho) \stackrel{\text{def}}{=} f(\rho)e^{-\rho}$ and $G(\rho) \stackrel{\text{def}}{=} g(\rho)e^{-\rho}$, thereby obtaining the equations for f and g,

$$g' - g + \frac{\kappa}{\rho}g = \left(\frac{a_1}{a} + \frac{\gamma}{\rho}\right)f, \qquad f' - f - \frac{\kappa}{\rho}f = \left(\frac{a_2}{a} - \frac{\gamma}{\rho}\right)g.$$

The Frobenius ansatz $f = \rho^s \sum_0^\infty c_n\rho^n$, and $g = \rho^s \sum_0^\infty b_n\rho^n$, leads to the recursion relations[26] for $n > 0$,

$$b_n(n+s+\kappa) - b_{n-1} = \gamma c_n + \frac{a_1}{a}c_{n-1} \tag{10.68}$$

$$c_n(n+s-\kappa) - c_{n-1} = -\gamma b_n + \frac{a_2}{a}b_{n-1}, \tag{10.69}$$

and for $n = 0$ to the indicial equations

$$(s+\kappa)b_0 - \gamma c_0 = 0, \qquad \gamma b_0 + (s-\kappa)c_0 = 0,$$

which require that $s^2 - \kappa^2 + \gamma^2 = 0$ and therefore for the solution regular at the origin,

$$s = +\sqrt{\kappa^2 - \gamma^2}.$$

Notice that, in contrast to the nonrelativistic Coulomb wave functions, the solutions are not polynomials, and they may even be singular at the origin (when $s < 1$).[27] Now multiply (10.68) by a, (10.69) by a_1, and add the

[25] **Do it.**

[26] **Derive them.**

[27] If $s < \frac{1}{2}$, the derivative of the wave function is not square integrable at the origin, so that the expectation value of the momentum does not exist. For $\kappa = 1$, to prevent this requires that $\gamma < \frac{1}{2}\sqrt{3}$, i.e., $Z < 119$. For heavier (point) nuclei the Dirac equation would lead to problems.

results to obtain

$$b_n[a(n+s+\kappa)-a_1\gamma] = c_n[a_1(n+s-\kappa)+a\gamma],$$

which, together with (10.68), yields the recursion relation

$$b_n\left[(n+2+\kappa)-\gamma\frac{a(n+s+\kappa)-a_1\gamma}{a_1(n+s-\kappa)+a\gamma}\right] = b_{n-1}\left[1+\frac{a_1}{a}\frac{a(n-1+s+\kappa)-a_1\gamma}{a_1(n-1+s-\kappa)+a\gamma}\right],$$

implying that for $n\to\infty$

$$\frac{b_n}{b_{n-1}} \simeq \frac{c_n}{c_{n-1}} \simeq \frac{2}{n}.$$

Therefore the tail end of the series behaves like $e^{2\rho}$, which means that the solutions increase exponentially unless the series terminates. (That is, for real values of a; for $E > Mc^2$ a is imaginary and there is no need for the series to break up: these are the scattering states.) It follows from the recursion relations that both series must then end at the same point, and if $b_{n'}$ is the last non-zero term, we must have

$$2a(n'+s) = (a_1-a_2)\gamma,$$

which, with the explicit values of a, a_1, a_2, and γ, leads to the energy levels[28]

$$E = Mc^2\left[1+\frac{\alpha^2Z^2}{(n'+\sqrt{\kappa^2-\alpha^2Z^2})^2}\right]^{-\frac{1}{2}}. \tag{10.70}$$

The ground-state energy of a hydrogenic atom is therefore

$$E_{\min} = Mc^2\sqrt{1-\alpha^2Z^2},$$

as compared to the nonrelativistic ground-state energy $E^{\mathrm{NR}}_{\min} = -\frac{1}{2}\alpha^2Z^2Mc^2$ (which, of course, does not include the rest energy Mc^2). According to (10.70) the energies are independent of the sign of κ, which is the analogue of the nonrelativistic l-degeneracy. For $\alpha Z \ll 1$ the energy eigenvalues may be expanded in powers of αZ, giving[29]

$$E = Mc^2\left[1-\frac{\alpha^2Z^2}{2n^2}-\frac{1}{2}\frac{\alpha^4Z^4}{|\kappa|n^3}+\frac{3}{8}\frac{\alpha^4Z^4}{n^4}+\ldots\right],$$

where $n \overset{\text{def}}{=} n'+|\kappa|$, or

$$E-Mc^2 = -\frac{\alpha^2Z^2Mc^2}{2n^2}-\frac{\alpha^4Z^4Mc^2}{2n^3}\left(\frac{1}{|\kappa|}-\frac{3}{4n}\right)+\ldots, \tag{10.71}$$

[28] **Show this.**
[29] **Check this.**

which agrees with the previously calculated value (4.70) together with (7.26) for the fine structure.

Note that κ can take on positive or negative values, *except* when $n' = 0$. In that particular case (10.68) says that $c_0/b_0 = -a/a_1 < 0$ and the indicial equation says $\gamma c_0/b_0 = \kappa + s = \kappa + \sqrt{\kappa^2 - \gamma^2}$; therefore only $\kappa < 0$ is allowed for $n' = 0$, which makes the ground state non-degenerate, with $n' = 0$ and $\kappa = -1$. The energy of the first excited states is obtained from (10.70) by setting $n' = 1$, and there still is a twofold degeneracy, with $\kappa = \pm 1$, just as in the fine-structure correction to the nonrelativistic energy. As mentioned at the end of Section 7.3.1, this degeneracy is finally broken by the Lamb shift and explained by the "radiative corrections" of quantum electrodynamics.

For the calculation of the hydrogenic wave functions by means of the Dirac equation, see [Rose] [Note, however, that Rose uses a representation of $\vec{\alpha}$ that differs from (10.61).]

10.4 Problems and Exercises

1. Prove the following equations for the reflection operators γ_μ in five dimensions:

$$\frac{1}{2}\{\gamma_\mu, \gamma_\nu\} = \delta_{\mu\nu},$$

$$\gamma_\mu^{-1}\Sigma_{\nu\lambda}\gamma_\mu = \begin{cases} -\Sigma_{\nu\lambda} & \text{for } \mu = \lambda \text{ or } \mu = \nu \\ \Sigma_{\nu\lambda} & \text{for } \mu \neq \lambda, \mu \neq \nu. \end{cases}$$

and

$$\frac{1}{2i}[\gamma_\mu, \Sigma_{\lambda\nu}] = \delta_{\mu\nu}\gamma_\lambda - \delta_{\mu\lambda}\gamma_\nu.$$

2. Prove (10.31).

3. Prove (10.33).

4. Prove (10.34).

5. Prove (10.36) for states of two particles that are symmetric under exchange.

6. Prove that if the matrix M commutes with all the matrices γ_μ, $\mu = 1, \ldots, 4$, then M must be a multiple of the unit matrix.

7. Show that the charge conjugate spinors u^c are spinors under proper Lorentz transformations.

8. From the relativistic relation between the velocity and the momentum of a particle, find the velocity probability distribution if the momentum is uniformly distributed. Show from this that if the momentum is uniformly distributed, then $\langle v \rangle = c$.

9. (a) Show that if u_0 is the relativistic spinor of an electron in its rest system, then $\bar{u}_0\gamma_j\gamma_0 u_0 = 0$ for $j = 1, 2, 3$.

   (b) Supposing that the spin of the electron in its rest system is in the x-direction, calculate $\langle \vec{\Sigma} \rangle$ in a reference frame in which the electron has the momentum p in the z-direction.

Part II

Mathematical Appendices

This page intentionally left blank

A
The Dirac Delta Function

The Dirac delta function is defined by the property that for any continuous function $f(x)$ on the interval (a, b) with $a < 0 < b$,

$$\int_a^b dx\, f(x)\delta(x) = f(0). \tag{A.1}$$

Of course, there really is no such function; $\delta(x)$ is, instead, a *generalized function* or a *distribution*.[1] It may be thought of as an idealization of an infinitely sharp spike of area 1, concentrated at the point 0, physically exemplified by a sharp pulse at the time $t = 0$. For example, we may define the function

$$\delta_\epsilon(x) \stackrel{\text{def}}{=} \frac{i}{2\pi}\left[\frac{1}{x+i\epsilon} - \frac{1}{x-i\epsilon}\right] = \frac{\epsilon/\pi}{x^2+\epsilon^2} \tag{A.2}$$

and the integral as the limit, to be taken *after* the integral is carried out,

$$\int_{-1}^{1} dx\, \delta(x)f(x) \stackrel{\text{def}}{=} \lim_{\epsilon\to 0}\int_{-1}^{1} dx\, \delta_\epsilon(x)f(x) = f(0)\lim_{\epsilon\to 0}\frac{2}{\pi}\cot^{-1}\epsilon = f(0).$$

In a similar manner, the *Cauchy principal value* is defined by

$$\mathcal{P}\frac{1}{x} \stackrel{\text{def}}{=} \frac{1}{2}\left[\frac{1}{x+i\epsilon} + \frac{1}{x-i\epsilon}\right] \tag{A.3}$$

[1] For more details, see, for example, [Lieb], Chapter 6.

in the sense that

$$\mathcal{P}\int_{-1}^{1} dx\,\frac{f(x)}{x} \stackrel{\text{def}}{=} \lim_{\epsilon\to 0}\int_{-1}^{1} dx\,f(x)\frac{1}{2}\left[\frac{1}{x+i\epsilon}+\frac{1}{x-i\epsilon}\right].$$

As a result we have the useful formula, in which it is understood that $0<\epsilon\ll 1$,

$$\frac{1}{x+i\epsilon} = \mathcal{P}\frac{1}{x} - i\pi\delta(x). \tag{A.4}$$

Alternatively, we may take

$$\underline{\delta}_\epsilon(x) \stackrel{\text{def}}{=} \frac{\sin(x/\epsilon)}{\pi x},$$

which is not sharply spiked at the origin but, instead, oscillates infinitely rapidly as $\epsilon\to 0$ everywhere except at the origin, which has the same effect. The Dirac delta function may also be regarded as the derivative of a unit step function, called the *Heaviside* function: $\theta(x)=0$ for $x<0$, while $\theta(x)=1$ for $x>0$. It is easily seen that $\theta'(x)=\delta(x)$, because for any $a<0<b$, by an integration by parts,

$$\begin{aligned}\int_a^b dx\,f(x)\theta'(x) &= f(b)\theta(b)-f(a)\theta(a)-\int_a^b dx\,f'(x)\theta(x)\\ &= f(b)-\int_0^b dx\,f'(x) = f(0).\end{aligned}$$

If the function $f(x)$ has a step-discontinuity at $x=0$, the effect of multiplying it by $\delta(x)$ is ambiguous and the result depends on the context. In most instances it is convenient to define

$$\int_a^b dx\,f(x)\delta(x) = \frac{1}{2}[f(0+)+f(0-)], \tag{A.5}$$

where $f(0+)$ and $f(0-)$ are the limits from above and below, respectively, and $a<0<b$.

Here are some properties of $\delta(x)$ that are very useful for formal manipulations. The first is that it is *even:*

$$\delta(-x)=\delta(x). \tag{A.6}$$

Furthermore, it follows from $\int_a^b dx\,\delta(cx) = \frac{1}{|c|}\int_{|c|a}^{|c|b} dy\,\delta(y) = \frac{1}{|c|}$ that for all real $c\neq 0$,

$$\delta(cx) = \frac{1}{|c|}\delta(x). \tag{A.7}$$

The derivative of a delta-function is defined by integration by parts:

$$\int_a^b dx\,\delta'(x)f(x) = -\int_a^b dx\,\delta(x)f'(x). \tag{A.8}$$

We also have

$$x\delta(x) = 0,$$

and

$$\delta'(-x) = -\delta'(x).$$

If $g(x)$ is monotonely increasing and passes through 0 between a and b, then a change of variables gives $\int_a^b dx\, f(x)\delta[g(x)] = \int dy\, f(x)\delta(y)/g'(x)$, where $g(x) = y$ so that $x = g^{-1}(y)$, from which we conclude that

$$\delta[g(x)] = \delta(y)/g'[g^{-1}(y)]. \tag{A.9}$$

The three-dimensional delta function $\delta^3(\vec{r}-\vec{r}')$, which is physically exemplified by the charge density of a unit point charge located at $\vec{r}'$, has to be handled carefully, depending on the coordinate system used. In Cartesian coordinates, it is simply given by

$$\delta^3(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z'),$$

but in spherical polar coordinates it has to be[2]

$$\delta^3(\vec{r} - \vec{r}') = \frac{\delta(r - r')}{r^2}\frac{\delta(\theta - \theta')}{\sin\theta}\delta(\varphi - \varphi'). \tag{A.10}$$

[2] **Why?**

This page intentionally left blank

B
Hilbert Space

B.1 Linear Vector Spaces

We begin by defining a *complex linear vector space* $\mathfrak{V}$: a collection of objects called *vectors,* with the following properties[1]:

1. To any pair of vectors Φ_1 and Φ_2 in $\mathfrak{V}$, there corresponds a third vector $\Phi \in \mathfrak{V}$, called the *sum* of Φ_1 and Φ_2, $\Phi = \Phi_1 + \Phi_2$, such that addition is (a) commutative, $\Phi_1 + \Phi_2 = \Phi_2 + \Phi_1$, and (b) associative, $\Phi_1 + (\Phi_2 + \Phi_3) = (\Phi_1 + \Phi_2) + \Phi_3$.

2. $\mathfrak{V}$ contains a unique vector 0 such that for all $\Phi \in \mathfrak{V}$, $\Phi + 0 = \Phi$.

3. For every $\Psi \in \mathfrak{V}$ there is a unique vector $-\Psi \in \mathfrak{V}$ such that $\Psi - \Psi \stackrel{\text{def}}{=} \Psi + (-\Psi) = 0$.

4. For every complex number[2] c (called a *scalar*) and every $\Psi \in \mathfrak{V}$ there is a vector $c\Psi \in \mathfrak{V}$, called the *product* of the two, such that $1\Psi = \Psi$ and multiplication is associative: $a(b\Psi) = (ab)\Psi$.

5. Multiplication by scalars is distributive over vector addition, $a(\Psi + \Phi) = a\Psi + a\Phi$ as well as over scalar addition, $(a + b)\Psi = a\Psi + b\Psi$.

---

[1] It will help your intuition always to keep in mind the analogues of these properties for ordinary real Euclidian vector spaces and to think geometrically.

[2] If these numbers are restricted to be real, $\mathfrak{V}$ is a *real linear vector space,* such as ordinary three-dimensional Euclidian space.

6. For any two vectors in $\mathfrak{V}$ an *inner product* is defined as a complex number such that

$$(\Psi, \Phi) = (\Phi, \Psi)^*, \qquad (\Psi, \Psi) \geq 0,$$

(where * denotes the complex conjugate), $(\Psi, \Psi) = 0$ if and only if $\Psi = 0$, and

$$(\Psi, c_1\Phi_1 + c_2\Phi_2) = c_1(\Psi, \Phi_1) + c_2(\Psi, \Phi_2).$$

It then follows that[3]

$$(c_1\Phi_1 + c_2\Phi_2, \Psi) = c_1^*(\Phi_1, \Psi) + c_2^*(\Phi_2, \Psi).$$

Any two vectors Ψ and Φ which are such that $(\Phi, \Psi) = 0$ are called *orthogonal* to one another. Similarly, two sets of vectors $\mathcal{S}_1$ and $\mathcal{S}_2$ are called mutually orthogonal if all the vectors in $\mathcal{S}_1$ are orthogonal to all the vectors in $\mathcal{S}_2$.

The *norm*, or "length," of a vector $\Psi \in \mathfrak{V}$ is defined by

$$\| \Psi \| = \sqrt{(\Psi, \Psi)},$$

so that $\| c\Psi \| = |c| \, \| \Psi \|$, and $\| \Psi \| = 0$ if and only if $\Psi = 0$. A vector whose norm equals unity is called *normalized;* such a vector is also called a *unit vector.*

The inner product and the norm satisfy two extremely useful inequalities: **Schwarz's inequality,**

$$|(\Phi, \Psi)| \leq \| \Phi \| \, \| \Psi \|, \tag{B.1}$$

where equality holds if and only if Φ and Ψ are multiples of one another or one of them vanishes, and the **triangle inequality,**

$$\| \Psi + \Phi \| \leq \| \Psi \| + \| \Phi \|, \tag{B.2}$$

in which equality holds if and only if there exists a real, non-negative number c such that either $\Psi = c\Phi$ or $\Phi = c\Psi$.

Schwarz's inequality is proved by a simple evaluation of the squared norm of the vector $\| \Phi \|^2 \, \Psi - (\Phi, \Psi)\Phi$:

$$\begin{aligned} 0 &\leq ([\| \Phi \|^2 \, \Psi - (\Phi, \Psi)\Phi], [\| \Phi \|^2 \, \Psi - (\Phi, \Psi)\Phi]) \\ &= \| \Phi \|^2 \, (\| \Phi \|^2 \, \| \Psi \|^2 - |(\Phi, \Psi)|^2), \end{aligned}$$

[3]In other words, the inner product is linear in its second member and "antilinear" in the first, a definition that is universally customary among physicists. In the mathematical literature you will find that mathematicians usually employ an inner product linear in the first and antilinear in the second member.

which proves it. Note that $\| \| \Phi \|^2 \Psi - (\Phi, \Psi)\Phi \| = 0$ if and only if Ψ is a multiple of Φ or one of them vanishes. As for the triangle inequality, I will ask you to prove it as an **exercise,** using Schwarz's inequality.

The vectors in a set $\{\Psi_i \neq 0,\ \ i = 1, \ldots, n\}$ are called *linearly dependent* if there exists a set of complex numbers $\{c_i,\ \ i = 1, \ldots, n\}$, not all of them equal to zero, such that $\sum_i c_i \Psi_i = 0$; if this equation holds for no set of complex numbers other than $\{c_i = 0,\ \ i = 1, \ldots, n\}$, then the n vectors are called *linearly independent.*

Suppose that $\mathfrak{V}$ is such that there is a maximal number of linearly independent vectors in it, i.e., given any set of non-zero vectors with more than n members, they must be linearly dependent. The number n is then called the *dimension* of $\mathfrak{V}$. It follows that, given any set of n linearly independent vectors $\{\Psi_i \in \mathfrak{V},\ \ i = 1, \ldots, n\}$, every given vector $\Phi \in \mathfrak{V}$ can be written as a linear combination of the form $\Phi = \sum_{i=1}^{n} c_i \Psi_i$, and the set $\{\Psi_i \in \mathfrak{V},\ \ i = 1, \ldots, n\}$ is called a *basis* in $\mathfrak{V}$. The most useful kinds of bases are the *orthonormal* ones, for which $(\Psi_i, \Psi_j) = \delta_{ij},\ i, j = 1, \ldots, n$.[4] In that case, the coefficients in the expansion $\Phi = \sum_i c_i \Psi_i$ are easily calculated, $c_i = (\Psi_i, \Phi)$, and the inner product of two vectors becomes the sum of products of these coefficients: if $\Phi_1 = \sum_i c_i^{(1)} \Psi_i$ and $\Phi_2 = \sum_i c_i^{(2)} \Psi_i$, then $(\Phi_1, \Phi_2) = \sum_i c_i^{(1)*} c_2^{(2)}$, and as a special case, $\| \Phi \|^2 = \sum_i |c_i|^2 = \sum_i |(\Psi_i, \Phi)|^2$, which is known as *Parseval's equality.*

A set of vectors in $\mathfrak{V}$ that forms a linear space by itself is called a *subspace.* For example, the set of linear combinations of $m < n$ linearly independent vectors $\{\Psi_i\}_{i=1}^{m}$ in an n-dimensional space $\mathfrak{V}$ form an m-dimensional proper subspace of $\mathfrak{V}$; it is called the space *spanned* by the vectors $\{\Psi_i\}_1^m$. On the other hand, the set of all vectors Ψ in $\mathfrak{V}$ such that $\| \Psi \| \leq 1$ is *not* a subspace.[5] The *orthogonal complement* $\mathcal{S}^\perp$ of a subspace $\mathcal{S}$ is the set of all vectors in $\mathfrak{H}$ that are orthogonal to all vectors in $\mathcal{S}$.[6]

Tensor products. Suppose $\mathfrak{V}^{\mathrm{I}}$ and $\mathfrak{V}^{\mathrm{II}}$ are two linear vector spaces with complete sets of orthonormal basis vectors $\{\Psi_i^{\mathrm{I}}\}_{i=1}^{n}$ and $\{\Psi_j^{\mathrm{II}}\}_{j=1}^{m}$, respectively. We may then form $n \times m$ new orthonormal vectors $\Psi_{ij} \stackrel{\mathrm{def}}{=} \Psi_i^{\mathrm{I}} \otimes \Psi_j^{\mathrm{II}}$ by defining their inner products $(\Psi_{ij}, \Psi_{kl}) \stackrel{\mathrm{def}}{=} (\Psi_i^{\mathrm{I}}, \Phi_k^{\mathrm{I}})(\Psi_j^{\mathrm{II}}, \Psi_l^{\mathrm{II}})$. The $n \times m$-dimensional linear vector space $\mathfrak{V}$ spanned by the vectors $\{\Psi_{ij}\}$ is called the *tensor product* of $\mathfrak{V}^{\mathrm{I}}$ and $\mathfrak{V}^{\mathrm{II}}$, written $\mathfrak{V} = \mathfrak{V}^{\mathrm{I}} \bigotimes \mathfrak{V}^{\mathrm{II}}$. The members of $\mathfrak{V}$ can all be expressed in the form of linear combinations of the kind $\Psi = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \Psi_i^{\mathrm{I}} \otimes \Psi_j^{\mathrm{II}}$.

Direct sums. The direct sum of the two spaces $\mathfrak{V}^{\mathrm{I}}$ and $\mathfrak{V}^{\mathrm{II}}$ is a space $\mathfrak{V} \stackrel{\mathrm{def}}{=} \mathfrak{V}^{\mathrm{I}} \bigoplus \mathfrak{V}^{\mathrm{II}}$ of dimension $n + m$ whose vectors are pairs $\Psi \stackrel{\mathrm{def}}{=} \{\Psi^{\mathrm{I}}, \Psi^{\mathrm{II}}\}$ with the inner product $(\Psi, \Phi) \stackrel{\mathrm{def}}{=} (\Psi^{\mathrm{I}}, \Phi^{\mathrm{I}})_{\mathfrak{V}^{\mathrm{I}}} + (\Psi^{\mathrm{II}}, \Phi^{\mathrm{II}})_{\mathfrak{V}^{\mathrm{II}}}$. Its first n basis

---

[4] δ_{ij} is the Kronecker symbol; its value equals 1 when $i = j$ and 0 otherwise.

[5] **Why not?**

[6] **Is it a subspace?**

vectors may be taken to be $\{\Psi_i^{\rm I}, 0\}$ and its second m are $\{0, \Psi_j^{\rm II}\}$, so that every vector $\Psi \in \mathfrak{V}$ has the form $\Psi = \sum_{i=1}^n a_i\{\Psi_i^{\rm I}, 0\} + \sum_{j=1}^m b_j\{0, \Psi_j^{\rm II}\}$, with the inner product $(\Psi, \Phi) = \sum_{i=1}^n a_i^* a_i' + \sum_{j=1}^m b_j^* b_j'$ if $\Phi = \sum_{i=1}^n \{a_i'\Psi_i^{\rm I}, 0\} + \sum_{j=1}^m b_j'\{0, \Psi_j^{\rm II}\}$.

B.2 Infinite-Dimensional Linear Vector Spaces

If there is no limit to the number of linearly independent vectors in $\mathfrak{V}$, its dimension is infinite, and certain new constraints arise from requirements of convergence. Specifically, the space is called *complete* **if Cauchy sequences converge** in it. A sequence $\mathcal{S} = \{\Psi_n\}_1^\infty$ of vectors in $\mathfrak{V}$ is called a *Cauchy sequence* if for every given $\epsilon > 0$ there exists an N such that for all $n, m \geq N$, $\| \Psi_n - \Psi_m \| < \epsilon$. Thus, if $\mathfrak{V}$ is complete and $\mathcal{S}$ is a Cauchy sequence $\{\Psi_n\}$ in it, then there exists a vector $\Psi \in \mathfrak{V}$ such that $\| \Psi_n - \Psi \| \to 0$ as $n \to \infty$. A *Hilbert space* is a complete linear vector space with an inner product. (Spaces of finite dimensionality are automatically complete.) The only Hilbert spaces we will be dealing with are complex. It is important to note that a subspace $\mathcal{S}$ of a Hilbert space $\mathfrak{H}$ is not necessarily a Hilbert space; $\mathcal{S}$ may not be complete, which means it may not contain all its limit points: a Cauchy sequence of vectors in $\mathcal{S}$ may converge to a vector $\Psi \in \mathfrak{H}$ but $\Psi \notin \mathcal{S}$. In order to make it into a Hilbert space it may have to be *completed* by adding all its limit points that were not already contained in it. The *tensor product* $\mathfrak{H}^{\rm I} \bigotimes \mathfrak{H}^{\rm II}$ of two Hilbert spaces $\mathfrak{H}^{\rm I}$ and $\mathfrak{H}^{\rm II}$ is defined to be the Hilbert space that is the completion of the tensor product of the two vector spaces as defined earlier for finite-dimensional vector spaces.

A Hilbert space $\mathfrak{H}$ is called *separable* if it contains a *countable set* of vectors that is *dense*[7] in $\mathfrak{H}$. All the Hilbert spaces of interest to physics are separable.

A *basis* in an infinite-dimensional Hilbert $\mathfrak{H}$ space cannot be defined quite as simply as in a space of finite dimension. An infinite sequence of linearly independent vectors $\{\Psi_n\}_1^\infty$ in $\mathfrak{H}$ is called a basis in $\mathfrak{H}$ if every $\Psi \in \mathfrak{H}$ can be expanded in the form $\Psi = \sum_n c_n \Psi_n$. The meaning of this, more precisely, is that if we call the partial sums $S_m \stackrel{\rm def}{=} \sum_{n=1}^m c_n \Psi_n$, then $\lim_{m\to\infty} \| \Psi - S_m \| = 0$. Again, the most useful bases are those that are orthonormal, because for these the expansion coefficients can be calculated simply by $c_n = (\Psi_n, \Psi)$.

A vector Ψ in $\mathfrak{H}$ is uniquely determined if we know its inner products (χ, Ψ) with all $\chi \in \mathfrak{H}$. This is so because then, specifically, all the coefficients $c_n = (\Psi_n, \Psi)$ in its expansion $\Psi = \sum_n c_n \Psi_n$ on the orthonormal basis

[7] A set $\mathcal{S}$ of vectors in a Hilbert space $\mathfrak{H}$ is *dense* in $\mathfrak{H}$ if for every given $\epsilon > 0$ and every $\Psi \in \mathfrak{H}$ there is a vector $\Phi \in \mathcal{S}$ such that $\| \Psi - \Phi \| < \epsilon$.

$\{\Psi_n\}$ are determined; therefore it is, in fact, sufficient to know all the inner products $c_n = (\Psi_n, \Psi)$ with a set of basis vectors.

Examples An example of a Hilbert space is $L^2(\mathbb{R})$, the space of square-integrable functions $f(x)$ on the real line. Here the inner product is defined by

$$(f, g) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} dx\, f^*(x) g(x),$$

and the norm, $\| f \|^2 \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} dx\, |f(x)|^2$. Another example is provided by the space l^2 of square-summable sequences $s = \{s_n\}_1^\infty$, for which the inner product is

$$(s, t) \stackrel{\text{def}}{=} \sum_1^\infty s_n^* t_n,$$

and the norm $\| s \|^2 \stackrel{\text{def}}{=} \sum_1^\infty |s_n|^2$. (The expansion of vectors in $\mathfrak{H}$ on a basis constitutes a one-to-one linear mapping of $\mathfrak{H}$ on l^2.) The space $L^1(\mathbb{R})$ of integrable functions, on the other hand, is *not* a Hilbert space. The Hilbert space $L^2(\mathbb{R}^3)$ of square-integrable functions $f(\vec{r})$ on the three-dimensional Euclidian space $\mathbb{R}^3$ is the tensor product $L^2(\mathbb{R}) \bigotimes L^2(\mathbb{R}) \bigotimes L^2(\mathbb{R})$ of three copies of $L^2(\mathbb{R})$.

B.3 Operators

A linear operator $\mathbf{M}$ on $\mathfrak{H}$ is a linear mapping that sends every vector Ψ in some subspace $\mathcal{D}(\mathbf{M}) \subseteq \mathfrak{H}$ into a unique vector Φ in a linear space $\mathcal{R}(\mathbf{M})$ which may or may not be a subspace of $\mathfrak{H}$: $\Phi = \mathbf{M}\Psi$; the linearity of the mapping means that $\mathbf{M}(c_1\Psi_1 + c_2\Psi_2) = c_1\mathbf{M}\Psi_1 + c_2\mathbf{M}\Psi_2$. The product of two operators is defined by associativity, $(\mathbf{MN})\Psi = \mathbf{M}(\mathbf{N}\Psi)$ [assuming that $\mathcal{D}(\mathbf{M}) \subseteq \mathcal{R}(\mathbf{N})$], and they may or may not commute. The space $\mathcal{D}(\mathbf{M})$ on which $\mathbf{M}$ is defined is called its *domain*, and the set $\mathcal{R}(\mathbf{M})$ of images of $\mathcal{D}(\mathbf{M})$ under the mapping $\mathbf{M}$ is called its *range;* thus, $\mathbf{M}$ maps its domain $\mathcal{D}(\mathbf{M})$ into its range $\mathcal{R}(\mathbf{M})$. Most, but not all, of the operators $\mathbf{M}$ of interest for us are such that both $\mathcal{D}(\mathbf{M}) \subseteq \mathfrak{H}$ and $\mathcal{R}(\mathbf{M}) \subseteq \mathfrak{H}$.

If $\mathcal{R}(\mathbf{M}) \subseteq \mathfrak{H}$ and the "matrix elements" $(\Upsilon, \mathbf{M}\Phi)$ of $\mathbf{M}$ are known for all Φ in the domain of $\mathbf{M}$ and all $\Upsilon \in \mathfrak{H}$, then the operator $\mathbf{M}$ is uniquely determined, because we can evaluate the numbers $c_n = (\Psi_n, \mathbf{M}\Phi)$, which are the expansion coefficients of the image $\Psi = \mathbf{M}\Phi$ of Φ on the orthonormal basis $\{\Psi_n\}$.

The *norm* of the operator $\mathbf{M}$ is defined by the upper limit of the norms of the images of all the unit vectors in $\mathfrak{H}$,[8]

$$\| \mathbf{M} \| \stackrel{\text{def}}{=} \sup_{\Psi \in \mathfrak{H}} \| \mathbf{M}\Psi \| / \| \Psi \|, \tag{B.3}$$

or equivalently,

$$\| \mathbf{M} \| \stackrel{\text{def}}{=} \sup_{\Phi, \Psi \in \mathfrak{H}} |(\Phi, \mathbf{M}\Psi)| / \| \Psi \| \| \Phi \| . \tag{B.4}$$

If this number is finite, the operator is called *bounded;* otherwise it is unbounded. A "bounded operator on $\mathfrak{H}$" is defined on all of $\mathfrak{H}$: $\mathcal{D}(\mathbf{M}) = \mathfrak{H}$, and its range is a subspace of $\mathfrak{H}$: $\mathcal{R}(\mathbf{M}) \subseteq \mathfrak{H}$; on the other hand, if $\mathbf{M}$ is unbounded, its domain $\mathcal{D}(\mathbf{M})$ is not all of $\mathfrak{H}$, i.e., $\mathcal{D}(\mathbf{M}) \subset \mathfrak{H}$.

As an example, take a *projection operator* P onto a one-dimensional subspace $\mathcal{S}$ spanned by the unit vector Φ, defined by $\mathsf{P}\Psi \stackrel{\text{def}}{=} (\Phi, \Psi)\Phi$. If $\mathcal{S}$ is n-dimensional, spanned by the orthonormal set $\{\Phi_i\}_1^n$, P is defined by $\mathsf{P}\Psi = \sum_{i=1}^n (\Phi_i, \Psi)\Phi_i$. In either case, P is both Hermitian (see below) and *idempotent:* $\mathsf{P}^2 = \mathsf{P}$. Such operators are defined on all of $\mathfrak{H}$, so $\mathcal{D}(\mathsf{P}) = \mathfrak{H}$, and their range is $\mathcal{S}$. The norm of a projection operator is always equal to one: $\| \mathsf{P} \| = 1$.[9]

The operator norm satisfies both *Schwarz's inequality,*

$$\| \mathbf{MN} \| \leq \| \mathbf{M} \| \| \mathbf{N} \|, \tag{B.5}$$

and the *triangle inequality,*[10]

$$\| \mathbf{M} + \mathbf{N} \| \leq \| \mathbf{M} \| + \| \mathbf{N} \| . \tag{B.6}$$

The *Hermitian conjugate* (also called *Hermitian adjoint*) $\mathbf{M}^\dagger$ of $\mathbf{M}$ is defined by

$$(\Phi, \mathbf{M}^\dagger \Psi) \stackrel{\text{def}}{=} (\Psi, \mathbf{M}\Phi)^*, \tag{B.7}$$

and the operator $\mathbf{M}$ is called *Hermitian* if $\mathbf{M} = \mathbf{M}^\dagger$. Such operators are singled out for representing physical observables because it immediately follows from (B.7) that *all expectation values of a Hermitian operator are real numbers:*

$$\Im\langle \mathbf{M} \rangle = 0 \quad \text{if} \quad \mathbf{M} = \mathbf{M}^\dagger. \tag{B.8}$$

If these expectation values are all positive, $\langle \mathbf{M} \rangle > 0$, $\mathbf{M}$ is called *positive definite*; if $\langle \mathbf{M} \rangle \geq 0$, it is called *positive semi-definite.*

---

[8] "sup" means "least upper bound."

[9] **Why?**

[10] **I'll let you prove these as exercises.**

To be mathematically careful, however, we have to pay attention to questions of domain here. Equation (B.7) makes sense only on the assumption that there are vectors Υ and $\Psi \in \mathfrak{H}$ such that for all $\Phi \in \mathcal{D}(\mathbf{M})$, $(\Psi, \mathbf{M}\Phi)^* = (\Phi, \Upsilon)$. If such pairs of vectors exists, then the vectors Ψ form the domain of $\mathbf{M}^\dagger$ and the vectors Υ its range. Note also that if the domain of $\mathbf{M}$ is a proper subspace of $\mathfrak{H}$, the right-hand side of (B.7) is not defined on all of $\mathfrak{H}$, and hence not all the matrix elements of $\mathbf{M}^\dagger$ are defined by this equation; specifically, its range is not completely defined. For example, we could, without changing the equation, add to $\mathbf{M}^\dagger$ any operator whose range is orthogonal to $\mathcal{D}(\mathbf{M})$, if such an operator exists. The strict definition of a Hermitian operator, of course, depends on such domain questions, and the equation $\mathbf{M}^\dagger = \mathbf{M}$ implies that $\mathcal{D}(\mathbf{M}) = \mathcal{D}(\mathbf{M}^\dagger)$. Only then do mathematicians call $\mathbf{M}$ *self-adjoint,* and only then are certain important theorems applicable. If the domain question is left open, $\mathbf{M}$ is called *symmetric* by mathematicians. It is such subtleties, and more, that we usually ignore as physicists, which does not mean, however, they are unimportant![11]

As an important example, take $\mathfrak{H}$ to be the space $L^2(a, b)$ of square-integrable functions of $a \leq x \leq b$ and consider the operator $\mathbf{p} \stackrel{\text{def}}{=} -id/dx$. Since the functions in L^2 are not all differentiable, the domain of $\mathbf{p}$ can be no larger than the subspace of *differentiable* functions in L^2. (However, this subspace is not a Hilbert space. To complete it, its limit points have to be added. The most convenient way of doing so is to define the derivative operator in terms of Fourier transforms.) Furthermore, since

$$\int_a^b dx\, f^*(x)[-i\frac{d}{dx}]g(x) = \int_a^b dx\,[-i\frac{d}{dx}f(x)]^* g(x) + i[f^*(a)g(a) - f^*(b)g(b)],$$

the operator is Hermitian only if its domain is restricted to the functions for which the end-point terms are absent. This can be accomplished, for example, by adopting boundary conditions such that all the functions in the domain of $\mathbf{p}$ vanish at the points a and b, or by assuming them to be *periodic*: $f(b) = f(a)$. If a is taken to be $-\infty$ and $b = \infty$, then the possible oscillating boundary terms are ignored. (Remember that a square-integrable function need not necessarily vanish at infinity.) Again, the best way of doing this is to start with the Fourier-transform space of L^2-functions $\widetilde{f}(k)$ and the (unbounded) multiplication operator k, whose domain of definition $\mathcal{D}(k)$ is the space of functions such that $k\widetilde{f}(k) \in L^2$. The inverse Fourier transform of the operator of multiplication by k is then the unbounded operator $-id/dx$ and its domain is the space of functions that are the Fourier transforms of functions in $\mathcal{D}(k)$, a proper subspace of $L^2(\mathbb{R})$.

Suppose a bounded operator $\mathbf{M}$ maps all of $\mathfrak{H}$ one-to-one onto all of $\mathfrak{H}$. In that case, it clearly has an inverse $\mathbf{M}^{-1}$, the operator that inverts the mapping, also defined one-to-one from $\mathfrak{H}$ onto $\mathfrak{H}$. The assumption that the

[11] For further mathematical details on these questions, see, for example, [Reed].

mapping of $\mathbf{M}$ is one-to-one, however, excludes the possibility that there is a vector $\Upsilon \neq 0$ in $\mathfrak{H}$ that is annihilated by $\mathbf{M}$: $\mathbf{M}\Upsilon = 0$, because if such an Υ exists, then $\mathbf{M}\Phi = \mathbf{M}(\Phi + \Upsilon)$, implying that the mapping is not one-to-one. Conversely, if there are two vectors Φ_1 and Φ_2 such that $\mathbf{M}\Phi_1 = \mathbf{M}\Phi_2$, then $\mathbf{M}(\Phi_1 - \Phi_2) = 0$. The set of non-zero vectors $\Upsilon \in \mathfrak{H}$ that are annihilated by $\mathbf{M}$ is called its *nullspace,*[12] which we shall denote by $\mathcal{N}(\mathbf{M})$; it is a subspace. Thus *the necessary and sufficient condition for a bounded operator to have an inverse is that its nullspace be trivial,* which means it contains only the zero-vector.

Just as the introduction of an orthonormal basis $\{\Upsilon_n\}_1^\infty$ in $\mathfrak{H}$ allows us to represent vectors Ψ by the sequence of their projections on these basis vectors, so we can represent operators by *matrices.* This is done by defining

$$\mathcal{M}_{mn} \stackrel{\text{def}}{=} (\Upsilon_m, \mathbf{M}\Upsilon_n), \tag{B.9}$$

or in Dirac notation, writing $\Upsilon_n = |n\rangle$,

$$\mathcal{M}_{mn} \stackrel{\text{def}}{=} \langle m|\mathbf{M}|n\rangle, \tag{B.10}$$

so that the representation of the image $\mathbf{M}\Psi$ of Ψ under the mapping $\mathbf{M}$ is given by the sequence

$$\langle m|\mathbf{M}|\,\rangle = \sum_n \mathcal{M}_{mn}\langle n|\,\rangle, \qquad m = 1, 2, \ldots,$$

where $|\,\rangle \stackrel{\text{def}}{=} \Psi$. Operator multiplication is thereby converted into *matrix multiplication:* if $\mathcal{K}$ is the matrix representing the product $\mathbf{MN}$, then $\mathcal{K}_{mn} = \sum_k \mathcal{M}_{mk}\mathcal{N}_{kn}$, which is most easily seen by using, between the two operators $\mathbf{MN}$, the symbolic statement

$$\sum_n |n\rangle\langle n| = \mathbb{1} \tag{B.11}$$

of the assumption that the sequence $\{\Upsilon_n = |n\rangle\}$ forms an orthonormal basis in $\mathfrak{H}$. [The meaning of (B.11) resides entirely in what it leads to when acting on an arbitrary vector $\Psi = |\,\rangle \in \mathfrak{H}$, which is $|\,\rangle = \sum_n |n\rangle\langle n|\,\rangle$.] It is, however, important to remember that the matrices we are dealing with here are *infinite-dimensional* and, because of convergence problems of the infinite series involved in their products, they lack some of the familiar properties of finite-dimensional matrices, as we shall see.

B.3.1 Hermitian operators and their spectra

The spectrum of a Hermitian operator $\mathbf{M}$ consists of two sets of points, one of which, but not both, may be empty. The first set, called the *point*

[12] It is also sometimes called its *kernel.*

spectrum or *discrete spectrum*, consists of those numbers α, if any, for which there exists a vector $\Psi^\alpha \in \mathfrak{H}$ such that[13]

$$(\mathbf{M} - \alpha)\Psi^\alpha = 0. \tag{B.12}$$

These numbers, called *eigenvalues* of $\mathbf{M}$, are necessarily *real*, because (B.12) implies $(\Psi, \mathbf{M}\Psi)/(\Psi, \Psi) = \alpha$, the left-hand side of which is real, according to (B.8). They form a denumerable, *discrete set of points* (which we shall not prove here).

The second part of the spectrum of a Hermitian operator consists of those numbers α, if any, with the property that for any given $\epsilon > 0$ there exists a unit vector $\Psi^\alpha_\epsilon \in \mathfrak{H}$ such that

$$\| (\mathbf{M} - \alpha)\Psi^\alpha_\epsilon \| < \epsilon, \tag{B.13}$$

but there exists no $\Psi^\alpha \in \mathfrak{H}$ such that $(\mathbf{M} - \alpha)\Psi^\alpha = 0$. These numbers form the *continuous spectrum* of $\mathbf{M}$, and we shall call them *quasi-eigenvalues.* The continuous spectrum of a Hermitian operator also lies on the real line, as is proved by a simple calculation for any complex $\alpha = \alpha_1 + i\alpha_2$ and $\| \Psi \| = 1$,

$$\| (\mathbf{M} - \alpha)\Psi \|^2 = \| (\mathbf{M} - \alpha_1)\Psi \|^2 + \alpha_2^2 \geq \alpha_2^2,$$

which implies that for a given $\alpha_2 \neq 0$, the left-hand side cannot be arbitrarily small. Thus *the entire spectrum of a Hermitian operator lies on the real line.* The "continuous spectrum" of a Hermitian operator indeed forms a continuum (which we shall also not prove here).[14]

If α is an eigenvalue of $\mathbf{M}$, there may be more than one corresponding eigenvector: the nullspace $\mathcal{N}(\mathbf{M} - \alpha)$ (called the *eigenspace* of $\mathbf{M}$ at the eigenvalue α) may be n-dimensional, with $n > 1$. The eigenvalue is then said to have *multiplicity* n or to be *n-fold degenerate,* and we can introduce n mutually orthogonal eigenvectors that span the eigenspace. The eigenvectors belonging to two *different* eigenvalues of a Hermitian operator, on the other hand, are automatically *mutually orthogonal.* [This can be seen by taking the inner product of $\mathbf{M}\Psi_1 = \alpha_1\Psi_1$ with Ψ_2 and subtracting the complex conjugate of the inner product of $\mathbf{M}\Psi_2 = \alpha_2\Psi_2$ with Ψ_1, which results in $0 = (\Psi_2, \mathbf{M}\Psi_1) - (\Psi_1, \mathbf{M}\Psi_2)^* = (\alpha_1 - \alpha_2)(\Psi_2, \Psi_1)$, since α_1 and α_2 are real. Therefore $(\Psi_1, \Psi_2) = 0$ if $\alpha_1 \neq \alpha_2$.]

Let us define P_{α_n} to be the projection on the nullspace $\mathcal{N}(\mathbf{M} - \alpha_n)$, so that $\mathsf{P}^2_{\alpha_n} = \mathsf{P}_{\alpha_n}$, $\mathsf{P}^\dagger_{\alpha_n} = \mathsf{P}_{\alpha_n}$, and $(\mathbf{M} - \alpha_n)\mathsf{P}_{\alpha_n} = 0$. The extremely

[13]Strictly speaking, $\mathbf{M} - \alpha$ should be written in the form $\mathbf{M} - \alpha\mathbf{1}$, but it is customary to omit the unit operator $\mathbf{1}$.

[14]By contrast, the "continuous spectrum," which mathematicians call the "essential spectrum," of non-Hermitian operators may contain isolated points, and the "discrete spectrum" (i.e., the eigenvalues) may form a continuum. In some instances the two kinds of spectra of a Hermitian operator may overlap, that is, some point eigenvalues may be "embedded in the continuum."

important **spectral theorem** asserts that, if the spectrum of the Hermitian operator $\mathbf{M}$ consists of discrete eigenvalues only, the direct sum of all the nullspaces $\mathcal{N}(\mathbf{M}-\alpha_n)$ is equal to all of $\mathfrak{H}$; this means that the *eigenvectors of* $\mathbf{M}$ *span the Hilbert space* $\mathfrak{H}$. As a result, the operator $\mathbf{M}$ can then be expanded in the form

$$\mathbf{M} = \sum_n \alpha_n \mathsf{P}_{\alpha_n} \tag{B.14}$$

and we have the "resolution of the identity"

$$\sum_n \mathsf{P}_{\alpha_n} = \mathbb{1}, \tag{B.15}$$

from which (B.14) follows by multiplication by $\mathbf{M}$. If we write $\{|\alpha_n\rangle\}$ for the set of orthonormal eigenvectors of $\mathbf{M}$ in the Dirac notation, (B.15) can also be written in the form[15]

$$\sum_n |\alpha_n\rangle\langle\alpha_n| = \mathbb{1}, \tag{B.16}$$

generally called a *completeness relation* and its immediate consequence, Parseval's equality for any vector $|\,\rangle \in \mathcal{H}$,

$$\langle\,|\,\rangle = \sum_n |\langle\alpha_n|\,\rangle|^2. \tag{B.17}$$

Here is an instance of an important difference between finite-dimensional matrices and those representing operators on an infinite-dimensional Hilbert space. In the n-dimensional case the eigenvalues α of the matrix $\mathcal{M}$ can be found from the fact that if the nullspace $\mathcal{N}(\mathcal{M}-\alpha)$ is nontrivial, i.e., if there exists a vector $\Psi \neq 0$ such that $(\mathbf{M}-\alpha)\Psi = 0$, α must solve the n^{th}-order algebraic equation $\det(\mathcal{M}-\alpha) = 0$. In the infinite-dimensional case, however, this determinant is not necessarily well defined. In the next chapter we shall discuss conditions under which such infinite-dimensional determinants exist and how they can be calculated. Note also that, whereas the proof of the spectral theorem in the finite-dimensional case is trivial when there is no degeneracy, and when there is, consists simply of the demonstration that the geometric and algebraic degeneracies are equal, its proof for operators on an infinite-dimensional Hilbert space is more intricate and we shall not give it here.

The case of a continuous spectrum is more difficult to handle in the abstract without introducing additional mathematical machinery. Let us begin by considering an example. Suppose our Hilbert space is the space $L^2(-a,a)$ of square-integrable functions from $-a$ to a: $f(x)$, $-a < x < a$.

---

[15] This notation is somewhat ambiguous. In the case of degeneracy, we will have to assume that when some α_n are equal, the corresponding eigenvectors are mutually orthogonal.

Let the operator $\mathbf{M}$ be $-\frac{d^2}{dx^2}$, defined on the domain of twice differentiable functions with periodic boundary conditions: $f(a) = f(-a)$, and $f'(a) = f'(-a)$; ; then the eigenvalues of $\mathbf{M}$ are the non-negative numbers $\lambda_n^2 = (\pi n/a)^2$, $n = 0, 1, \ldots$, and each eigenvalue is twofold degenerate: to each eigenvalue λ_n^2 there correspond two normalized eigenfunctions $f_{\pm n}(x) = \frac{1}{\sqrt{2a}} e^{\pm i\lambda_n x}$. The Fourier theorem tells us that these functions form an orthonormal basis in $L^2(-a, a)$: every function $f(x) \in L^2(-a, a)$ can be expanded in the form $f(x) = \frac{1}{\sqrt{2a}} \sum_{-\infty}^{\infty} c_n e^{\pi i n x/a}$, and the coefficients are given by $c_n = \frac{1}{\sqrt{2a}} \int_{-a}^{a} dx\, f(x) e^{-\pi i n x/a}$. Let us next take the Hilbert space to be $L^2(\mathbb{R})$, the space of square-integrable functions from $-\infty$ to $+\infty$, with the same operator $\mathbf{M} = -\frac{d^2}{dx^2}$, defined for functions whose second derivatives are also square-integrable. In that case, the spectrum consists of the continuum of all non-negative numbers λ^2, and to each quasi-eigenvalue there again correspond two quasi-eigenfunctions $(2\pi)^{-1/2} e^{\pm i\lambda x}$, but these functions are *not* square-integrable: they are not in the Hilbert space.[16] Nevertheless, the Fourier integral theorem says that any function $f(x) \in L^2(\mathbb{R})$ can be represented as an integral of the form

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty d\lambda^2 \, \left[g_1(\lambda) e^{i\lambda x} + g_2(\lambda) e^{-i\lambda x} \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\lambda \, g(\lambda) e^{i\lambda x}, \tag{B.18}$$

where the function $g(\lambda)$ is given by

$$g(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, f(x) e^{-i\lambda x}. \tag{B.19}$$

The two formulas (B.18) and (B.19) can be combined in the symbolic form[17]

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \, e^{i\lambda(x-x')} = \delta(x - x'), \tag{B.20}$$

which yields (B.18) and (B.19) on multiplication by $f(x)$ and integrating. Furthermore, the "normalization" of the quasi-eigenfunctions $f_\lambda(x) = (2\pi)^{-1/2} e^{\pm i\lambda x}$ is such that

$$\int dx\, f_\lambda^*(x) f_{\lambda'}(x) = \delta(\lambda - \lambda'). \tag{B.21}$$

---

[16] In some vague sense the spectrum of this operator may be thought of as arising from the operator on $L^2(-a, a)$ by letting $a \to \infty$, in which case its eigenvalues become more and more closely spaced. Note, however, that the normalized eigenfunctions vanish in the limit.

[17] Caution: The use of the Dirac delta function in the context of a Hilbert space like L^2 is inherently ambiguous, because functions in L^2 are not defined pointwise and need not be continuous. For example, the 0-function in L^2, which is such that $\int dx\, |f(x)|^2 = 0$, can differ from 0 at infinitely many points.

For a general Hermitian operator $\mathbf{M}$ with a continuous spectrum $\mathfrak{S}$ (and no point spectrum), the points of which we shall denote by α, the situation is analogous. To every $\Psi \in \mathfrak{H}$ there corresponds a square-integrable complex-valued function $f_\Psi(\alpha)$,

$$\Psi \leftrightarrow f_\Psi(\alpha),$$

defined on $\mathfrak{S}$, such that

$$\mathbf{M}\Psi \leftrightarrow f_{\mathbf{M}\Psi}(\alpha) = \alpha f_\Psi(\alpha), \tag{B.22}$$

and Parseval's equality holds,

$$\| \Psi \|^2 = \int_{\mathfrak{S}} d\alpha \, |f_\Psi(\alpha)|^2,$$

with the properties that if $\Psi \leftrightarrow f_\Psi(\alpha)$ and $\Phi \leftrightarrow f_\Phi(\alpha)$, then for all complex a and b, $a\Psi + b\Phi \leftrightarrow f_{a\Psi+b\Phi}(\alpha) = af_\Psi(\alpha) + bf_\Phi(\alpha)$ and

$$(\Phi, \Psi) = \int_{\mathfrak{S}} dA \, f^*_\Phi(\alpha) f_\Psi(\alpha).$$

The function $f_\Psi(\alpha)$ corresponding to $\Psi \stackrel{\text{def}}{=} |\,\rangle$ will be denoted by $f(\alpha) \stackrel{\text{def}}{=} \langle\alpha|\,\rangle$, while $f^*(\alpha) \stackrel{\text{def}}{=} \langle\,|\alpha\rangle$. As a function of Ψ it has the linearity properties of an inner product of the vector $\Psi = |\,\rangle$ with the *quasi-vector* $|\alpha\rangle$, and the latter is such that, according to (B.22), $\mathbf{M}|\alpha\rangle = \alpha|\alpha\rangle$, and

$$\int_{\mathfrak{S}} d\alpha \, |\alpha\rangle\langle\alpha| = \mathbb{1}, \tag{B.23}$$

as well as

$$\langle\alpha|\alpha'\rangle = \delta(\alpha - \alpha'), \tag{B.24}$$

which implies that, analogous to Fourier integrals,

$$|\,\rangle = \int_{\mathfrak{S}} d\alpha \, |\alpha\rangle\langle\alpha|\,\rangle = \int_{\mathfrak{S}} d\alpha \, |\alpha\rangle f(\alpha). \tag{B.25}$$

The quasi-eigenvector $|\alpha\rangle$ is not in $\mathfrak{H}$, because a square-integrable function need not have a finite value at every point, so there is no assurance that for a given α the inner product $\langle\alpha|\,\rangle$ has to be finite for all $|\,\rangle \in \mathfrak{H}$. (In fact, for every $\alpha \in \mathfrak{S}$ there must exist an $|\,\rangle \in \mathfrak{H}$ such that $\langle\alpha|\,\rangle$ is not finite, because otherwise α would be an eigenvalue of $\mathbf{M}$ rather than a quasi-eigenvalue.)[18]

Let $\{\phi_n\}$ be an orthonormal basis in $\mathfrak{H}$; then

$$\delta_{nm} = (\phi_n, \phi_m) = \int_{\mathfrak{S}} d\alpha \, f^*_{\phi_n}(\alpha) f_{\phi_m}(\alpha) \tag{B.26}$$

[18] See also the formulation in terms of *rigged Hilbert spaces*; e.g., [Bohm, A.].

as well as

$$\begin{aligned}(\Psi,\Phi) &= \sum_n (\Psi,\phi_n)(\phi_n,\Phi)\\ &= \int_{\mathfrak{S}} dA \int_{\mathfrak{S}} d\alpha'\, f^*_\Psi(\alpha) f_\Phi(\alpha') \sum_n f_{\phi_n}(\alpha) f^*_{\phi_n}(\alpha')\\ &= \int_{\mathfrak{S}} d\alpha\, f^*_\Psi(\alpha) f_\Phi(\alpha),\end{aligned}$$

from which we conclude that

$$\sum_n f_{\phi_n}(\alpha) f^*_{\phi_n}(\alpha') = \delta(\alpha-\alpha'). \tag{B.27}$$

To every function $f \in L^2(\mathfrak{S})$, conversely, there corresponds a vector $\Psi \in \mathfrak{H}$. The expansion coefficients c_n of this Ψ on a basis $\{\phi_n\}$ are given by

$$c_n = \int_{\mathfrak{S}} d\alpha\, f^*_{\phi_n}(\alpha) f(\alpha),$$

which by (B.26) implies that

$$\| \Psi \|^2 = \sum_n |c_n|^2 = \int_{\mathfrak{S}} d\alpha\, |f(\alpha)|^2.$$

To prove that $\mathfrak{S}$ indeed constitutes the continuous spectrum of the operator $\mathbf{M}$, choose any $\alpha \in \mathfrak{S}$, any normalised $\Psi \in \mathfrak{H}$, and let $\Phi_\epsilon \in \mathfrak{H}$ be the normalised vector in $\mathfrak{H}$ corresponding to

$$f_\epsilon(\beta) \stackrel{\text{def}}{=} \sqrt{\frac{\epsilon/\pi}{\epsilon^2+(\beta-\alpha)^2}}.$$

Then

$$\begin{aligned}|(\Psi,(\mathbf{M}-\alpha)\Phi_\epsilon)|^2 &= \left| \int_{\mathfrak{S}} d\beta\, (\beta-\alpha) f_{\Phi_\epsilon}(\beta) f^*_\Psi(\beta) \right|^2\\ &\le \int_{\mathfrak{S}} d\beta'\, |f_\Psi(\beta')|^2 \int_{\mathfrak{S}} d\beta\, \frac{(\beta-\alpha)^2\epsilon/\pi}{\epsilon^2+(\beta-\alpha)^2} \le \epsilon\end{aligned}$$

which implies that α is in the continuous spectrum of $\mathbf{M}$ [unless there exists an eigenvector $\Upsilon \in \mathfrak{H}$, with $(\mathbf{M}-\alpha)\Upsilon = 0$, which is not the case by construction].

If the spectrum of $\mathbf{M}$ is partly discrete and partly continuous, then the expansions consist in part of sums and in part of integrals.

For a general, abstract formulation of the spectral theorem for Hermitian operators, we proceed as follows. Let $\mathbf{M}$ be a given Hermitian operator whose spectrum $\mathfrak{S}$ consists of

the countable point spectrum $\mathfrak{S}_p = \{\alpha_1, \alpha_2, \ldots\}$, with corresponding eigenvectors $\Psi_n \stackrel{\text{def}}{=} |\alpha_n\rangle$, and the continuous spectrum $\mathfrak{S}_c$, $\mathfrak{S} = \mathfrak{S}_p \cup \mathfrak{S}_c$. Define the *spectral projection*, which is a projection operator function $\mathsf{P}(\alpha)$, $\alpha \in \mathfrak{S}$, that is monotonely non-decreasing [which means that $\mathsf{P}(\alpha)\mathsf{P}(\alpha') = \mathsf{P}(\alpha')\mathsf{P}(\alpha) = \mathsf{P}(\alpha)$ if $\alpha \leq \alpha'$] from 0 to $\mathbf{1}$, and such that $[\mathbf{M}, \mathsf{P}(\alpha)] = 0$ for all $\alpha \in \mathfrak{S}$ and $\mathbf{M} = \int_{\mathfrak{S}} d\mathsf{P}(\alpha)\alpha$ as a Stieltjes integral. This spectral projection is defined so that for all $\Psi, \Phi \in \mathfrak{H}$,[19]

$$(\Psi, \mathsf{P}(\alpha)\Phi) = \sum_{\alpha_n \in \mathfrak{S}_p} \theta(\alpha - \alpha_n)(\Psi, \Psi_n)(\Psi_n, \Phi) + \int_{\mathfrak{S}_c} d\alpha'\, \theta(\alpha - \alpha') f^*_\Psi(\alpha') f_\Phi(\alpha').$$

The spectral theorem consists of the equation (which we are not going to pove)

$$\int_{\mathfrak{S}} d\mathsf{P}(\alpha) = \mathbf{1}. \tag{B.28}$$

In terms of the notation $(\Psi_n, \Psi) = \langle \alpha_n | \, \rangle$ and $f_\Psi(\alpha) = \langle \alpha | \, \rangle$, the spectral projection may be written in the form

$$\mathsf{P}(\alpha) = \sum_{\alpha_n \in \mathfrak{S}_p} \theta(\alpha - \alpha_n) |\alpha_n\rangle \langle \alpha_n| + \int_{\mathfrak{S}_c} d\alpha' \theta(\alpha - \alpha') |\alpha'\rangle \langle \alpha'|,$$

and the spectral theorem is the "resolution of the identity"

$$\int_{\mathfrak{S}_c} d\alpha \, |\alpha\rangle \langle \alpha| + \sum_{\alpha_n \in \mathfrak{S}_p} |\alpha_n\rangle \langle \alpha_n| = \mathbf{1}, \tag{B.29}$$

meaning that for all $\Psi \in \mathfrak{H}$,

$$\Psi = |\,\rangle = \int_{\mathfrak{S}_c} d\alpha \, |\alpha\rangle \langle \alpha| \, \rangle + \sum_{\alpha_n \in \mathfrak{S}_p} |\alpha_n\rangle \langle \alpha_n| \, \rangle, \tag{B.30}$$

as well as the Parseval equality

$$\| \, \Psi \, \|^2 = \langle \, | \, \rangle = \int_{\mathfrak{S}_c} d\alpha |\, \langle \alpha | \, \rangle|^2 + \sum_{\alpha_n \in \mathfrak{S}_p} |\langle \alpha_n | \, \rangle|^2. \tag{B.31}$$

B.3.2 More on operators

An operator $\mathbf{M}$ that satisfies the equation $\mathbf{M}^\dagger \mathbf{M} = \mathbb{1}$ is *isometric,* which means it preserves the lengths of all $\Psi \in \mathfrak{H}$, $\| \, \mathbf{M}\Psi \, \| = \| \, \Psi \, \|$; furthermore, if two vectors Ψ and Φ are orthogonal, so are their images under the mapping $\mathbf{M}$: if $(\Psi, \Phi) = 0$, then $0 = (\Psi, \mathbf{M}^\dagger \mathbf{M} \Phi) = (\mathbf{M}\Psi, \mathbf{M}\Phi)$. However, if the range of $\mathbf{M}$ is a proper subspace of $\mathfrak{H}$, let $\chi \neq 0$ be in the orthogonal complement[20] $\mathcal{R}^\perp(\mathbf{M})$ of $\mathcal{R}(\mathbf{M})$; then for all $\Psi \in \mathfrak{H}$ we have $(\mathbf{M}^\dagger \chi, \Psi) = (\chi, \mathbf{M}\Psi) = 0$, and hence $\mathbf{M}^\dagger \chi = 0$, as well as conversely: if $\mathbf{M}^\dagger \chi = 0$, then $(\chi, \Phi) = 0$ for all $\Phi \in \mathcal{R}(\mathbf{M})$; therefore, $\mathcal{N}(\mathbf{M}^\dagger) = \mathcal{R}^\perp(\mathbf{M})$. In such a case we must have $\mathbf{M}\mathbf{M}^\dagger \neq \mathbb{1}$, because the left-hand side annihilates any

---

[19] **Show that the so-defined function $\mathsf{P}(\alpha)$ has all the properties mentioned.**

[20] The existence of a vector orthogonal to $\mathcal{R}(\mathbf{M})$ implies that the range is not dense in $\mathfrak{H}$.

vector in $\mathcal{N}(\mathbf{M}^\dagger)$, whereas the right-hand side does not. So while for finite-dimensional matrices $\mathbf{M}^\dagger\mathbf{M} = \mathbb{1}$ implies $\mathbf{M}\mathbf{M}^\dagger = \mathbb{1}$, for operators on an infinite-dimensional space these two equations are independent. If, on the other hand, $\mathbf{M}^\dagger\mathbf{M} = \mathbf{M}\mathbf{M}^\dagger = \mathbb{1}$, then $\mathbf{M}$ is called *unitary*. So for a unitary operator, $\mathbf{M}^\dagger = \mathbf{M}^{-1}$. *A unitary operator* U *transforms any orthonormal basis* $\{\Psi_n\}$ *in* $\mathfrak{H}$ *into another orthonormal basis* $\{\Phi_n\}$ *by* $\Phi_n = \mathsf{U}\Psi_n$.

Suppose now that the operator $\mathbf{M}$ is represented as a matrix $\mathcal{M}$ with elements $\mathcal{M}_{mn} = (\Psi_m, \mathbf{M}\Psi_n)$ on some given orthonormal basis $\{\Psi_n\}$. Then the new matrix $\mathcal{M}'$ representing $\mathbf{M}$ on the new basis $\{\Phi_n = \mathsf{U}\Psi_n\}$ has the matrix elements

$$\mathcal{M}'_{mn} = (\Phi_m, \mathbf{M}\Phi_n) = (\Psi_m, \mathsf{U}^\dagger\mathbf{M}\mathsf{U}\Psi_n) = \sum_{kl} \mathcal{U}^*_{km}\mathcal{M}_{kl}\mathcal{U}_{ln}, \tag{B.32}$$

where

$$\mathcal{U}_{ln} \stackrel{\text{def}}{=} (\Psi_l, \mathsf{U}\Psi_n) \tag{B.33}$$

is the matrix representing U on he old basis $\{\Psi_n\}$, in terms of which the new vectors are given by

$$\Phi_n = \sum_m \Psi_m \mathcal{U}_{mn}. \tag{B.34}$$

Conversely, any matrix $\mathcal{M}_{mn}$ given in the basis $\{\Psi_n\}$, together with the transformation property (B.32) defines an operator $\mathbf{M}$ on $\mathfrak{H}$. Note that if $\mathbf{M}$ is Hermitian, then the matrix $\mathcal{M}$ is also Hermitian: the Hermitian property is preserved under unitary transformations.

If $\mathbf{M}$ is Hermitian and its spectrum consists of discrete points only, its eigenvectors χ_n, corresponding to the eigenvalues α_n, can be assumed to form an orthonormal basis in $\mathfrak{H}$, so that $\mathbf{M}\chi_n = \alpha_n\chi_n$ and as a result $\mathcal{M}'_{mn} = (\chi_m, \mathbf{M}\chi_n) = \alpha_n\delta_{mn}$. This means that $\mathcal{M}'$ is *diagonal* and the unitary transformation defined by (B.33) has *diagonalized* the matrix $\mathcal{M}$: *given a matrix* $\mathcal{M}$ *representing a Hermitian operator* $\mathbf{M}$*, finding its eigenvalues and eigenvectors is equivalent to diagonalizing* $\mathcal{M}$. Moreover, it follows from (B.32) and (B.34) that if $\mathcal{U}$ is the diagonalizing matrix, then

$$\sum_n \mathcal{M}_{mn}\mathcal{U}_{nl} = \alpha_l\mathcal{U}_{ml}, \tag{B.35}$$

which means that the columns of the matrix $\mathcal{U}$ consist of the components of the eigenvectors of $\mathbf{M}$ on the old basis, the l^{th} column forming the eigenvector belonging to the l^{th} eigenvalue.

A very useful notion is that of the **trace** of an operator. If the operator $\mathbf{X}$ is given as a matrix X_{nm} in some representation, its trace is defined by $\operatorname{tr}\mathbf{X} \stackrel{\text{def}}{=} \sum_n X_{nn}$, that is, the sum of its diagonal elements. An important property of the trace is its *invariance* under a change of basis. This is easily proved: $\operatorname{tr} X^{(2)} = \sum_n X^{(2)}_{nn} = \sum_{nkl} U^{(12)*}_{kn} X^{(1)}_{kl} U^{(12)}_{ln} = \sum_n X^{(1)}_{nn} = \operatorname{tr} X^{(1)}$,

by the unitarity property of $U^{(12)}$. Therefore it makes sense to refer to the trace of the *operator*, rather than specifying its matrix representation. If $\mathbf{X}$ is Hermitian, we may represent it on the basis of its eigenvectors, in which case the diagonal elements are its eigenvalues. So the trace is the *sum of its eigenvalues.* (In cases of degeneracy, each eigenvalue has to be counted as many times as its degeneracy indicates.)

These things work, formally, equally well for a quasi-basis and the case of quasi-eigenvalues, provided that the integral kernel $\mathcal{M}(\beta,\gamma)$ that takes the place of the matrix $\mathcal{M}_{mn}$, and which is formed by means of quasi-vectors, $\mathcal{M}(\beta,\gamma) = \langle\beta|\mathbf{M}|\gamma\rangle$, exists, which, even for a bounded operator $\mathbf{M}$, cannot always be taken for granted.

The trace of an operator may be evaluated also using a quasi-basis, because

$$\int dB\,\langle B|\mathbf{X}|B\rangle = \sum_{mn}\int dB\,U_m^{(12)*}(B)X_{mn}U(12)_n(B) = \sum_n X_{nn}, \quad \text{(B.36)}$$

or,

$$\operatorname{tr}\mathbf{X} = \int dB\,\langle B|\mathbf{X}|B\rangle. \quad \text{(B.37)}$$

Equation (B.37) shows, however, that if the spectrum of the Hermitian operator $\mathbf{X}$ is continuous, then its trace cannot be finite; it is *not* the integral over its quasi-eigenvalues.

Finally, there is the important fact that if $\mathbf{M}$ and $\mathbf{N}$ are two Hermitian operators that *commute* with one another, $[\mathbf{M},\mathbf{N}] = 0$, then there exists an orthonormal basis of *common eigenvectors* for them, which implies that *they can be simultaneously diagonalized.* This is proved as follows.

Suppose that $\mathbf{M}\chi = \alpha\chi$; then $\alpha\mathbf{N}\chi = \mathbf{N}\mathbf{M}\chi = \mathbf{M}\mathbf{N}\chi$, which means that $\mathbf{N}\chi$ is an eigenvector of $\mathbf{M}$ with the same eigenvalue α. Therefore, if α is not degenerate, we must have $\mathbf{N}\chi = \beta\chi$ and χ is an eigenvector of $\mathbf{N}$ as well. On the other hand, if α is m-fold degenerate, then $\mathbf{N}\chi$ must lie in the m-dimensional eigenspace of $\mathbf{M}$ at α, a space we may assume to be spanned by the orthonormal basis $\{\chi_r\}_1^m$, so that $\mathbf{N}\chi_r = \sum_n \chi_n c_{nr}$, where $c_{nr} = (\chi_n, \mathbf{N}\chi_r)$ is a Hermitian $m\times m$ matrix. Diagonalizing this matrix by introducing a new orthonormal basis in the eigenspace of $\mathbf{M}$ then has the effect of ending up with an orthonormal basis consisting of eigenvectors of both $\mathbf{M}$ and $\mathbf{N}$. □

A *complete set of commuting Hermitian operators* on $\mathfrak{H}$ is a set of such operators all of whose eigenvectors can be *unambiguously* labeled by their eigenvalues. If an eigenvalue of one of these operators is degenerate, the eigenspace is spanned by a basis of eigenstates of the other operators in the set. For example, in $L^2(\mathbb{R})$ we may take the operator $\mathbf{M} \overset{\text{def}}{=} -\frac{d^2}{dx^2}$, whose quasi-eigenvalues, as we saw, are twofold degenerate, and add the *parity* operator $\mathbf{P}$ defined by $\mathbf{P}f(x) \overset{\text{def}}{=} f(-x)$, which commutes with $\mathbf{M}$. If $f_{\pm\lambda}(x) = e^{\pm i\lambda x}$ are two linearly independent quasi-eigenfunctions of $\mathbf{M}$

with the eigenvalue λ^2, then the common quasi-eigenfunctions of $\mathbf{M}$ and $\mathbf{P}$ are the even and odd combinations of $f_{\pm\lambda}$, which are $\cos(\lambda x)$ and $\sin(\lambda x)$. These functions, then, form a complete set of common quasi-eigenfunctions of the two Hermitian operators $\mathbf{M}$ and $\mathbf{P}$, with $\cos(\lambda x)$ belonging to the pair of quasi-eigenvalues $(\lambda^2, +1)$ and $\sin(\lambda x)$ to the pair $(\lambda^2, -1)$.

B.4 Problems and Exercises

1. Prove the triangle inequality (B.2) for vectors in a Hilbert space, using Schwarz's inequality.

2. Prove that every finite-dimensional linear vector space (with an inner product) is complete.

3. Prove that (B.3) and (B.4) are equivalent.

4. Prove Schwarz's inequality (B.5) for the norm of the product of two linear operators on a Hilbert space.

5. Prove the triangle inequality (B.6) for the norm of the sum of two linear operators on a Hilbert space.

6. Prove the inequality $(\Psi, \mathbf{M}\mathsf{P}(\alpha)\Psi) \leq \alpha \parallel \Psi \parallel^2$ for all $\Psi \in \mathfrak{H}$, if $\mathsf{P}(\alpha)$ is the spectral projection for $\mathbf{M}$.

7. Take $\mathfrak{H}$ to be the space of square-integrable functions $f(x)$ on the real line, $L^2(I\!R)$, and let $\mathbf{M}$ be the operator of multiplication by the square-integrable function $h(x)$, defined on all of $\mathfrak{H}$. Prove that $\mathcal{R}(\mathbf{M}) \subseteq L^1(I\!R)$. [Note that $L^1(I\!R)$ is not a subspace of $L^2(I\!R)$.]

8. Prove that the nullspace $\mathcal{N}(\mathbf{M})$ of a linear operator $\mathbf{M}$ on $\mathfrak{H}$ is a subspace of $\mathfrak{H}$.

9. Show that if $\mathbf{M}$ is a Hermitian operator, then its norm is equal to the absolute value of the upper or lower end of its spectrum, whichever is the larger.

10. Show that the operator defined by the left-hand side of (B.16) is a projection.

11. Let $\mathfrak{H}$ be the space $L^2(I\!R)$ of square-integrable functions $f(x)$, and consider the operator $\mathbf{x}$ of multiplication by x. Show that every real number x is a quasi-eigenvalue of $\mathbf{x}$ as defined by (B.13).

12. Let $\mathfrak{H}$ be the space $L^2(I\!R)$ of square-integrable functions $f(x)$, and consider the operator $\mathbf{p} \stackrel{\text{def}}{=} -i\frac{d}{dx}$. Show that every real number p is a quasi-eigenvalue of $\mathbf{p}$ as defined by (B.13).

C
Linear Integral Equations

C.1 Fredholm Methods

The integral equations we are dealing with in quantum mechanics are usually Fredholm equations, or they can be transformed into such. A *Fredholm equation of the first kind* is a homogeneous equation of the form[1]

$$f(\vec{r}) = \lambda \int d^3r'\, K(\vec{r},\vec{r}')f(\vec{r}'),$$

which we shall write in the operator form $f = \lambda \mathbf{K} f$, or

$$(\mathbb{1} - \lambda \mathbf{K})f = 0,$$

and its inhomogeneous version, called a *Fredholm equation of the second kind,*

$$f(\vec{r}) = g(\vec{r}) + \lambda \int d^3r'\, K(\vec{r},\vec{r}')f(\vec{r}'), \tag{C.1}$$

or, in operator form,

$$(\mathbb{1} - \lambda \mathbf{K})f = g.$$

The inhomogeneity g is assumed to be square-integrable, the operator (or integral kernel) $\mathbf{K}$ is assumed to be *compact*, and we are looking for solutions f that are also square-integrable; so we are working on the Hilbert space of square-integrable functions.

[1] We are taking the variables $\vec{r}$ and $\vec{r}'$ to be in three-dimensional Euclidian space, but they could be in other spaces.

Compactness of an operator means that it can be arbitrarily closely approximated by an operator of finite rank, i.e., by an operator whose range is of finite dimensions and which is therefore representable as a finite-dimensional matrix. [If $\mathbf{K}$ is of finite rank, it can be written as a finite sum of the form $K(\vec{r},\vec{r}') = \sum_1^N F_n(\vec{r})G_n(\vec{r}').$] Compact operators share many of the simple properties of finite matrices, particularly the fact that, except for the origin, their spectra consist of point eigenvalues only. Each of these eigenvalues has finite degeneracy, and they accumulate at the origin. Furthermore, if α is an eigenvalue of $\mathbf{K}$, then α^* is an eigenvalue of $\mathbf{K}^\dagger$. (If $\mathbf{K}$ is compact, so is $\mathbf{K}^\dagger$.)

If α is an eigenvalue of $\mathbf{K}$, then for $\lambda = 1/\alpha$ the homogeneous version of (C.1) has a nontrivial solution $h \in L^2$, and since h can be added to any solution of (C.1), the solution of the inhomogeneous equation can, of course, not be unique. But can the inhomogeneous equation still have solutions? This works just as for matrices: suppose h is an eigenfunction of $\mathbf{K}$, $\mathbf{K}h = \alpha h$;, and h' is an eigenfunction of $\mathbf{K}^\dagger$ with the eigenvalue α^*. Taking the inner product of the inhomogeneous equation (C.1) with h' then immediately leads to the consequence $(h', g) = 0$. Therefore, *a necessary condition for the inhomogeneous equation (C.1) to have a solution is that the inhomogeneity g be orthogonal to all solutions of the adjoint homogeneous equation $h' = \lambda^*\mathbf{K}^\dagger h'$.* That this condition is also sufficient can be seen by restricting $\mathbf{K}$ to the orthogonal complement of its eigenspace at $1/\lambda$.

The result of this reasoning is the **Fredholm alternative**: *If the homogeneous version of (C.1) has no nontrivial solution in L^2, then the inhomogeneous equation (C.1) has a unique solution in L^2; if the homogeneous form of (C.1) has a nontrivial solution in L^2, then (C.1) has a solution in L^2 if and only if g is orthogonal to all solutions of the adjoint homogeneous equation $h' = \lambda^*\mathbf{K}^\dagger h'$.*

Assuming, then, that $1/\lambda$ is not an eigenvalue of $\mathbf{K}$, the next question is how to solve (C.1), i.e., how to construct the operator $\mathbf{M}^{-1} \stackrel{\text{def}}{=} (\mathbb{1} - \lambda\mathbf{K})^{-1}$, in terms of which the solution of (C.1) is given by $f = \mathbf{M}^{-1}g$. The **Fredholm method** is to proceed just as for finite-dimensional matrices, which is as follows. The matrix $\mathbf{M}^{-1}$ is a matrix $\mathbf{N}$, whose transpose consists of the cofactors of $\mathbf{M}$, divided by the determinant of $\mathbf{M}$, $\Delta \stackrel{\text{def}}{=} \det \mathbf{M}$, so that

$$\mathbf{NM} = \mathbf{MN} = \mathbb{1}\Delta. \tag{C.2}$$

Now both Δ and $\mathbf{N}$ may be expanded in power series in λ,

$$\begin{aligned} \mathbf{N} &= \textstyle\sum_{n=0}^{\infty} \lambda^n \mathbf{N}_n, \qquad \mathbf{N}_0 = \mathbb{1}, \\ \Delta &= \textstyle\sum_{n=0}^{\infty} \lambda^n \Delta_n, \qquad \Delta_0 = 1. \end{aligned} \tag{C.3}$$

For finite-dimensional matrices, of course, these power series terminate and become polynomials. Next use the familiar differentiation rule for determi-

nants,

$$\frac{d\Delta}{d\lambda} = \mathrm{tr}\mathbf{N}\frac{d\mathbf{M}}{d\lambda} = -\mathrm{tr}\mathbf{N}\mathbf{K}, \tag{C.4}$$

and combine it with (C.2):

$$-\mathbb{1}\mathrm{tr}\mathbf{N}\mathbf{K} = -\mathbf{N}\mathbf{K} + \frac{d\mathbf{N}}{d\lambda}\mathbf{M} = -\mathbf{K}\mathbf{N} + \mathbf{M}\frac{d\mathbf{N}}{d\lambda}.$$

Inserting the expansions (C.3) in this equation, we obtain the recursion relations for $n \geq 1$

$$\mathbf{N}_n = \mathbf{K}\mathbf{N}_{n-1} - \frac{1}{n}\mathbb{1}\mathrm{tr}\mathbf{K}\mathbf{N}_{n-1} = \mathbf{N}_{n-1}\mathbf{K} - \frac{1}{n}\mathbb{1}\mathrm{tr}\mathbf{K}\mathbf{N}_{n-1}, \tag{C.5}$$

and

$$\Delta = 1 - \sum_{n=1}^{\infty}\frac{\lambda^n}{n}\mathrm{tr}\mathbf{K}\mathbf{N}_{n-1}. \tag{C.6}$$

These expansions may next be used to construct the inverse in the form

$$M^{-1} = \frac{1}{\Delta}\mathbf{N} \stackrel{\mathrm{def}}{=} \mathbb{1} + \frac{\lambda}{\Delta}\mathbf{Y} \tag{C.7}$$

by the series

$$\begin{aligned}
\mathbf{Y} &= \mathbf{K}\mathbf{N} = \sum_{n=0}^{\infty}\lambda^n\mathbf{Y}_n, \\
\mathbf{Y}_n &= \mathbf{K}\mathbf{Y}_{n-1} + \mathbf{K}\Delta_n = \mathbf{Y}_{n-1}\mathbf{K} + \mathbf{K}\Delta_n, \;\; n > 0, \;\; \mathbf{Y}_0 = \mathbf{K}, \qquad \text{(C.8)} \\
\Delta &= \sum_{n=0}^{\infty}\lambda^n\Delta_n, \qquad \Delta_n = -\frac{1}{n}\mathrm{tr}\mathbf{Y}_{n-1}, \quad \Delta_0 = 1. \qquad \text{(C.9)}
\end{aligned}$$

The function $\Delta(\lambda)$ is called the *Fredholm determinant*, and $\mathbf{Y}$ is called the *first Fredholm minor.*

Other expressions for Δ and $\mathbf{Y}$ are obtained by defining the determinants

$$\mathfrak{y}_{n+1}(\vec{r},\vec{r}';\vec{r}_1,\ldots,\vec{r}_n) \stackrel{\mathrm{def}}{=} \begin{vmatrix} K(\vec{r},\vec{r}') & K(\vec{r},\vec{r}_1) & \ldots & K(\vec{r},\vec{r}_n) \\ K(\vec{r}_1,\vec{r}') & K(\vec{r}_1,\vec{r}_1) & \ldots & K(\vec{r}_1,\vec{r}_n) \\ \vdots & \vdots & \ldots & \vdots \\ K(\vec{r}_n,\vec{r}') & K(\vec{r}_n,\vec{r}_1) & \ldots & K(\vec{r}_n,\vec{r}_n) \end{vmatrix}, \tag{C.10}$$

and in terms of them, we find that

$$Y(\vec{r},\vec{r}') = K(\vec{r},\vec{r}') + \sum_{n=1}^{\infty}\frac{(-\lambda)^n}{n!}\int d^3r_1\cdots d^3r_n\,\mathfrak{y}_{n+1}(\vec{r},\vec{r}';\vec{r}_1,\ldots,\vec{r}_n), \tag{C.11}$$

as well as

$$\Delta = 1 + \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} \int d^3r_1 \cdots d^3r_n \, \mathfrak{y}_n(\vec{r}_1, \vec{r}_1; \vec{r}_2, \ldots, \vec{r}_n). \tag{C.12}$$

Now we have to answer two questions: 1) What are the requirements on $\mathbf{K}$ for all the quantities in (C.8), (C.9), (C.11), and (C.12) to make sense? 2) Under what conditions do the power series in (C.8), (C.9), (C.11), and (C.12) converge? The answer to the first question is that $\mathrm{tr}\mathbf{K}^n$ has to be finite for $n = 1, 2, \ldots$, and that will be assured if $\mathrm{tr}\mathbf{K} = \int d^3r \, K(\vec{r}, \vec{r})$ is finite, and, in addition, $\mathbf{K}$ is in the *Hilbert-Schmidt class*, which means $\mathrm{tr}\mathbf{K}\mathbf{K}^\dagger < \infty$, or, more explicitly,

$$\int d^3r d^3r' \, |K(\vec{r}, \vec{r}')|^2 < \infty. \tag{C.13}$$

Being in the Hilbert-Schmidt class ensures that $\mathbf{K}$ is compact; in fact, although it is not a necessary condition for compactness, the Hilbert-Schmidt test is the simplest one to perform, and many of the kernels of integral equations of physical interest pass it. On the other hand, in some important instances the trace of $\mathbf{K}$ fails to exist, so that the first terms in the two Fredholm expansions are not well defined. In that case there is a simple remedy: in (C.11), replace $\mathfrak{y}_{n+1}$ by

$$\mathfrak{y}'_{n+1}(\vec{r}, \vec{r}'; \vec{r}_1, \ldots, \vec{r}_n) \stackrel{\text{def}}{=} \begin{vmatrix} K(\vec{r}, \vec{r}') & K(\vec{r}, \vec{r}_1) & \ldots & K(\vec{r}, \vec{r}_n) \\ K(\vec{r}_1, \vec{r}') & 0 & \ldots & K(\vec{r}_1, \vec{r}_n) \\ \vdots & \vdots & \ldots & \vdots \\ K(\vec{r}_n, \vec{r}') & K(\vec{r}_n, \vec{r}_1) & \ldots & 0 \end{vmatrix}, \tag{C.14}$$

and in (C.12) replace $\mathfrak{y}_n$ by

$$\mathfrak{y}''_n(\vec{r}_1, \ldots, \vec{r}_n) \stackrel{\text{def}}{=} \begin{vmatrix} 0 & K(\vec{r}_1, \vec{r}_2) & \ldots & K(\vec{r}_1, \vec{r}_n) \\ K(\vec{r}_2, \vec{r}_1) & 0 & \ldots & K(\vec{r}_1, \vec{r}_n) \\ \vdots & \vdots & \ldots & \vdots \\ K(\vec{r}_n, \vec{r}_1) & K(\vec{r}_n, \vec{r}_2) & \ldots & 0 \end{vmatrix}. \tag{C.15}$$

The so defined *modified first Fredholm minor*,

$$Y'(\vec{r}, \vec{r}') \stackrel{\text{def}}{=} K(\vec{r}, \vec{r}') + \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} \int d^3r_1 \cdots d^3r_n \, \mathfrak{y}'_{n+1}(\vec{r}, \vec{r}'; \vec{r}_1, \ldots, \vec{r}_n), \tag{C.16}$$

and the *modified Fredholm determinant*,

$$\det_2(\mathbb{1} - \lambda K) \stackrel{\text{def}}{=} \Delta' = 1 + \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} \int d^3r_1 \cdots d^3r_n \, \mathfrak{y}''_n(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_n), \tag{C.17}$$

neither of which contains $K(\vec{r},\vec{r})$, can be used just like $\mathbf{Y}$ and Δ; that is,

$$\mathbf{M}^{-1} = \mathbb{1} + \frac{\lambda}{\Delta'}\mathbf{Y}'.$$

In fact, it turns out that if tr$\mathbf{K}$ exists, then

$$\mathbf{Y}' = \mathbf{Y}e^{\lambda \mathrm{tr}\mathbf{K}}, \qquad \Delta' = \Delta e^{\lambda \mathrm{tr}\mathbf{K}}, \tag{C.18}$$

so that $\mathbf{Y}/\Delta = \mathbf{Y}'/\Delta'$ is an identity. On the other hand, since neither $\mathbf{Y}'$ nor Δ' contain tr$\mathbf{K}$, they will exist even when tr$\mathbf{K}$ does not.

As for the second question, *the series in (C.3), (C.11), and (C.12) converge absolutely for all complex values of* λ.[2] Therefore, the modified Fredholm procedure can be used for all integral kernels in the Hilbert-Schmidt class.

If, for a specific value $\lambda = 1/\alpha$, the Fredholm determinant (or its modified form) vanishes,

$$\det[\mathbb{1} - (1/\alpha)\mathbf{K}] = 0, \tag{C.19}$$

then the homogeneous form of (C.1) has a nontrivial solution; that is to say, α is an eigenvalue of $\mathbf{K}$. So $\Delta(\lambda)$ [or $\Delta'(\lambda)$] serves both to construct the solution of (C.1) and also to find the eigenvalues of the operator $\mathbf{K}$.

C.2 The Born Series

Another possibility of solving the integral equation (C.1) is to expand the solution in a power series in λ—this is called a *Neumann series* by mathematicians and a *Born series* by physicists—i.e., to expand

$$(\mathbb{1} - \lambda\mathbf{K})^{-1} = \sum_{n=0}^{\infty} \lambda^n \mathbf{K}^n,$$

which amounts to solving (C.1) by iteration. What is the radius of convergence of this power series? If a function $F(\lambda)$ is holomorphic in a neighborhood of the origin and its singularity closest to 0 is λ_0, then the power-series expansion of F converges absolutely for all $|\lambda| < |\lambda_0|$. Now, it follows from the Fredholm method that the expansion of the numerator function $\mathbf{Y}$ in $(\mathbb{1} - \lambda\mathbf{K})^{-1}$ has an infinite radius of convergence; therefore, the relevant question here is, what is the radius of convergence R of the power series expansion of $1/\Delta(\lambda)$? Since the zero of Δ that is closest to the origin is $\lambda_0 = 1/\alpha_0$, where α_0 is the eigenvalue of K with the largest modulus, the answer is simply $R = 1/|\alpha_0|$. In contrast to the Fredholm procedure, which

[2] We shall not prove this here; for a proof, see, for example, [Newton 82], Section 9.3, where (C.11), (C.12), and (C.18) are also proved.

converges for all values of λ, the Born series therefore always has a finite radius R of convergence; for $|\lambda| > R$ it diverges. While this expansion is simpler than the Fredholm method, its general lack of convergence is a serious drawback.

C.3 Application to the Lippmann-Schwinger Equation

In the case of the Lippmann-Schwinger equation, none of the assumptions made in the general theory given above are satisfied. However, if (4.122) is multiplied by $|2MV(\vec{r})/\hbar^2|^{1/2}e^{-i\vec{k}\cdot\vec{r}}$, one obtains the equation

$$\phi(\vec{k},\vec{r}) = |2MV(\vec{r})\hbar^2|^{1/2} + \int d^3r'\, K(\vec{k};\vec{r},\vec{r}\,')\phi(\vec{k},\vec{r}\,'), \tag{C.20}$$

where

$$K(\vec{k};\vec{r},\vec{r}\,') \stackrel{\text{def}}{=} -2M|V(\vec{r})V(\vec{r}\,')|^{1/2}v(\vec{r}\,')\frac{e^{i[k|\vec{r}-\vec{r}\,'|-\vec{k}\cdot(\vec{r}-\vec{r}\,')]}}{4\pi\hbar^2|\vec{r}-\vec{r}\,'|} \tag{C.21}$$

and $v(\vec{r}) \stackrel{\text{def}}{=} V(\vec{r})/|V(\vec{r})|$. Therefore, if $V \in L^1(\mathbb{R}^3)$, then the inhomogeneous term is square-integrable, and if V is in the *Rollnik class,* defined by[3]

$$\int d^3r d^3r'\, \frac{|V(\vec{r})V(\vec{r}\,')|}{|\vec{r}-\vec{r}\,'|^2} < \infty, \tag{C.22}$$

then $\mathbf{K}$ is Hilbert-Schmidt,

$$\mathrm{tr}\mathbf{K}\mathbf{K}^\dagger = \int d^3r d^3r'\, |K(\vec{k};\vec{r},\vec{r}\,')|^2 < \infty$$

for all k such that $\Im k \geq 0$. On the other hand, $K(\vec{k};\vec{r},\vec{r})$ is infinite, so it is appropriate to apply the modified Fredholm method to solve (C.20) and

$$\Delta(k) \stackrel{\text{def}}{=} \det_2[\mathbb{1} - \mathbf{K}(k)]$$

is well defined and an analytic function of k that is holomorphic in the upper half plane (since each term in its Fredholm expansion is analytic in k). As a result, the solution ϕ of (C.20) exists, is unique and square-integrable for all $\Im k \geq 0$, except for values of $k = k_n$ such that $\Delta(k_n) = 0$. For such k_n, the homogeneous version of (C.20) has a nontrivial square-integrable solution,

[3]For example, if for some $\varepsilon, a, C > 0$ and all $\vec{r}$, $V(\vec{r}) < C(a+r)^{-3-\varepsilon}$, then V is in the Rollnik class. **Prove this.**

which, for $\Im k > 0$, leads to a square-integrable solution of (4.122) and thus a bound state. It follows that the zeros of $\Delta(k)$ in the upper half plane must be confined to the imaginary axis, so that the bound-state energies $E_n^{\mathrm{bd}} = \hbar^2 k_n^2/2M$ are real. For potentials in the Rollnik class, there can be no positive point eigenvalues, and hence no bound states of positive energy.

This page intentionally left blank

D
Special Functions

D.1 Spherical Harmonics

The partial differential equation (4.9) can be separated by setting $Y(\theta, \varphi) = f(\theta)g(\varphi)$, which leads to the equation

$$\frac{\sin\theta}{f}\frac{d}{d\theta}\left(\sin\theta\frac{df}{d\theta}\right) - \lambda\sin^2\theta = -\frac{1}{g}\frac{d^2g}{d\varphi^2}.$$

From this we conclude that both sides must be equal to a constant m^2:

$$\frac{d^2g}{d\varphi^2} = -m^2g, \tag{D.1}$$

and

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{df}{d\theta}\right) - \frac{m^2}{\sin^2\theta}f - \lambda f = 0. \tag{D.2}$$

Equation (D.1) is immediately solved by

$$g = e^{im\varphi},$$

and if Y is to be single-valued as φ increases by 2π, m must be either zero or a positive or negative integer: $m = 0, \pm1, \pm2, \ldots$. Equation (D.2), on the other hand, can be rewritten in the form of Legendre's differential equation by setting $\cos\theta = x$,

$$\frac{d}{dx}(1-x^2)\frac{d}{dx}f(x) - \left(\lambda + \frac{m^2}{1-x^2}\right)f(x) = 0. \tag{D.3}$$

Note that the differential equation (D.3) remains unchanged if x is replaced by $-x$. Therefore, the solutions may be chosen to be either even or odd functions of x.

D.1.1 Legendre polynomials

Consider first the case of $m = 0$. The Frobenius *ansatz* $f(x) = x^\sigma \sum_n a_n x^n$ then leads to the recursion relation[1]

$$(\sigma + n + 1)(\sigma + n + 2)a_{n+2} = [(\sigma + n)(\sigma + n + 1) + \lambda]a_n \tag{D.4}$$

with the indicial equation $\sigma(\sigma - 1) = 0$, which assures that $a_{-2} = 0$ and the series starts with $n = 0$. Therefore either $\sigma = 0$ or $\sigma = 1$; the choice of $\sigma = 0$ yields an even solution, whose expansion in powers of x^2 begins with the constant term a_0 and

$$a_{n+2} = \frac{n(n+1) + \lambda}{(n+1)(n+2)} a_n \quad n = 0, 2, \ldots;$$

that of $\sigma = 1$ leads to an odd solution, for which

$$a_{n+2} = \frac{(n+2)(n+1) + \lambda}{(n+2)(n+3)} a_n \quad n = 0, 2, \ldots,$$

the power series in x^2 being multiplied by x. For $n \to \infty$ in both cases, $a_{n+2}/a_n \sim 1$, so that the tail end of the series goes like $\sum_k x^{2k}$, which diverges as $x \to \pm 1$. In order for the solution to be well-behaved at $x = \pm 1$ the series must therefore be required to terminate and become a polynomial, which can happen only if, for some even integer n either $\lambda = -n(n+1)$ or $\lambda = -(n+1)(n+2)$. So we must have $\lambda = -l(l+1)$, $l = 0, 1, 2, \ldots$, and the solution is a polynomial of degree l, called the *Legendre polynomial* $P_l(x)$, with the parity $(-1)^l$:

$$P_l(-x) = (-1)^l P_l(x). \tag{D.5}$$

They are conventionally normalized by setting $P_l(1) = 1$.

A useful way of generating the Legendre polynomials is by expanding the *generating function*

$$T(x, y) \stackrel{\text{def}}{=} \frac{1}{\sqrt{1 - 2xy + y^2}}, \tag{D.6}$$

in a power series in y for $|y| < 1$,

$$T(x, y) = \sum_{l=0}^{\infty} y^l P_l(x). \tag{D.7}$$

[1] **Show this.**

That the coefficients are indeed the Legendre polynomials is proved by means of the partial differential equation[2]

$$\left[(1-x^2)\frac{\partial^2}{\partial x^2} - 2x\frac{\partial}{\partial x} + y\frac{\partial^2}{\partial y^2}y\right] T(x,y) = 0. \tag{D.8}$$

Insertion of the expansion (D.7) in (D.8) shows that the coefficient of y^l must satisfy Legendre's equation (D.3) with $m = 0$; furthermore, for $x = 1$ the expansion (D.7) must go over into that of $1/(1-y)$, which requires that $P_l(1) = 1$. Note that (D.7), written in a different form, means that for $|\mathbf{x}| < |\mathbf{y}|$ and $\mathbf{x}\cdot\mathbf{y} = |\mathbf{x}||\mathbf{y}|\cos\theta$,

$$\frac{1}{|\mathbf{x}-\mathbf{y}|} = \sum_{l=0}^{\infty} |\mathbf{x}|^l |\mathbf{y}|^{-l-1} P_l(\cos\theta). \tag{D.9}$$

The function T also satisfies the equation[3]

$$(1-x^2)\frac{\partial T}{\partial x} = yT + y(y-x)\frac{\partial T}{\partial y}, \tag{D.10}$$

from which we obtain by inserting (D.7),

$$(x^2-1)P_l' = -lP_{l-1} + lxP_l. \tag{D.11}$$

Similarly, the equation[4]

$$(1-xy)\frac{\partial T}{\partial x} - y^2\frac{\partial T}{\partial y} = yT \tag{D.12}$$

leads to

$$P_l' = xP_{l-1}' + lP_{l-1}. \tag{D.13}$$

Other recursion relations for the Legendre polynomials are

$$(l+1)P_{l+1} - (2l+1)xP_l + lP_{l-1} = 0, \tag{D.14}$$

$$xP_l' - P_{l-1}' = lP_l, \tag{D.15}$$

$$P_{l+1}' - P_{l-1}' = (2l+1)P_l, \tag{D.16}$$

the last two of which follow from the others.

The following representation, called *Schläfli's integral*, sometimes comes in handy:

$$P_l(x) = \frac{1}{2^l}\frac{1}{2\pi i}\oint dz \frac{(z^2-1)^l}{(z-x)^{l+1}}, \tag{D.17}$$

[2] **Check this.**
[3] **Check this.**
[4] **Check this.**

where the contour integral in the complex plane runs counter clock wise around the point x. It is proved by inserting it in the left-hand side of (D.3) for $m = 0$ and finding

$$(1-x^2)P_l'' - 2xP_l' + l(l+1)P_l = \frac{1-l}{2^l(2\pi i)} \oint dz \frac{d}{dz} \frac{(z^2-1)^{l+1}}{(z-x)^{l+2}} = 0.$$

Furthermore it is easily seen that for $x = 1$ the right-hand side of (D.17) equals 1.

Equation (D.17) can also be written in the form

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx}\right)^l \frac{1}{2\pi i} \oint dz \frac{(z^2-1)^l}{z-x},$$

and this integral is readily carried out, with the result

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx}\right)^l (x^2-1)^l, \tag{D.18}$$

which is known as *Rodrigues's formula.* It follows from this formula that $P_l(x)$ has exactly l zeros between -1 and $+1$.[5] An alternative way of writing it is

$$P_l(\cos\theta) = \frac{(-1)^l}{2^l l!} \left(\frac{d}{d\cos\theta}\right)^l (\sin\theta)^{2l}. \tag{D.19}$$

That two Legendre polynomials of different degrees are mutually orthogonal,

$$\int_{-1}^{1} dx\, P_l(x) P_{l'}(x) = 0, \qquad l \neq l',$$

follows directly from multiplying Legendre's differential equation (D.3) for P_l by $P_{l'}$, integrating, and subtracting the same equation for $P_{l'}$, multiplied by P_l and integrated, using an integration by parts. The normalization integral is most simply calculated by using the generating function and the previous result:

$$\begin{aligned}\int_{-1}^{1} dx \frac{1}{\sqrt{1-2xy+x^2}^2} &= \frac{1}{x}\log\frac{1+y}{1-y} = 2\sum_l \frac{y^{2l}}{2l+1} \\ &= \sum_{ll'} y^{l+l'} \int_{-1}^{1} dx\, P_l(x) P_{l'}(x) = \sum_l y^{2l} \int_{-1}^{1} dx\, P_l^2(x),\end{aligned}$$

from which we conclude that

$$\int_{-1}^{1} dx\, P_l(x) P_{l'}(x) = \frac{2}{2l+1}\delta_{ll'}. \tag{D.20}$$

[5] **Why?**

The Legendre polynomials form an orthogonal basis on the Hilbert space $L^2(-1,1)$. Furthermore, every polynomial of degree n can be expressed as a linear combination of Legendre polynomials of order $l \leq n$.

D.1.2 Associated Legendre functions

The associated Legendre functions P_l^m are defined for $0 \leq m \leq l$ in terms of the Legendre polynomials by

$$P_l^m(x) \stackrel{\text{def}}{=} (1-x^2)^{m/2} \left(\frac{d}{dx}\right)^m P_l(x), \tag{D.21}$$

or

$$P_l^m(\cos\theta) \stackrel{\text{def}}{=} (\sin\theta)^m \left(\frac{d}{d\cos\theta}\right)^m P_l(\cos\theta). \tag{D.22}$$

[One of the exercises for this Appendix requires you to show that we must always have $|m| \leq l$, and (D.21) gives a nonzero result only in that case.] For $m < 0$, it is convenient to define

$$P_l^{-|m|}(\cos\theta) \stackrel{\text{def}}{=} (-1)^m \frac{(l-|m|)!}{(l+|m|)!} P_l^{|m|}(\cos\theta), \tag{D.23}$$

so that the following formula, derived from (D.21) and Rodrigues's (D.18),

$$P_l^m(\cos\theta) = \frac{(-1)^{l+m}}{2^l l!} \frac{(l+m)!}{(l-m)!} (\sin\theta)^{-m} \left(\frac{d}{d\cos\theta}\right)^{l-m} (\sin\theta)^{2l}, \tag{D.24}$$

holds for $-l \leq m \leq l$. The function $P_l^m(x)$ has the parity $(-1)^{l+m}$:

$$P_l^m(-x) = (-1)^{l+m} P_l^m(x). \tag{D.25}$$

D.1.3 Spherical harmonics

We are now ready to assemble the functions $P_l^m(\cos\theta)$ and $e^{im\varphi}$ into spherical harmonics, adopting a convenient phase convention (which varies from one author to another):

$$Y_l^m(\widehat{n}) \stackrel{\text{def}}{=} Y_l^m(\theta,\varphi) = (-1)^m i^l \left[\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}\right]^{1/2} e^{im\varphi} P_l^m(\cos\theta), \tag{D.26}$$

where θ and φ are the polar angles of the unit vector $\widehat{n}$. These functions have the complex conjugation property

$$Y_l^{m*} = (-1)^{l+m} Y_l^{-m}, \tag{D.27}$$

the parity $(-1)^l$

$$Y_l^m(-\widehat{n}) = (-1)^l Y_l^m(\widehat{n}), \tag{D.28}$$

and they form an orthonormal basis in the Hilbert space of square-integrable functions on the unit sphere:

$$\int d\Omega\, Y_l^{m*}(\widehat{n}) Y_{l'}^{m'}(\widehat{n}) = \delta_{ll'}\delta_{mm'}, \tag{D.29}$$

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_l^m(\widehat{n}) Y_l^{m*}(\widehat{n}') = \delta^2(\widehat{n}, \widehat{n}'), \tag{D.30}$$

where $d\Omega$ is the solid-angle element and $\delta^2(\widehat{n}, \widehat{n}')$ is the Dirac delta-function on the unit sphere such that

$$\int d\Omega' \delta^2(\widehat{n}, \widehat{n}') f(\widehat{n}') = f(\widehat{n}).$$

The meaning of the symbolic (D.30) is that every square-integrable function $f(\widehat{n})$ on the unit sphere can be expanded in the form

$$f(\widehat{n}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_l^m Y_l^m(\widehat{n}), \text{ where } a_l^m = \int d\widehat{n}\, Y_l^{m*}(\widehat{n}) f(\widehat{n}). \tag{D.31}$$

The Legendre polynomial $P_l(\widehat{n} \cdot \widehat{n}') = P_l(\cos\theta)$, where θ is the angle between $\widehat{n}$ and $\widehat{n}'$, can be expressed in terms of spherical harmonics of $\widehat{n}$ and $\widehat{n}'$ of the same order:

$$4\pi \sum_{m=-l}^{l} Y_l^m(\widehat{n}) Y_l^{m*}(\widehat{n}') = (2l+1) P_l(\widehat{n} \cdot \widehat{n}'), \tag{D.32}$$

which implies that any well-behaved rotationally invariant function of $\widehat{n}$ and $\widehat{n}'$ can be expanded in terms of spherical harmonics of equal m-values:

$$f(\widehat{n}, \widehat{n}') = \sum_{l} \sum_{m=-l}^{l} c_l Y_l^m(\widehat{n}) Y_l^{m*}(\widehat{n}'),$$

with coefficients c_l that are independent of m.

D.1.4 Zonal harmonics

Instead of forming complex spherical harmonics, it is also possible to assemble *real* solutions of Eq. (4.9). This is done by choosing as the two linearly independent solutions of (D.1), $g = \sin(m\varphi)$ and $g = \cos(m\varphi)$ rather than $g = e^{\pm im\varphi}$. The resulting real functions

$$Z_l^{sm}(\theta, \varphi) = \sqrt{\frac{2l+1}{2\pi} \frac{(l-m)!}{(l+m)!}} \sin(m\varphi) P_l^m(\cos\theta), \tag{D.33}$$

$$Z_l^{cm}(\theta, \varphi) = \sqrt{\frac{2l+1}{2\pi} \frac{(l-m)!}{(l+m)!}} \cos(m\varphi) P_l^m(\cos\theta), \tag{D.34}$$

are called *zonal harmonics.* Their nodal surfaces are circular cones about the z-axis (the zeros of the Legendre functions) and planes through the z-axis (the zeros of the trigonometric functions).

D.2 Spherical Bessel Functions

Dividing Eq. (4.19) by k^2 and setting $kr \stackrel{\text{def}}{=} x$ transforms it into the simpler equation

$$u'' - \frac{l(l+1)}{x^2}u + u = 0, \tag{D.35}$$

which, upon defining $U(x) \stackrel{\text{def}}{=} x^{-1/2}u(x)$, turns into *Bessel's equation,*

$$x^2U'' + xU' + [x^2 - (l + \frac{1}{2})^2]U = 0.$$

Therefore, the radial functions that solve (4.10) for $\lambda = -l(l+1)$ differ by a factor of $x^{-1/2}$ from Bessel functions of half-integral order. Since $x = 0$ is a regular singular point of this differential equation, there is a regular and an irregular solution. The regular solution is conventionally defined in terms of a Bessel function by

$$j_l(x) \stackrel{\text{def}}{=} \sqrt{\frac{\pi}{2x}}J_{l+\frac{1}{2}}(x) \tag{D.36}$$

and is called a *spherical Bessel function.* Two kinds of irregular solutions are the *spherical Neumann function,*

$$n_l(x) \stackrel{\text{def}}{=} (-1)^{l+1}\sqrt{\frac{\pi}{2x}}J_{-l-\frac{1}{2}}(x), \tag{D.37}$$

and the *spherical Hankel function,*

$$h_l^{(1)}(x) \stackrel{\text{def}}{=} \sqrt{\frac{\pi}{2x}}H_{l+\frac{1}{2}}^{(1)}(x) = j_l(x) + in_l(x). \tag{D.38}$$

The functions j_l and n_l are both real and of the general form $P(x)\sin x + Q(x)\cos x$, where P and Q are polynomials in $1/x$, while $h_l^{(1)}(x)$ has the form of a polynomial in $1/x$ times e^{ix} such that $h_l^{(1)}(ix)$ is real.

The solutions of (D.35) differ from the functions defined above by a factor of x, and they are called *Riccati-Bessel, Riccati-Neumann,* and *Riccati-Hankel functions*, respectively:

$$u_l(x) \stackrel{\text{def}}{=} xj_l(x),$$

$$v_l(x) \stackrel{\text{def}}{=} xn_l(x),$$

and

$$w_l^{(+)}(x) \stackrel{\text{def}}{=} i(-1)^l x h_l^{(1)}(x) = (-1)^{l+1}[v_l(x) - iu_l(x)].$$

Making the Frobenius *ansatz* $u(x) = x^\sigma \sum_n a_n x^n$ near the origin to solve (D.35), we obtain[6] the indicial equation $\sigma(\sigma-1) = l(l+1)$, whose solutions are $\sigma = l+1$ and $\sigma = -l$. Thus the regular solution goes as x^{l+1} near the origin, and the irregular solution goes as x^{-l}. More precisely,

$$u_l(x) = \frac{x^{l+1}}{(2l+1)!!} + \dots, \tag{D.39}$$

where $(2l+1)!! \stackrel{\text{def}}{=} 1 \cdot 3 \cdot 5 \cdots (2l+1)$, and

$$v_l(x) = -(2l-1)!! x^{-l} + \dots. \tag{D.40}$$

The function $u_l(x)$ has the parity $(-1)^{l+1}$ and $v_l(x)$ has the parity $(-1)^l$. [The functions $x^{-l-1}u_l(x)$ and $x^l v_l(x)$ can be expanded in power series in x^2.] The asymptotic behavior of these functions as $x \to \infty$ is as follows:

$$u_l(x) \simeq \sin(x - \frac{1}{2}\pi l), \;\; v_l(x) \simeq -\cos(x - \frac{1}{2}\pi l), \;\; w_l^{(+)}(x) \simeq e^{ix + \frac{1}{2}\pi l}. \tag{D.41}$$

For $l = 0$ these functions are simply $u_0 = \sin x$, $v_0(x) = -\cos x$, and $w_0(x) = e^{ix}$.

D.3 Hermite Polynomials

The Hermite polynomials satisfy the differential equation on the real axis, $-\infty < x < \infty$,

$$H'' - 2xH' + (\lambda - 1)H = 0. \tag{D.42}$$

In order for a solution of this equation not to grow like e^{x^2} at infinity (and thereby overwhelm the factor $e^{-\frac{1}{2}x^2}$ needed to form a solution of the Schrödinger equation for a simple harmonic oscillator), λ must be an odd integer, $\lambda = 2n+1$, $n = 0, 1, \dots$, in which case H is a polynomial of degree n and with the parity of n.[7] [Note that (D.42) is invariant under reflection, so you expect its solutions to be odd or even functions.] They can be defined by the generating function

$$S(t,x) \stackrel{\text{def}}{=} e^{-t^2+2tx} = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x), \tag{D.43}$$

[6] **Show it.**

[7] **Prove these assertions.**

To show that the coefficients H_n satisfy (D.42), insert the expansion in the differential equation

$$\frac{\partial}{\partial x} S = 2tS;$$

comparing coefficients leads to the conclusion that

$$H_n'(x) = 2nH_{n-1}(x), \tag{D.44}$$

whereas its insertion in the equation

$$\frac{\partial}{\partial t} S = 2(x-t)S$$

leads to

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x). \tag{D.45}$$

Differentiation of (D.44) and use of (D.45) then leads to (D.42). Furthermore, since

$$\frac{\partial^n}{\partial x^n} S = (2t)^n S,$$

the expansion of which in powers of t starts with t^n, we conclude that $(d/dx)^n H_m(x) = 0$ for $m < n$, so that the coefficient $H_m(x)$ must be a polynomial of order m. The normalization of the Hermite polynomials is fixed by (D.43).

Equation (3.30) implies that $H_{n+1}(x) = e^{\frac{1}{2}x^2}(x - d/dx)e^{-\frac{1}{2}x^2}H_n(x)$, and therefore

$$H_{n+1}(x) = \left(2x - \frac{d}{dx}\right) H_n(x), \tag{D.46}$$

which can also be written in the form

$$H_{n+1}(x) = -e^{x^2} \frac{d}{dx}\left(e^{-x^2} H_n(x)\right), \tag{D.47}$$

implying that

$$H_n(x) = (-1)^n e^{x^2} \left(\frac{d}{dx}\right)^n e^{-x^2}. \tag{D.48}$$

The first few of these polynomials are given by

$$\begin{aligned} H_0 &= 1 \\ H_1 &= 2x \\ H_2 &= 4x^2 - 2. \end{aligned}$$

The Hermite polynomials are mutually orthogonal with the weight function e^{-x^2}. This can be proved by means of the generating function (D.43):

$$\int_{-\infty}^{\infty} dx\, e^{-t^2+2tx} e^{-t'^2+2t'x} e^{-x^2} = e^{2tt'} \int_{-\infty}^{\infty} dx\, e^{-[x^2-2x(t+t')+(t+t')^2]}$$

$$= e^{2tt'} \int_{-\infty}^{\infty} dx\, e^{-x^2} = e^{2tt'} \sqrt{\pi} = \sqrt{\pi} \sum_{0}^{\infty} \frac{(2tt')^n}{n!}$$
$$= \sum_{nm} \frac{t^n t'^m}{n!m!} \int_{-\infty}^{\infty} dx\, e^{-x^2} H_n(x) H_m(x),$$

from which we conclude that

$$\int_{-\infty}^{\infty} dx\, e^{-x^2} H_n(x) H_m(x) = \delta_{nm} \sqrt{\pi} 2^n n!. \tag{D.49}$$

In a similar manner we find that

$$\int_{-\infty}^{\infty} dx\, x^2 e^{-x^2} H_n^2(x) = \sqrt{\pi} 2^n n! (n + \frac{1}{2}). \tag{D.50}$$

The value of $H_n(x)$ at the origin can also be obtained from the generating function, and we find that

$$H_n(0) = \begin{cases} 0 & \text{for odd } n \\ (-1)^{n/2} n!/(n/2)! & \text{for even } n \end{cases} \tag{D.51}$$

Finally, the Hermite polynomials form a complete set in the sense

$$\sum_{n=0}^{\infty} \frac{\exp[-\frac{1}{2}(x^2 - x'^2)]}{\sqrt{\pi} 2^n n!} H_n(x) H_n(x') = \delta(x - x'), \tag{D.52}$$

meaning that for all $f \in L^2(I\!R)$,

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{\sqrt{\pi} 2^n n!} e^{-\frac{1}{2}x^2} c_n H_n(x), \tag{D.53}$$

where

$$c_n = \int_{-\infty}^{\infty} dx\, f(x) e^{-\frac{1}{2}x^2} H_n(x). \tag{D.54}$$

D.4 The Hypergeometric Function

The hypergeometric function is defined by the *hypergeometric series*

$$F(a, b; c; z) \stackrel{\text{def}}{=} 1 + \frac{ab}{c} z + \frac{a(a+1)b(b+1)}{2!c(c+1)} z^2 + \dots \tag{D.55}$$

and it satisfies the differential equation

$$z(1 - z)F'' + [c - (a + b + 1)z]F' - abF = 0. \tag{D.56}$$

Setting $z = x/b$ and $F(z) = G(x)$, we obtain

$$x\left(1-\frac{x}{b}\right)G'' + \left(c - \frac{a+b+1}{b}x\right)G' - aG = 0,$$

which, in the limit as $b \to \infty$, goes over into the equation

$$xG'' + (c-x)G' - aG = 0, \tag{D.57}$$

whose regular solution

$$\begin{aligned} F(a|\,c\,|x) &\stackrel{\text{def}}{=} \lim_{b\to\infty} F(a,b;c;\frac{x}{b}) = 1 + \frac{a}{c}x + \frac{a(a+1)}{2!c(c+1)}x^2 + \dots \\ &= \sum_{s=0}^{\infty} \frac{\Gamma(a+s)\Gamma(c)}{\Gamma(c+s)\Gamma(a)}\frac{x^s}{s!} \end{aligned} \tag{D.58}$$

is called the *confluent hypergeometric function.*

The asymptotic behavior of this function for large x is as follows:

$$\begin{aligned} F(a|\,c\,|x) \;\simeq\; & \Gamma(c)\left\{\frac{1}{\Gamma(a)}e^{x+(a-c)\log x}\left(1 + \frac{(1-a)(c-a)}{x} + \dots\right)\right. \\ & \left. + \frac{1}{\Gamma(c-a)}e^{-a\log(-x)}\left(1 + \frac{a(a-c+1)}{x} + \dots\right)\right\} \end{aligned} \tag{D.59}$$

D.5 Laguerre Polynomials

The Laguerre polynomials are defined by the generating function

$$U(t,x) \stackrel{\text{def}}{=} \frac{e^{-tx/(1-t)}}{1-t} = \sum_{n=0}^{\infty}\frac{t^n}{n!}L_n(x), \tag{D.60}$$

which shows that

$$L_n(0) = n!. \tag{D.61}$$

Since U satisfies the partial differential equation[8]

$$t\frac{\partial}{\partial t}U + (1-x)\frac{\partial}{\partial x}U + x\frac{\partial^2}{\partial x^2}U = 0,$$

the coefficients solve the differential equation

$$x\frac{d^2}{dx^2}L_n(x) + (1-x)\frac{d}{dx}L_n(x) + nL_n(x) = 0, \tag{D.62}$$

[8] **Check this.**

which has a regular singular point at $x = 0$. The generating function leads to the following recurrence relations,

$$L'_n = nL'_{n-1} - nL_{n-1}, \tag{D.63}$$

$$L_{n+1} = (2n + 1 - x)L_n - n^2 L_{n-1}, \tag{D.64}$$

and they can be explicitly expressed in the form

$$L_n(x) = e^x \frac{d^n}{dx^n}(e^{-x}x^n) = \left(\frac{d}{dx} - 1\right)^n x^n, \tag{D.65}$$

which shows them to be polynomials of order n.

The *associated Laguerre polynomials*[9] are defined by

$$L_n^m(x) = (-1)^m \frac{d^m}{dx^m} L_{n+m}(x), \tag{D.66}$$

which shows that they are also polynomials of degree n, and they satisfy the differential equation (sometimes called *Laplace's equation*)

$$x\frac{d^2}{dx^2}L_n^m(x) + (m + 1 - x)\frac{d}{dx}L_n^m(x) + nL_n^m(x) = 0. \tag{D.67}$$

They are special cases of confluent hypergeometric functions defined by (D.58),

$$L_n^m(x) = \frac{[(n + m)!]^2}{n!m!} F(-n \,|\, m + 1 \,|\, x), \tag{D.68}$$

and their generating function is

$$\frac{e^{-tx/(1-t)}}{(1 - t)^{m+1}} = \sum_{n=0}^{\infty} \frac{t^n}{(n + m)!} L_n^m(x), \tag{D.69}$$

from which their normalization integral can be obtained as

$$\int_0^\infty dx\, x^p e^{-x} L_n^p(x) L_m^p(x) = \delta_{mn} \frac{[(p + n)!]^3}{n!}. \tag{D.70}$$

When n is not an integer, these functions are still defined by (D.68), but they are not polynomials, and the integrals in (D.70) diverge.

[9] Some authors use $(-1)^m L_{n-m}^m$ to denote the polynomial here denoted by L_n^m.

D.6 Problems and Exercises

1. Derive the recursion relation (D.4) for the coefficients in the power series expansion $f(x) = x^\sigma \sum_0^\infty a_n x^n$ of a solution of Legendre's differential equation for $m = 0$.

2. Let $f(x)$ be a solution of Legendre's equation for $m > 0$ and define $g(x) = (1 - x^2)^{-m/2} f(x)$. Find the differential equation satisfied by g, expand it in the form $g(x) = x^\sigma \sum_0^\infty b_n x^n$, and find the recursion relation for the coefficients b_n. Show thereby again that in order for g to be continuous at $x = \pm 1$, the expansion must break up, and this happens if and only if l is an integer such that $l \geq m$.

3. Show that the function (D.6) satisfies the partial differential equation (D.8) and that, if $T(x, y)$ is expanded in a power series in y, the coefficients, as functions of x, must satisfy Legendre's equation for $m = 0$.

4. Show that $T(x, y)$ satisfies the partial differential equation (D.10) and use this equation to derive (D.11).

5. Show that $T(x, y)$ satisfies the partial differential equation (D.12) and use it to derive (D.13).

6. Show that the function P_l^m defined by (D.21) satisfies Legendre's equation. (Hint: Use Schläfli's integral representation.)

7. Let P_l be the lth Legendre polynomial. Show that $\vec{\mathbf{L}}^2 P_l(\vec{r} \cdot \vec{r}') = \vec{\mathbf{L}}'^2 P_l(\vec{r} \cdot \vec{r}')$ and $(\vec{\mathbf{L}} + \vec{\mathbf{L}}') P_l(\vec{r} \cdot \vec{r}') = 0$. Use these two equations to prove (D.32).

8. Show that the Hermite polynomial $H_n(x)$ has n real zeros. (Hint: Use (D.47).)

9. Show that
$$H_n(x) = \left(2x - \frac{d}{dx}\right)^n 1.$$

10. Prove (D.50).

11. Prove (D.51).

12. Prove that the associated Laguerre polynomial $L_n^m(x)$ has n positive zeros.

This page intentionally left blank

E
Group Representations

E.1 Groups

A group $\mathcal{G}$ is a collection of mathematical objects, such as transformations of a physical system or of a reference frame, for which an associative binary operation called *multiplication* is defined [i.e., $(AB)C = A(BC)$] that satisfies the following axioms:

1. The collection is closed, i.e., if $A \in \mathcal{G}$ and $B \in \mathcal{G}$ then $AB \in \mathcal{G}$.
2. $\mathcal{G}$ contains a unit element E such that for all $A \in \mathcal{G}$, $EA = AE = A$.
3. For every $A \in \mathcal{G}$ there is a unique *inverse* A^{-1} such that $A^{-1}A = AA^{-1} = E$.

There is no general requirement for multiplication to be commutative, $AB = BA$; if all elements of a group $\mathcal{G}$ commute, $\mathcal{G}$ is called *Abelian.* The number of elements in $\mathcal{G}$ is called its *order.* Any subset of a group $\mathcal{G}$ that forms a group by itself (with the multiplication law inherited from $\mathcal{G}$) is called a *subgroup*; its order must be a divisor of the order of $\mathcal{G}$ (Lagrange's theorem). Two groups are called *isomorphic* if there is a one-to-one mapping between them that preserves multiplication—i.e., if $A \mapsto A'$ and $B \mapsto B'$, then $AB \mapsto A'B'$. If the mapping is not one-to-one, so that more than one element of $\mathcal{G}$ is mapped on the same element of $\mathcal{G}'$, they are called *homomorphic.*

The set of permutations of n objects is an example of a group of finite order; this group is called S_n, and its order is $n!$. Cayley's theorem asserts that every group of order n is isomorphic to a subgroup of S_n.

If $A, X \in \mathcal{G}$, then the element XAX^{-1} is called *conjugate* to A: $XAX^{-1} \sim A$. Conjugation is an *equivalence relation:* it is reflexive ($A \sim A$), transitive (if $A \sim B$ and $B \sim C$, then $A \sim C$), and symmetric (if $A \sim B$, then $B \sim A$). Elements of a group that are all conjugate to one another form a *class.* The unit element E is, of course, in a class by itself, and the whole group $\mathcal{G}$ can be divided into disjoint classes. A subgroup that contains all the classes of its elements is called an *invariant*, or *normal*, or *self-conjugate* subgroup, or also a *normal divisor.*

The group S_3 of permutations of three objects may serve as an example. Calling $E = (123)$, $A = (132)$, $B = (213)$, $C = (321)$, $D = (231)$, $F = (312)$, where $(132)[abc] = [acb]$, etc., the multiplication table of the group is as follows:

$$\begin{array}{cccccc} E & A & B & C & D & F \\ A & E & D & F & B & C \\ B & F & E & D & C & A \\ C & D & F & E & A & B \\ D & C & A & B & F & E \\ F & B & C & A & E & D \end{array}$$

This group contains three classes: E; D, F; A, B, C and four subgroups, all of which are Abelian: E, A; E, B; E, C; E, D, F, the last being self-conjugate.[1]

If $\mathcal{G}_1$ and $\mathcal{G}_2$ are two groups of orders n_1 and n_2 with the elements $E_1, A_1, B_1, \ldots$ and $E_2, A_2, B_2, \ldots$, respectively, we can form the *direct product* group $\mathcal{G}_1 \times \mathcal{G}_2$ of order $n_1 n_2$ with the elements $A_1 \times E_2, A_1 \times A_2, A_1 \times B_2, \ldots B_1 \times E_2, \ldots$ and the obvious multiplication law. An example would be the permutations of pairs $(abc, \alpha\beta\gamma\delta)$ of three balls and four cubes, such that

$$(abc, \alpha\beta\gamma\delta) \to (acb, \beta\alpha\gamma\delta) = (132) \times (2134),$$

etc. Then $E_1 \times E_2, E_1 \times A_2, E_1 \times B_2, \ldots$ is isomorphic to $\mathcal{G}_2$, and $E_1 \times E_2, A_1 \times E_2, B_1 \times E_2, \ldots$ is isomorphic to $\mathcal{G}_1$, so that both $\mathcal{G}_1$ and $\mathcal{G}_2$ are (invariant) subgroups of $\mathcal{G}_1 \times \mathcal{G}_2$, and each element in $\mathcal{G}_1 \times \mathcal{G}_2$ is uniquely expressible as a product of elements in $\mathcal{G}_1$ and $\mathcal{G}_2$: $A_1 \times B_2 = (A_1 \times E_2)(E_1 \times B_2)$. Furthermore, the elements in the two subgroups $\mathcal{G}_1$ and $\mathcal{G}_2$ commute. Conversely, if $\mathcal{G}_1$ and $\mathcal{G}_2$ are two mutually commuting subgroups of $\mathcal{G}$, such that each element of $\mathcal{G}$ can be uniquely written as a product of an element of $\mathcal{G}_1$ and an element of $\mathcal{G}_2$, then we say that $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2$.

[1] **Show all of this as an exercise.**

E.2 Representations

A representation $\Gamma(\mathcal{G})$ of the group $\mathcal{G}$ is a mapping of the elements R of $\mathcal{G}$ to a set of square matrices $\mathcal{D}(R)$, such that group multiplication corresponds to matrix multiplication. The n-dimensional linear vector space on which the $n \times n$ matrices act is called the *carrier space* of the representation.

There are two kinds of representations: (a) *isomorphic* or *faithful* ones, i.e., those that are one-to-one, so that each matrix corresponds to one and only one group element, and (b) *homomorphic* or *unfaithful* ones, in which the same matrix may represent more than one group element. It is customary to assume that all the matrices are non-singular and $\mathcal{D}(E) = \mathbf{1}$, where $\mathbf{1}$ is the unit matrix, so that $\mathcal{D}(A^{-1}) = \mathcal{D}^{-1}(A)$. If a representation is unfaithful, more than one group element must be mapped into $\mathbf{1}$.[2] Every group has the trivial (unfaithful) representation that associates each element with the unit matrix.

Here is our first important theorem:

Theorem E.1 *A matrix representation (using non-singular matrices) of a group of finite order can always be made unitary by a canonical transformation.*

To prove this, call the matrices in the representation $\mathcal{A}_g$ and define the Hermitian matrix

$$H \stackrel{\text{def}}{=} \sum_{g\in\mathcal{G}} \mathcal{A}_g \mathcal{A}_g^\dagger,$$

where the summation runs over all the elements of the group. The matrix H can be diagonalized by a unitary matrix U:

$$UHU^{-1} \stackrel{\text{def}}{=} h = \sum_{g\in\mathcal{G}} (U\mathcal{A}_g U^{-1})(U\mathcal{A}_g^\dagger U^{-1}) = \sum_g \mathcal{A}'_g \mathcal{A}'^\dagger_g, \qquad \mathcal{A}'_g \stackrel{\text{def}}{=} U\mathcal{A}_g U^{-1},$$

so that the diagonal elements are given by

$$h_{ii} = \sum_k \sum_{g\in\mathcal{G}} (\mathcal{A}'_g)_{ik} (\mathcal{A}'^\dagger_g)_{ki} = \sum_k \sum_{g\in\mathcal{G}} |(\mathcal{A}'_g)_{ik}|^2 > 0,$$

because $\det \mathcal{A}'_g = \det \mathcal{A}_g \neq 0$. Therefore, we can take the inverse of the square root:

$$\mathbf{1} = h^{-1/2} \sum_g \mathcal{A}'_g \mathcal{A}'^\dagger_g h^{-1/2}.$$

The new matrices

$$\mathcal{A}''_g \stackrel{\text{def}}{=} h^{-1/2} \mathcal{A}'_g h^{1/2} = (h^{-1/2}U)\mathcal{A}_g (h^{-1/2}U)^{-1}$$

are now unitary:

$$\begin{aligned} \mathcal{A}''_g \mathcal{A}''^\dagger_g &= h^{-1/2} \mathcal{A}'_g h \mathcal{A}'^\dagger_g h^{-1/2} = h^{-1/2} \sum_f \mathcal{A}'_g \mathcal{A}'_f \mathcal{A}'^\dagger_f \mathcal{A}'^\dagger_g h^{-1/2} \\ &= h^{-1/2} \sum_f (\mathcal{A}'_g \mathcal{A}'_f)(\mathcal{A}'_g \mathcal{A}'_f)^\dagger h^{-1/2} = h^{-1/2} \sum_{f'} \mathcal{A}'_{f'} \mathcal{A}'^\dagger_{f'} h^{-1/2} = \mathbf{1}, \end{aligned}$$

---

[2] **Prove this.**

because $(\mathcal{A}'_g\mathcal{A}'_f)$ runs through the entire group. □

Any two representations that are related by a canonical transformation, $\mathcal{D}'(R) = S\mathcal{D}(R)S^{-1}$ for all $R \in \mathcal{G}$, are called *equivalent,* written as $\mathcal{D}' \approx \mathcal{D}$; so the theorem says that every representation is equivalent to a unitary one. We may therefore restrict ourselves always to unitary representations.

Suppose now you take two representations of a group $\mathcal{G}$, with matrices $\mathcal{A}_g$ and $\mathcal{B}_g$, and form the new matrices

$$\begin{pmatrix} \mathcal{A}_g & 0 \\ 0 & \mathcal{B}_g \end{pmatrix}$$

with the blocks $\mathcal{A}_g$ and $\mathcal{B}_g$ on the main diagonal, thus forming a new representation. If this construction is followed by a canonical transformation, its origin is no longer apparent.[3] A representation that can be transformed by a canonical transformation (the same for all the group elements) into such block form (the same kinds of blocks for all the elements) is called *fully reducible*[4]; otherwise it is *irreducible.* You can see the significance of this distinction: irreducible representations are the only new ones; a fully reducible representation is just equivalent to one pieced together out of old ones by a direct sum.

For example, the group S_3, whose multiplication table was given above, has the unfaithful two-dimensional representation that assigns the unit matrix to the elements E, D, and F, while A, B, and C are represented by the matrix

$$\mathcal{C} = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix}.$$

The Hermitian matrix $\mathcal{C}$ can be diagonalized by a canonical transformation, which, of course, leaves the unit matrix representing E, D, and F unchanged. Therefore, this representation is reducible; since the eigenvalues of $\mathcal{C}$ are ± 1, it is made up of two one-dimensional representations, one of which is the trivial representation that assigns the number 1 to all six group elements, and the other is $E, D, F = 1$ and $A, B, C = -1$.

Now, think about the carrier space $\mathcal{S}$ of a given representation $\Gamma(\mathcal{G})$, and let $\mathcal{U}$ be a subspace of $\mathcal{S}$. If for every vector $v \in \mathcal{U}$ and every $\mathcal{D}_g \in \Gamma(\mathcal{G})$ we have $\mathcal{D}_g v \in \mathcal{U}$, then $\mathcal{U}$ is called an *invariant subspace* of $\mathcal{S}$ for the group $\mathcal{G}$. The question of whether Γ is reducible is equivalent to the question whether its carrier space has any invariant subspaces for $\mathcal{G}$. If so, we can find an orthogonal basis such that each basis vector lies entirely in one invariant

---

[3] Note that a simultaneous rearrangement of rows and corresponding collumns can be accomplished by a canonical tranformation.

[4] An n-dimensional representation is called *reducible* if it is eqivalent to one in which all the matrices have an $m \times m$, $m < n$, block of zeros in the lower left-hand corner. A reducible unitary representation is necessarily fully reducible.

subspace, i.e., different subsets of these basis vectors span different invariant subspaces; the matrices in Γ are then in block form and the representation has been reduced.[5]

The following lemma is a useful tool for deciding whether a given representation is reducible:

Lemma E.2 *A matrix M that commutes with all the matrices in a given irreducible representation Γ of a group is necessarily a multiple of the unit matrix; and conversely, if the only matrix that commutes with all the matrices in Γ is a multiple of the unit matrix, Γ is irreducible.*

Proof: If the representation is reducible, then there certainly exists a matrix other than $c\mathbf{1}$ that commutes with all the matrices in it. When the representation is in block form, we need only take a diagonal matrix that has different constants on the diagonal for each of the blocks (i.e., it is a different multiple of the units matrix for each of the blocks separately). Therefore, if the only matrix that commutes with them all is of the form $c\mathbf{1}$, then the representation is irreducible. To prove the converse, i.e., that if the representation Γ is irreducible, then there exists no matrix other than a multiple of the unit matrix that commutes with all the matrices in Γ, we assume that the matrix M that commutes with all the matrices in Γ can be diagonalized by a canonical transformation S. (First, make the representation unitary; in this form it follows that both M and $M^\dagger$ commute with all the matrices, and therefore so do the Hermitian matrices $M + M^\dagger$ and $i(M - M^\dagger)$, which can be diagonalized.) Now apply the canonical transformation S to all the matrices in Γ too. But a diagonal matrix M commutes with another matrix only if the latter is in block form, with blocks corresponding to the eigenvalues of M that are equal. Therefore all the matrices in the representation must now be in block form, and the representation is reducible, *unless* all the eigenvalues of M are the same, which makes M a multiple of the unit matrix. □

Lemma E.2 can now be used to prove the following important proposition, known as *Schur's lemma*:

Lemma E.3 *Let $\mathcal{D}^{(1)}(R)$ and $\mathcal{D}^{(2)}(R)$ be the matrices of two irreducible representations of $\mathcal{G}$ of dimensions n_1 and n_2, respectively, and let M be a rectangular matrix so that for all $R \in \mathcal{G}$, $\mathcal{D}^{(1)}(R)M = M\mathcal{D}^{(2)}(R)$. If $n_1 \neq n_2$, then $M = 0$; if $n_1 = n_2$, either $M = 0$ or else it has an inverse, in which case the two representations are equivalent.*

To prove this, assume, without loss of generality, that the representations are unitary. It then follows from $M^\dagger \mathcal{D}^{(2)\dagger}(R) = \mathcal{D}^{(1)\dagger}(R) M^\dagger$ that $M^\dagger \mathcal{D}^{(2)-1}(R) = \mathcal{D}^{(1)-1}(R) M^\dagger$, so that $M^\dagger \mathcal{D}^{(2)}(R^{-1}) = \mathcal{D}^{(1)}(R^{-1}) M^\dagger$ and therefore for all $R \in \mathcal{G}$, $M^\dagger \mathcal{D}^{(2)}(R) =$

[5] A matrix that leaves a subspace invariant does not necessarily leave its orthogonal complement invariant. So it is not necessarily in block form, or fully reduced. However, a unitary matrix does; this is why for a unitary representation, *reducible* implies *fully reducible.*

$\mathcal{D}^{(1)}(R)M^\dagger$, which implies $MM^\dagger\mathcal{D}^{(2)}(R) = M\mathcal{D}^{(1)}(R)M^\dagger = \mathcal{D}^{(2)}(R)MM^\dagger$. Therefore, by Lemma E.2, $MM^\dagger = c\mathbf{1}$. Hence, if $n_1 = n_2$, so that M is a square matrix, either $c = 0$, in which case $\sum_k |M_{ik}|^2 = 0$ and hence $M = 0$, or $c \neq 0$, in which case $M^{-1} = \frac{1}{c}M^\dagger$. If $n_2 \neq n_1$, say, $n_2 > n_1$, we make M square by filling up the square matrix with zeros on the right, thereby forming the square matrix N, so that $NN^\dagger = MM^\dagger = c\mathbf{1}$. But now $\det N = 0$, so it has no inverse, which makes it impossible that $c \neq 0$. It follows that $c = 0$ and hence $M = N = 0$. □

The following orthogonality theorem plays an important role in the construction of irreducible representations.

Theorem E.4 *Let $\mathcal{D}^{(1)}(R)$ and $\mathcal{D}^{(2)}(R)$ be matrices of two inequivalent irreducible representations of a group $\mathcal{G}$ of order h. Then*

$$\sum_{R\in\mathcal{G}} \mathcal{D}^{(1)}_{\mu\nu}(R)\mathcal{D}^{(2)}_{\lambda\kappa}(R^{-1}) = 0$$

for all $\mu, \nu, \lambda, \kappa$, and

$$\sum_{R\in\mathcal{G}} \mathcal{D}^{(1)}_{\mu\nu}(R)\mathcal{D}^{(1)}_{\lambda\kappa}(R^{-1}) = \delta_{\mu\kappa}\delta_{\nu\lambda}h/n_1,$$

where n_1 is the dimension of the representation $\mathcal{D}^{(1)}$.

Of course, if the representation is unitary, then $\mathcal{D}_{\lambda\kappa}(R^{-1}) = (\mathcal{D}^{-1})_{\lambda\kappa}(R) = \mathcal{D}^*_{\kappa\lambda}(R)$.

Proof: Define the following matrix:

$$M \stackrel{\text{def}}{=} \sum_{R\in\mathcal{G}} \mathcal{D}^{(1)}(R)X\mathcal{D}^{(2)}(R^{-1}),$$

where X is an arbitrary $n_1 \times n_2$-matrix. Now

$$M\mathcal{D}^{(2)}(S) = \sum_{R\in\mathcal{G}} \mathcal{D}^{(1)}(R)X\mathcal{D}^{(2)}(R^{-1}S) = \sum_{T\in\mathcal{G}} \mathcal{D}^{(1)}(ST)X\mathcal{D}^{(2)}(T^{-1}) = \mathcal{D}^{(1)}(S)M.$$

Therefore by Schur's lemma: if $n_1 \neq n_2$, then $M = 0$; choosing $X_{ij} = \delta_{i\nu}\delta_{j\lambda}$ now yields the first equation of the theorem. If $n_1 = n_2$, then, since the two representations are assumed to be inequivalent, again $M = 0$, and we get the first equation. If we take the two representations to be equal, as in the second equation, then, by Lemma (E.2), M must be a multiple of the unit matrix, and the same choice of X as before yields

$$M_{\mu\kappa} = \sum_R \mathcal{D}^{(1)}_{\mu\nu}(R)\mathcal{D}^{(1)}_{\lambda\kappa}(R^{-1}) = c(\nu,\lambda)\delta_{\mu\kappa},$$

the trace of which gives

$$c(\nu,\lambda)n_1 = \sum_\mu M_{\mu\mu} = \sum_{R\mu} \mathcal{D}^{(1)}_{\mu\nu}(R)\mathcal{D}^{(1)}_{\lambda\mu}(R^{-1}) = \sum_R \mathcal{D}^{(1)}_{\lambda\nu}(E) = h\delta_{\lambda\nu}.$$

□

E.2.1 Character

The *character* of a group element in a given representation is the *trace* of the matrix representing it: $\chi(R) \stackrel{\text{def}}{=} \text{tr}\mathcal{D}(R)$. By the character of a representation we mean the collection of the numbers that are the traces of all the matrices. Note that the character is a *class function,* since $\chi(ABA^{-1}) = \text{tr}\mathcal{D}(ABA^{-1}) = \text{tr}\mathcal{D}(A)\mathcal{D}(B)D^{-1}(A) = \text{tr}\mathcal{D}(B) = \chi(B)$. For similar reasons, equivalent representations have the same characters. The character of an irreducible representation is called *primitive.*

Taking traces of the two formulas in Theorem E.4 leads to the orthogonality theorem for primitive characters:

Theorem E.5

$$\sum_C \sqrt{N_C/h}\,\chi^{(i)}(C)\,\sqrt{N_C/h}\,\chi^{(j)*}(C) = \delta_{ij},$$

where the sum runs over the classes of the group, N_C is the number of elements in the class C, $\chi^{(i)}$ and $\chi^{(j)}$ denote the characters of irreducible representations i and j, respectively, and h is the order of the group.

(This theorem implies, incidentally, that for all but the trivial representation, $\sum_R \chi(R) = 0$.[6]) We may therefore form an h-dimensional complex linear vector space, in which the different primitive characters form mutually orthogonal vectors; in fact, the dimensionality of the space need only be equal to the number of classes in the group. Hence it follows immediately that the number of irreducible representations of a given group $\mathcal{G}$ cannot be larger than the number of classes in $\mathcal{G}$. Indeed, the following, stronger proposition holds. (Its proof will be given a little later, when we have all the needed tools.)

Theorem E.6 *The number of inequivalent irreducible representations of a group $\mathcal{G}$ is equal to the number of classes in $\mathcal{G}$.*

Since this theorem, together with Theorem E.5, implies that the primitive characters form an orthogonal basis in a linear vector space whose dimensionality equals the number of classes in $\mathcal{G}$, it is equivalent to the statement that

$$\sum_i \chi^{(i)}(C)\chi^{(i)*}(C') = \delta_{CC'}h/N_C, \tag{E.1}$$

(where the sum runs over all the inequivalent irreducible representations, that is, over all the primitive characters) which says they form a *complete set.* Indeed, Theorem E.5 and Eq. (E.1) together are equivalent to Theorem E.6.

---

[6] **Why?**

Theorem E.5 allows us to determine how many times each irreducible representation $\Gamma^{(i)}$ of a group $\mathcal{G}$ is contained in a given representation Γ, which is written in the form

$$\Gamma = k_1\Gamma^{(1)} \oplus k_2\Gamma^{(2)} \oplus \ldots = \bigoplus_i k_i\Gamma^{(i)},$$

meaning that when Γ is reduced to block form, the block corresponding to the irreducible representation $\Gamma^{(i)}$ appears k_i times. Taking the trace of this implies $\chi(R) = \sum_i k_i\chi^{(i)}(R)$, multiplication of which by $\chi^{(j)*}(R)$ and summing over R then yields, according to Theorem E.5,

$$k_j = \frac{1}{h}\sum_C N_C\chi(C)\chi^{(j)*}(C). \tag{E.2}$$

It follows from this that the number of times an individual irreducible representation is contained in a given representation is independent of the way the reduction was performed; it is unique. The theorem also implies that a necessary and sufficient condition for two irreducible representations to be equivalent is that their characters are the same. (Clearly, if they are equivalent, their characters are the same; if they are inequivalent, the theorem says that their characters are *orthogonal*.) The same then follows for all representations, since their content of irreducible ones is unique.

Suppose now that $\Gamma = \bigoplus_i k_i\Gamma^{(i)}$, so that $\chi(R) = \sum_i k_i\chi^{(i)}(R)$, and from the orthogonality theorem,

$$\sum_{R\in\mathcal{G}} |\chi(R)|^2 = \sum_C N_C|\chi(C)|^2 = \sum_i k_i^2 \sum_C N_C|\chi^{(i)}(C)|^2 = h\sum_i k_i^2, \tag{E.3}$$

which immediately implies that *if for a given representation* $\sum_R |\chi(R)|^2 > h$, *then it must be reducible.*

The **regular representation** is another useful tool. It is defined by the following matrices:[7]

$$\mathcal{D}_{\mu\nu}(R_i) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if} \quad R_\mu^{-1}R_\nu = R_i \\ 0 & \text{otherwise.} \end{cases} \tag{E.4}$$

What this amounts to is writing down the $h \times h$ multiplication table for the group, but with the inverses of the elements in the left column. For the

[7] **Prove that this defines a representation.**

group S_3 this looks as follows:

	E	A	B	C	D	F
E^{-1}	E	A	B	C	D	F
A^{-1}	A	E	D	F	B	C
B^{-1}	B	F	E	D	C	A
C^{-1}	C	D	F	E	A	B
D^{-1}	F	B	C	A	E	D
F^{-1}	D	C	A	B	F	E

The 6×6 matrix representing A then has the entry 1 wherever A appears and zeros everywhere else, etc.

How many times does the regular representation contain the i^{th} irreducible representation? We saw that $hk_i = \sum_R \chi^{\rm reg}(R)\chi^{(i)*}(R)$; but for the regular representation $\chi^{\rm reg}(R) = 0$ unless $R = E$, in which case $\chi^{\rm reg}(E) = h$ and $\chi^{(i)}(E) = n_i$. Therefore $k_i = n_i$: *the regular representation contains each irreducible representation* $\Gamma^{(i)}$ *as many times as the dimensionality of* $\Gamma^{(i)}$. Since the left-hand side of the equation $\chi^{\rm reg}(C) = \sum_i n_i \chi^{(i)}(C)$ vanishes unless $C = E$, in which case it equals h, we have the general formula:

$$\sum_i n_i \chi^{(i)}(C) = \begin{cases} 0 & \text{if} \quad C \neq E \\ h & \text{if} \quad C = E \end{cases} . \tag{E.5}$$

Using (E.3) for the regular representation leads to the result that $h^2 = h\sum_i k_i^2 = h\sum_i n_i^2$, which implies another general formula,

$$\sum_i n_i^2 = h. \tag{E.6}$$

We are now ready to prove Theorem E.6.

Proof of Theorem E.6: In any given irreducible representation of a given group $\mathcal{G}$ with the matrices $\mathcal{D}(R)$, form the following class function:

$$\Omega(C) \stackrel{\rm def}{=} \sum_{R \in C} \mathcal{D}(R), \tag{E.7}$$

C being a class in $\mathcal{G}$. Multiplying two of them, we obtain

$$\Omega(C)\Omega(C') = \sum_{R \in C} \sum_{S \in C'} \mathcal{D}(R)\mathcal{D}(S) = \sum_{R \in C} \sum_{S \in C'} \mathcal{D}(RS), \tag{E.8}$$

where RS runs through whole classes as R and S do, with the result that

$$\Omega(C)\Omega(C') = \sum_{C''} a_{CC'C''}\Omega(C''), \tag{E.9}$$

in which the coefficients $a_{CC'C''}$ are non-negative integers that simply count the number of times the same class C'' appears in the sum.

Note that for any $A \in \mathcal{G}$

$$\Omega(C)\mathcal{D}(A) = \sum_{R \in C} \mathcal{D}(R)\mathcal{D}(A) = \mathcal{D}(A) \sum_{R \in C} \mathcal{D}(A^{-1}RA) = \mathcal{D}(A) \sum_{R \in C} \mathcal{D}(R) = \mathcal{D}(A)\Omega(C).$$

Since $\Omega(C)$ commutes with all the matrices in the group, we conclude by Lemma E.2 that

$$\Omega(C) = \alpha_C \mathbf{1},$$

the trace of which yields $\mathrm{tr}\Omega(C) = N_C\chi(C) = \alpha_C n$, where n is the dimension of the representation and N_C is the number of elements in the class C; thus, $\alpha_C = N_C\chi(C)/n$,

$$\Omega(C) = \frac{N_C}{n}\chi(C)\mathbf{1},$$

so that according to (E.9)

$$\Omega(C)\Omega(C') = \frac{N_C N_{C'}}{n^2}\chi(C)\chi(C')\mathbf{1} = \sum_{C''} a_{CC'C''}\frac{N_{C''}}{n}\chi(C'')\mathbf{1},$$

or, for the i^{th} irreducible representation,

$$\sum_{C''} n_i \frac{a_{CC'C''}N_{C''}}{N_C N_{C'}}\chi^{(i)}(C'') = \chi^{(i)}(C)\chi^{(i)}(C').$$

Summing this formula over i, that is, over all the irreducible representations, and using (E.5), we obtain

$$\sum_i \chi^{(i)}(C)\chi^{(i)}(C') = \sum_{C''}\frac{a_{CC'C''}N_{C''}}{N_C N_{C'}}\sum_i n_i\chi^{(i)}(C'') = \frac{a_{CC'E}N_{C''}}{N_C N_{C'}}h.$$

Now, if A and B are in the same class C, then A^{-1} and B^{-1} are also in the same class, which we'll call C^{-1}; so we have

$$\Omega(C)\Omega(C^{-1}) = a_{CC^{-1}E}\Omega(E) + \ldots.$$

But you can get E only if $R = S^{-1}$ in (E.8), and that happens N_C times; therefore

$$a_{CC^{-1}E} = N_C = N_{C^{-1}}.$$

On the other hand, if $C' \neq C^{-1}$, then one will never get E, so that

$$a_{CC'E} = 0 \qquad \text{for} \qquad C' \neq C^{-1}.$$

The fact that $\chi(C^{-1}) = \chi^*(C)$ then leads to (E.1), and therefore proves the theorem.□

E.2.2 Real representations

Here we want to study conditions under which a given irreducible representation Γ of a group $\mathcal{G}$, if it is not already real, is equivalent to a real representation, i.e., when there exists a canonical transformation that makes all the matrices in Γ real. A look at the character of Γ allows us to make an immediate distinction: if all the characters are real, then $\Gamma \approx \Gamma^*$; if they are not, $\Gamma \not\approx \Gamma^*$. Clearly then, if Γ is equivalent to a real representation, its character has to be real. We can therefore distinguish between three classes of representations:

1. The characters are all real, so that $\Gamma \approx \Gamma^*$; in that case, either

   a) Γ is equivalent to a real representation, in which event the representation is called *of type 1* (or "integer"), or else

   b) Γ is not equivalent to a real representation, in which event it is designated *of type 2* (or "half-integer").

2. The characters are *not* all real (so that $\Gamma \not\approx \Gamma^*$), in which case the representation is called *of type 3.*

Lemma E.7 *A given irreducible representation* Γ *is of type 1 if and only if there exists a symmetric matrix* X *such that the matrix* K *defined by*

$$K \stackrel{\text{def}}{=} \sum_{R\in\mathcal{G}} \widetilde{\mathcal{D}}(R) X \mathcal{D}(R) \tag{E.10}$$

is $\neq 0$*; it is of type 2 if and only if there exists an anti-symmetric* X *such that* $K \neq 0$*; it is of type 3, if for all* X*,* $K = 0$*. If a representation is of odd dimension and all its characters are real, it is necessarily of type 1.*

Proof: Assume that Γ is irreducible and unitary, and suppose we have case (1), so that $\mathcal{D} \approx \mathcal{D}^* = \widetilde{\mathcal{D}}^{-1}$. There then exists a unitary matrix M such that for all $R \in \mathcal{G}$, $M\mathcal{D}(R)M^{-1} = \mathcal{D}^*(R) = \widetilde{\mathcal{D}}^{-1}(R)$, or

$$\widetilde{\mathcal{D}}(R) M \mathcal{D}(R) = M.$$

The transpose of this equation reads

$$\widetilde{\mathcal{D}}(R) \widetilde{M} \mathcal{D}(R) = \widetilde{M},$$

and its inverse reads

$$\mathcal{D}^{-1}(R) M^{-1} \widetilde{\mathcal{D}}^{-1}(R) = M^{-1}.$$

Multiplying these two equations leads to $M^{-1}\widetilde{M} = \mathcal{D}^{-1}(R)M^{-1}\widetilde{M}\mathcal{D}(R)$, or

$$\mathcal{D}(R) M^{-1} \widetilde{M} = M^{-1} \widetilde{M} \mathcal{D}(R).$$

Because Γ is irreducible, Schur's lemma implies that $M^{-1}\widetilde{M}$ must be a multiple of the unit matrix, or

$$\widetilde{M} = cM,$$

whose transpose is $M = c\widetilde{M} = c^2 M$, and therefore $c^2 = 1$, and M must be either symmetric or anti-symmetric. Since $\det \widetilde{M} = \det M$ and $\det(-M) = (-1)^n \det M$ if the dimensionality of M is n, $c = -1$ *can occur only for a representation of even dimension;* if the dimension of Γ is odd, M must be symmetric.

Suppose now that $c = +1$, so that M is symmetric and unitary; in that case the representation Γ *is of type 1.* To prove this, we define the unitary, symmetric matrix N such that $N^2 = M$,[8] and we have $N^2\mathcal{D}(R)N^{-2} = \mathcal{D}^*(R)$, which implies

$$\mathcal{D}'(R) \stackrel{\text{def}}{=} N\mathcal{D}(R)N^{-1} = N^{-1}\mathcal{D}^*(R)N = N^*\mathcal{D}^*(R)N^{*-1} = (N\mathcal{D}(R)N^{-1})^* = \mathcal{D}'^*(R).$$

The representation $\Gamma' \approx \Gamma$ is thus real and the representation is of type 1. Moreover, because $\mathcal{D}$ and N are both unitary, so is $\mathcal{D}'$; therefore $\mathcal{D}'$ is *real, orthogonal.*

For any given irreducible, unitary representation Γ, next define a matrix K by (E.10), where X is an arbitrary square matrix. We then have

$$\mathcal{D}^*(R)K = \widetilde{\mathcal{D}}(R^{-1})K = \sum_{S\in\mathcal{G}} \widetilde{\mathcal{D}}(SR^{-1})X\mathcal{D}(S) = \sum_{T\in\mathcal{G}} \widetilde{\mathcal{D}}(T)X\mathcal{D}(TR) = K\mathcal{D}(R)$$

for all R in the group. In case (2), i.e., if the representation is of type 3, it follows by Schur's lemma that $K = 0$, no matter what X is, whereas in case (1), i.e, if the

---

[8] **Prove that such an** N **always exists.**

representation is of type 1 or 2, K is invertible and connects $\mathcal{D}^*$ with $\mathcal{D}$ in the sense that

$$\mathcal{D}^*(R) = K\mathcal{D}(R)K^{-1}.$$

Similarly, we find $\mathcal{D}^*(R)\widetilde{K} = \widetilde{K}\mathcal{D}(R)$, and hence $\mathcal{D}(R)K^\dagger = K^\dagger\mathcal{D}^*(R)$, from which it follows that

$$K^\dagger K\mathcal{D}(R) = \mathcal{D}(R)K^\dagger K.$$

In case (1) we can therefore conclude from Schur's lemma that, unless $K = 0$, we must have $K^\dagger K = a\mathbf{1}$ for some $a > 0$. We can then multiply X by $a^{-1/2}$ so that the new K is unitary, $K^{-1} = K^\dagger$. The matrix K defined by (E.10) therefore has precisely the property assumed above for the matrix M. Thus, *if the representation is of type 3, K must vanish for all X; if for some choice of X, K fails to vanish and is symmetric, the representation is of type 1.*

Conversely, assume that the representation Γ is of type 1, so that there exists a unitary transformation T that makes it real:

$$\mathcal{D}(R) = T\mathcal{D}'(R)T^{-1}, \qquad \mathcal{D}'(R) = \mathcal{D}'^*(R).$$

In that case

$$K = \sum_{R\in\mathcal{G}} \widetilde{\mathcal{D}}(R)X\mathcal{D}(R) = \widetilde{T}^{-1}\sum_{R\in\mathcal{G}} \widetilde{\mathcal{D}}'(R)\widetilde{T}XT\mathcal{D}'(R)T^{-1},$$

and

$$K' \stackrel{\text{def}}{=} \widetilde{T}KT = \sum_{R\in\mathcal{G}} \widetilde{\mathcal{D}}'(R)Y\mathcal{D}'(R),$$

where $Y \stackrel{\text{def}}{=} \widetilde{T}XT$. The new K' has the same symmetry as K: if $\widetilde{K} = cK$, then $\widetilde{K}' = \widetilde{T}\widetilde{K}T = c\widetilde{T}KT = cK'$. Written out in detail, we have, according to Theorem E.4, if h is the order of the group and n the dimension of the representation,

$$K'_{\mu\nu} = \sum_{\alpha\beta}\sum_{R\in\mathcal{G}} \mathcal{D}'_{\alpha\mu}(R)Y_{\alpha\beta}\mathcal{D}'_{\beta\nu}(R) = \frac{h}{n}\sum_{\alpha\beta} Y_{\alpha\beta}\delta_{\alpha\beta}\delta_{\mu\nu} = \frac{h}{n}\delta_{\mu\nu}\mathrm{tr}Y,$$

which implies that K' is symmetric, and so is K. Therefore, a necessary and sufficient condition for the representation to be of type 1 is that there exist a matrix X such the K of (E.10) fails to vanish and is symmetric. Moreover, if one such X exists, every matrix X leads to a K that either vanishes or is symmetric.

The formula (E.10) shows that the symmetric part $\frac{1}{2}(K + \widetilde{K})$ of K is produced by the symmetric part of X, and its anti-symmetric part is produced by the anti-symmetric part of X. Since, for any given X, K is either symmetric or anti-symmetric, we need to test only symmetric matrices X: if and only if there exists a symmetric X that leads to a $K \neq 0$, Γ is of type 1. □

E.2.3 Kronecker products

Suppose that the matrices $\mathcal{D}^{(i)}(R)$ form the i^{th} irreducible representation of a group $\mathcal{G}$. We can then form the *Kronecker product* (also called the *direct product*) of two of these representations, $\mathcal{D}^{(i\times j)}(R) \stackrel{\text{def}}{=} \mathcal{D}^{(i)}(R) \otimes \mathcal{D}^{(j)}(R)$, defined by[9]

$$\mathcal{D}^{(i\times j)}_{\mu\nu,\lambda\kappa}(R) \stackrel{\text{def}}{=} \mathcal{D}^{(i)}_{\mu\lambda}(R)\mathcal{D}^{(j)}_{\nu\kappa}(R).$$

[9] **Prove that this is a representation of the group.**

(Here the rows are numbered by the two indices $\mu\nu$ and the columns by the two indices $\lambda\kappa$. They can, of course, be renumbered by single indices, but that would complicate the notation.) This representation will, in general, be reducible and contain the l^{th} irreducible representation k_{ijl} times, so that we have the *Clebsch-Gordan series,*

$$\Gamma^{(i\times j)} \stackrel{\text{def}}{=} \Gamma^{(i)} \otimes \Gamma^{(j)} = \bigoplus_l k_{ijl}\Gamma^{(l)}, \tag{E.11}$$

or, for the matrices,

$$\mathcal{D}^{(i\times j)}(R) = \bigoplus_l k_{ijl}\mathcal{D}^{(l)}(R).$$

The coefficients k_{ijl} are easily found by the use of Theorem E.5,

$$k_{ijl} = \frac{1}{h}\sum_{R\in\mathcal{G}} \chi^{(i)}(R)\chi^{(j)}(R)\chi^{(l)*}(R). \tag{E.12}$$

For example, the number of times the trivial representation is contained in $\Gamma^{(i\times j)}$ is

$$k_{ij1} = \frac{1}{h}\sum_{R\in\mathcal{G}} \chi^{(i)}(R)\chi^{(j)}(R),$$

which is zero unless $\Gamma^{(i)}$ and $\Gamma^{(j)*}$ are equivalent, in which case it equals 1.

Note that if $\bar{k}_{ijl}$ is defined to be the number of times $\Gamma^{(l)*}$ is contained in $\Gamma^{(i)} \otimes \Gamma^{(j)}$, i.e.,

$$\Gamma^{(i)} \otimes \Gamma^{(j)} = \bigoplus_l \bar{k}_{ijl}\Gamma^{(l)*},$$

then $\bar{k}_{ijl}$ is completely symmetric in all its indices:

$$\bar{k}_{ijl} = \frac{1}{h}\sum_{R} \chi^{(i)}(R)\chi^{(j)}(R)\chi^{(l)}(R). \tag{E.13}$$

E.2.4 Carrier spaces

Let us look at all this from the perspective of the carrier space. Suppose that the group $\mathcal{G}$, such as a group of symmetries of a physical system, is implemented by unitary operators $\{\mathbf{O}_R\}_{R\in\mathcal{G}}$ on a Hilbert space $\mathfrak{H}$, so that $\mathbf{O}_R\mathbf{O}_S = \mathbf{O}_{RS}$, and the unitary matrices $\mathcal{D}^{(i)}(R)$ form the i^{th} irreducible representation of $\mathcal{G}$. Define a set of operators

$$\mathbf{P}^{(i)}_{\mu\nu} \stackrel{\text{def}}{=} \frac{n_i}{h}\sum_{R\in\mathcal{G}} \mathcal{D}^{(i)*}_{\mu\nu}(R)\mathbf{O}_R, \tag{E.14}$$

and you find,[10] by the use of Theorem E.4, that

$$\mathbf{P}^{(i)}_{\mu\nu}\mathbf{P}^{(j)}_{\lambda\kappa} = \delta_{ij}\delta_{\lambda\nu}\mathbf{P}^{(i)}_{\mu\kappa}. \tag{E.15}$$

The operators

$$\mathbf{P}^{(i)}_{\mu} \stackrel{\text{def}}{=} \mathbf{P}^{(i)}_{\mu\mu} = \frac{n_i}{h}\sum_{R\in\mathcal{G}} \mathcal{D}^{(i)*}_{\mu\mu}(R)\mathbf{O}_R \tag{E.16}$$

then have the property

$$\mathbf{P}^{(i)}_{\mu}\mathbf{P}^{(j)}_{\nu} = \delta_{ij}\delta_{\mu\nu}\mathbf{P}^{(i)}_{\mu}. \tag{E.17}$$

Furthermore,

$$\mathbf{P}^{(i)\dagger}_{\mu\nu} = \frac{n_i}{h}\sum_{R\in\mathcal{G}} \mathcal{D}^{(i)*}_{\nu\mu}(R^{-1})\mathbf{O}_{R^{-1}} = \mathbf{P}^{(i)}_{\nu\mu},$$

as a result of which the $\mathbf{P}^{(i)}_{\mu}$ are Hermitian: they are orthogonal projections onto mutually orthogonal subspaces. We also find

$$\begin{aligned}
\mathbf{O}_S\mathbf{P}^{(i)}_{\mu\nu} &= \frac{n_i}{h}\sum_{R\in\mathcal{G}} \mathcal{D}^{(i)*}_{\mu\nu}(R)\mathbf{O}_{SR} = \frac{n_i}{h}\sum_{T\in\mathcal{G}} \mathcal{D}^{(i)*}_{\mu\nu}(S^{-1}T)\mathbf{O}_T \\
&= \frac{n_i}{h}\sum_{\lambda} \mathcal{D}^{(i)*}_{\mu\lambda}(S^{-1})\sum_{T} \mathcal{D}^{(i)*}_{\lambda\nu}(T)\mathbf{O}_T \\
&= \sum_{\lambda} \mathcal{D}^{(i)}_{\lambda\mu}(S)\mathbf{P}^{(i)}_{\lambda\nu}.
\end{aligned} \tag{E.18}$$

Now take any normalized eigenvector of $\mathbf{P}^{(i)}_{\mu}$,[11] so that

$$\mathbf{P}^{(i)}_{\mu}\psi^{(i)}_{\mu} = \psi^{(i)}_{\mu},$$

and form $n_i - 1$ partners of it by

$$\psi^{(i)}_{\nu} \stackrel{\text{def}}{=} \mathbf{P}^{(i)}_{\nu\mu}\psi^{(i)}_{\mu}.$$

The so defined orthonormal[12] set of n_i vectors $\{\psi^{(i)}_{\kappa}\}_{\kappa=1,\ldots,n_i}$ then has the property

$$\begin{aligned}
\mathbf{O}_S\psi^{(i)}_{\nu} &= \mathbf{O}_S\mathbf{P}^{(i)}_{\nu\mu}\psi^{(i)}_{\mu} = \sum_{\lambda} \mathcal{D}^{(i)}_{\lambda\nu}(S)\mathbf{P}^{(i)}_{\lambda\mu}\psi^{(i)}_{\mu} \\
&= \sum_{\lambda} \mathcal{D}^{(i)}_{\lambda\nu}(S)\psi^{(i)}_{\lambda}.
\end{aligned}$$

---

[10] **Prove this.**

[11] There generally is more than one; that is, the range of the projection $\mathbf{P}^{(i)}_{\mu}$ is generally not one-dimensional.

[12] **Prove that they are orthonormal.**

Hence under the group $\mathcal{G}$, expressed in terms of the operators $\{\mathbf{O}_R\}_{R\in\mathcal{G}}$, the $\psi_\nu^{(i)}$ transform among themselves: they form an invariant subspace, which is the carrier space of the i^{th} irreducible representation. One says that $\psi_\nu^{(i)}$ *belongs to the* ν^{th} *row of the* i^{th} *irreducible representation.* Each of the partners $\psi_\nu^{(i)}$ of the original $\psi_\mu^{(i)}$ has the property $\mathbf{P}_\nu^{(i)}\psi_\nu^{(i)} = \psi_\nu^{(i)}$,[13] so it does not matter which of these vectors you start with. Furthermore, if $\mathbf{P}_\nu^{(i)}\psi_\nu^{(i)} = \psi_\nu^{(i)}$, then

$$\psi_\nu^{(i)} = \mathbf{P}_\nu^{(i)}\psi_\nu^{(i)} = \mathbf{P}_{\nu\lambda}^{(i)}\mathbf{P}_{\lambda\nu}^{(i)}\psi_\nu^{(i)} = \mathbf{P}_{\nu\lambda}^{(i)}\psi_\lambda^{(i)}.$$

Hence $\mathbf{P}_\nu^{(i)}\psi_\nu^{(i)} = \psi_\nu^{(i)}$ is a necessary and sufficient condition for $\psi_\nu^{(i)}$ to have $n_i - 1$ partners such that $\psi_\lambda^{(i)} = \mathbf{P}_{\lambda\nu}^{(i)}\psi_\nu^{(i)}$, and they belong to the λ^{th} row of the i^{th} irreducible representation. Define $\mathcal{S}^{(i,m)}$ to be the n_i-dimensional invariant subspace spanned by all the partners of one of the $\psi_\mu^{(i)}$. (The label m refers to the fact that the range of the projection $\mathbf{P}_\mu^{(i)}$ may not be one-dimensional, and there are other, possibly infinitely many, eigenvectors of $\mathbf{P}_\mu^{(i)}$ with which we could have started the procedure of this paragraph.)

Note that whenever a set of vectors in $\mathfrak{H}$ transform among themselves, so that

$$\mathbf{O}_R\psi_\nu = \sum_\mu \mathcal{D}_{\mu\nu}(R)\psi_\mu,$$

the matrices $\mathcal{D}_{\mu\nu}(R)$ form a representation of the group $\mathcal{G}$.[14]

Finally, we define the projection operators

$$\mathsf{P}^{(i)} \stackrel{\text{def}}{=} \sum_\mu \mathbf{P}_\mu^{(i)} = \frac{n_i}{h}\sum_{R\in\mathcal{G}} \chi^{(i)*}(R)\mathbf{O}_R = \frac{n_i}{h}\sum_C \chi^{(i)*}(C)\mathcal{K}(C), \qquad \text{(E.19)}$$

where $\mathcal{K}(C) \stackrel{\text{def}}{=} \sum_{R\in C}\mathbf{O}_R$ is the analogue on $\mathfrak{H}$ of the matrix $\Omega(C)$ defined in (E.7). The operator $\mathsf{P}^{(i)}$ is idempotent, $\mathsf{P}^{(i)2} = \mathsf{P}^{(i)}$,[15] and such that if $\psi^{(i)} = \sum_\nu a_\nu\psi_\nu^{(i)}$, then $\mathsf{P}^{(i)}\psi^{(i)} = \psi^{(i)}$.[16] So any vector in an invariant subspace $\mathcal{S}^{(i,m)}$ belonging to the i^{th} irreducible representation is an eigenvector of $\mathsf{P}^{(i)}$. One says that $\psi^{(i)}$ *belongs to the* i^{th} *irreducible representation.* The space $\mathcal{S}^{(i)}$ on which $\mathsf{P}^{(i)}$ projects is the direct sum $\mathcal{S}^{(i)} = \bigoplus_m \mathcal{S}^{(i,m)}$.

Lemma E.8 *The mutually orthogonal*[17] *projection operators* $\mathsf{P}^{(i)}$ *add up to the unit operator; i.e., their eigenvectors span the entire Hilbert space* $\mathfrak{H}$.

---

[13] **Prove this.**
[14] **Prove this.**
[15] **Prove this.**
[16] **Prove this.**
[17] **Prove that they are mutually orthogonal.**

Proof: We have

$$\sum_i \mathsf{P}^{(i)} = \frac{1}{h}\sum_{R\in\mathcal{G}} \mathbf{O}_R \sum_i n_i \chi^{(i)*}(R) = \frac{1}{h}\sum_R \mathbf{O}_R\, \chi^{\mathrm{reg}*}(R),$$

where χ^{reg} is the character of the *regular representation*. But $\chi^{\mathrm{reg}}(R) = 0$ unless $R = E$, in which case $\chi^{\mathrm{reg}}(E) = h$. Therefore

$$\sum_i \mathsf{P}^{(i)} = \mathbf{O}_E = \mathbf{1},$$

where $\mathbf{1}$ is the unit operator on $\mathfrak{H}$. □

This lemma implies that $\mathfrak{H}$ can be decomposed into a direct sum of mutually orthogonal subspaces $\mathcal{S}^{(i)}$, each belonging to one of the irreducible representations of $\mathcal{G}$, and each $\mathcal{S}^{(i)}$ in turn can be decomposed into a direct sum of (generally infinitely many) invariant subspaces $\mathcal{S}^{(i,m)}$ of dimension n_i (the dimension of the i^{th} irreducible representation) such that $\mathcal{S}^{(i,m)}$ is spanned by n_i mutually orthogonal vectors that are all partners of one another, and every vector $\Psi \in \mathfrak{H}$ can be written as a superposition of the form $\Psi = \sum_{im\mu} a_{im\mu}|i,m,\mu\rangle$, where $|i,m,\mu\rangle \in \mathcal{S}^{(i,m)}$ and $|i,m,\mu\rangle$ belongs to the μ^{th} row of the i^{th} irreducible representation, all the vectors $|i,m,\mu\rangle$, $\mu = 1,\ldots,n_i$, with fixed i and m, being partners of one another.

Suppose now that the Hermitian operator $\mathbf{H}$ on $\mathfrak{H}$ is invariant under the group $\mathcal{G}$, i.e., it commutes with all the operators $\mathbf{O}_R$, $R \in \mathcal{G}$, and hence also with all the projections $\mathbf{P}^{(i)}_\mu$. There then exists a complete orthonormal set of eigenfunctions or quasi-eigenfunctions $|E,i,m,\mu\rangle$ labeled by the eigenvalues or quasi-eigenvalues E of $\mathbf{H}$ and by the row μ of the i^{th} irreducible representation of $\mathcal{G}$. What is more, if $\mathbf{H}|E,i,m,\mu\rangle = E|E,i,m,\mu\rangle$, then also

$$\begin{aligned} E|E,i,m,\nu\rangle &= E\mathbf{P}^{(i)}_{\nu\mu}|E,i,m,\mu\rangle = \mathbf{P}^{(i)}_{\nu\mu}\mathbf{H}|E,i,m,\mu\rangle = \mathbf{H}\mathbf{P}^{(i)}_{\nu\mu}|E,i,m,\mu\rangle \\ &= \mathbf{H}|E,i,m,\nu\rangle, \end{aligned}$$

so that all the partners of a row of an irreducible representation have the same eigenvalue or quasi-eigenvalue E and thus E must have the degeneracy of the dimension of that representation. Moreover, according to (E.15),

$$\begin{aligned} \langle E,i,m,\nu|\mathbf{H}|E,j,n,\mu\rangle &= \langle E,i,m,\nu|\mathbf{P}^{(i)}_\nu \mathbf{H}\mathbf{P}^{(j)}_\mu|E,j,n,\mu\rangle \\ &= \langle E,i,m,\nu|\mathbf{H}\mathbf{P}^{(i)}_\nu \mathbf{P}^{(j)}_\mu|E,j,n,\mu\rangle = 0 \end{aligned}$$

if $i \neq j$. As a result, and in view of Lemma E.8, we have

Theorem E.9 *Let the Hermitian operator* $\mathbf{H}$ *on a Hilbert space* $\mathfrak{H}$ *be invariant under a group* $\mathcal{G}$ *of transformations whose irreducible representations have dimensions* $n_1, n_2, \ldots$; *then* $\mathfrak{H}$ *is spanned by a set of orthonormal vectors* $|E,i,m,\mu\rangle$, *such that* $|E,i,m,\mu\rangle$ *belongs to the* μ^{th} *row of the* i^{th} *irreducible representation of* $\mathcal{G}$, E *is an eigenvalue of* $\mathbf{H}$, $\mathbf{H}|E,i,m,\mu\rangle =$

$E|E,i,m,\mu\rangle$, and for each fixed i, the values of E are equal for all $1 \leq \mu \leq n_i$; thus, the eigenvalue is n_i-fold degenerate. Furthermore, all matrix elements of $\mathbf{H}$ connecting two different irreducible representations vanish. (If E is a quasi-eigenvalue, the wording of this statement has to be changed accordingly.)

If there is no further degeneracy, i.e., if any two so-labeled eigenfunction $|E,i,m,\mu\rangle$ and $|E',j,n,\nu\rangle$ with $i \neq j$ have $E \neq E'$, then the degeneracy is called *normal;* any additional degeneracy is called *accidental.*

E.2.5 Clebsch-Gordan coefficients

If the normalized vectors belonging to the i^{th} irreducible representation of a group $\mathcal{G}$ are $\{\psi_\mu^{(i)}\}_{\mu=1,\ldots,n_i}$ and we wish to decompose the vectors $\psi_\mu^{(i)} \otimes \psi_\nu^{(j)}$ in the carrier space of the Kronecker product of the i^{th} and the j^{th} irreducible representation into a sum of vectors that belong to individual rows of irreducible representations, the l^{th} representation will appear k_{ijl} times, and we have

$$\psi_\sigma^{(l\kappa)} = \sum_{\mu=1}^{n_i} \sum_{\nu=1}^{n_j} \langle i\,\mu, j\,\nu | l\,\kappa\,\sigma\rangle \psi_\mu^{(i)} \psi_\nu^{(j)}, \quad \kappa = 1,\ldots,k_{ijl}, \tag{E.20}$$

where the $\langle i\,\mu, j\,\nu | l\,\kappa\,\sigma\rangle$ are called *Clebsch-Gordan coefficients.* Since clearly there must be as many basis functions of one kind as of the other, it follows that

$$\sum_l k_{ijl} n_l = n_i n_j.$$

Assuming the irreducible representations to be unitary, so that the basis functions are orthonormal, the matrix of Clebsch-Gordan coefficients is unitary,

$$\sum_{l\kappa\sigma} \langle i\,\mu, j\,\nu | l\,\kappa\,\sigma\rangle \langle l\,\kappa\,\sigma | i\,\mu', j\,\nu'\rangle = \delta_{\mu\mu'}\delta_{\nu\nu'}, \tag{E.21}$$

as well as

$$\sum_{\mu\nu} \langle l\,\kappa\,\sigma | i\,\mu, j\,\nu\rangle \langle i\,\mu, j\,\nu | l'\,\kappa'\,\sigma'\rangle = \delta_{ll'}\delta_{\kappa\kappa'}\delta_{\sigma\sigma'}, \tag{E.22}$$

where

$$\langle l\,\kappa\,\sigma | i\,\mu, j\,\nu\rangle = \langle i\,\mu, j\,\nu | l\,\kappa\,\sigma\rangle^*,$$

and the relation (E.20) can be inverted to read

$$\psi_\mu^{(i)} \psi_\nu^{(j)} = \sum_{l\kappa\sigma} \langle l\,\kappa\,\sigma | i\,\mu, j\,\nu\rangle \psi_\sigma^{(l\kappa)}. \tag{E.23}$$

Since $\mathcal{D}_{\mu\nu}^{(i)}(R) = (\psi_\mu^{(i)}, \mathbf{O}_R \psi_\nu^{(i)})$, these relations lead to corresponding relations between the Kronecker products of the representation matrices,

namely,

$$\mathcal{D}^{(i)}_{\mu\lambda}(R)\mathcal{D}^{(j)}_{\nu\iota}(R) = \sum_{l\kappa\sigma\sigma'} \langle i\,\mu, j\,\nu|l\,\kappa\,\sigma'\rangle \mathcal{D}^{(l)}_{\sigma'\sigma}(R)\langle l\,\kappa\,\sigma|i\,\lambda, j\,\iota\rangle. \tag{E.24}$$

This equation implies that the Clebsch-Gordan coefficients make up the matrix that reduces the Kronecker product to block form. However, if some of the blocks are equal, so that the same irreducible representation occurs more than once, the reduction is ambiguous. We have assumed in these formulas that the various repetitions $\psi^{(i\kappa)}_\mu$ transform equally, and not under equivalent but unequal representations. The complicating issue of repetitions of the same irreducible representations becomes simplified for groups that are *simply reducible,* which many of the groups of physical interest are. Here is the definition of a *simply reducible* group: 1) for each $R \in \mathcal{G}$ there exists an $A \in \mathcal{G}$ such that $R^{-1} = ARA^{-1}$; in other words, every member of the group is in the same class as its inverse; 2) the Kronecker product representation of two irreducible representations contains no irreducible representation more than once. The first of these implies that all the primitive characters are *real.*[18]

Use of Theorem E.4 together with (E.24) leads to

$$\sum_{R\in\mathcal{G}} \mathcal{D}^{(i)}_{\mu\nu}(R)\mathcal{D}^{(j)}_{\nu\iota}(R)\mathcal{D}^{(k)*}_{\rho\alpha}(R) = \frac{h}{n_k}\sum_\kappa \langle \iota\,\mu, j\,\nu|k\,\kappa\,\rho\rangle\langle k\,\kappa\,\alpha|\iota\,\lambda, j\;\iota\rangle, \tag{E.25}$$

and for simply reducible groups this reduces to the following:

$$\sum_R D^{(i)}_{\mu\nu}(R)D^{(j)}_{\nu\iota}(R)D^{(k)*}_{\rho\alpha}(R) = \frac{h}{n_k}\langle \iota\,\mu, j\,\nu|k\,\rho\rangle\langle k\,\alpha|\iota\,\lambda, j\,\iota\rangle, \tag{E.26}$$

where the index κ now no longer appears. If the question is redefined to be: how many times does the Kronecker product contain $\Gamma^{(k)*}$, which we shall indicate by a bar over the k, and we lump the factor $\sqrt{n_k}$ into the CG-coefficient, the results are called *3j-symbols*:

$$\langle \bar{j}_3\mu_3|j_2\mu_2, j_1\mu_1\rangle n^{-1/2}_{j_3} \stackrel{\text{def}}{=} \begin{pmatrix} j_1 & j_2 & j_3 \\ \mu_1 & \mu_2 & \mu_3 \end{pmatrix}. \tag{E.27}$$

It follows that

$$\sum_R D^{(j_1)}_{\mu_1\mu_1'}(R)D^{(j_2)}_{\mu_2\mu_2'}(R)D^{(j_3)}_{\mu_3\mu_3'}(R) = h \begin{pmatrix} j_1 & j_2 & j_3 \\ \mu_1 & \mu_2 & \mu_3 \end{pmatrix}^* \begin{pmatrix} j_1 & j_2 & j_3 \\ \mu_1' & \mu_2' & \mu_3' \end{pmatrix}. \tag{E.28}$$

This formula implies that

$$\left|\begin{pmatrix} j_1 & j_2 & j_3 \\ \mu_1 & \mu_2 & \mu_3 \end{pmatrix}\right|$$

[18] **Why?**

is symmetric under the interchange of any two of its columns. Phases can always be chosen in such a way that the $3j$-symbols are even under even permutations and odd under odd permutations of its columns.

E.3 Lie Groups

Many groups of interest to physics have infinitely many elements, labeled by a set of p continuous, real parameters, as in $R(a)$, where a denotes the entire set, $a = \{a_1, \ldots, a_p\}$;[19] one usually chooses $R(0) = E$, if by 0 we mean $\{0, 0, \ldots, 0\}$. The multiplication law is given by

$$R(b)R(a) = R(c), \qquad \text{where } c = \phi(a, b), \tag{E.29}$$

so that the function ϕ takes the place of the multiplication table; because of the associativity law of the group, it must have the property

$$\phi(\phi(a, b), d) = \phi(a, \phi(b, d)). \tag{E.30}$$

Let us denote the parameter for $[R(a)]^{-1}$ by $\bar{a}$, so that $R(a)R(\bar{a}) = E$, and the equation $\phi(a, \bar{a}) = \phi(\bar{a}, a) = 0$ is solved by $\bar{a} = f(a)$. If the functions ϕ and f are *analytic*, the group is called a **Lie group;** it is called *compact* if the p-dimensional parameter space is bounded and closed.

E.3.1 Coordinate transformations

Most of the groups of physical interest are transformations of coordinates on an n-dimensional space (which need not necessarily be the real, physical space). Here are some examples:

• The *linear group* in n dimensions, called GL(n): a group of linear transformations of the form

$$x_i' = \sum_{j=1}^{n} M_{ji} x_j, \quad i = 1, \ldots, n, \quad \text{with} \quad \det M_{ij} \neq 0.$$

This is the group of $n \times n$ nonsingular, real matrices; it has n^2 parameters varying from $-\infty$ to ∞ and is thus not compact.

• The *special linear group* SL(n), which is the group of *unimodular* $n \times n$ matrices, i.e., such that $\det M = 1$. It has $n^2 - 1$ parameters.

• The *orthogonal group* O(n). These are the real coordinate transformations such that $\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i'^2$, which means that the matrices have to be orthogonal,

$$\sum_{i=1}^{n} M_{ik} M_{ij} = \delta_{kj}.$$

[19] It is assumed that these are all essential and cannot be reduced to a smaller set.

They have $n(n-1)/2$ parameters.[20] The orthogonality condition implies that $\det M = \pm 1$, but only the matrices with $\det M = 1$ are continuously connected to the unit matrix. The orthogonal, *unimodular* matrices form the orthogonal group SO(n) of pure rotations; if reflections are to be included, O(n) is a *mixed* continuous group, with an additional discrete parameter. For example, O(2) is a mixed one-parameter group (one angle), and O(3) is a mixed three-parameter group (3 Euler angles), and they include reflections.

• The *unitary group* U(n) is made up of unitary $n \times n$ matrices,

$$\sum_i U_{ki} U^*_{ji} = \delta_{kj}.$$

These are the complex coordinate transformations such that $\sum_{i=1}^n |x_i|^2 = \sum_{i=1}^n |x'_i|^2$. Each matrix has n^2 free real parameters,[21] and since each element must be ≤ 1, the parameter space is closed and U(n) is compact; it has O(n) as a subgroup.[22]

• The *special unitary group* SU(n) is a (n^2-1)-parameter subgroup of U(n) consisting of the unitary matrices with $\det U = +1$.

• The rigid motions in three-dimensional Euclidian space,

$$x'_i = \sum_{j=1}^3 T_{ji} x_j + a_i, \qquad i = 1, 2, 3,$$

where the matrices $\{T_{ji}\}$ are real orthogonal; the rigid motions form a six-parameter group, whose parameter space is not closed.

• If we add the dilations, $x'_i = a x_i$, we have the seven-parameter *similitude group.*

• The addition of inversions on spheres (see Figure E.1) to the group of rigid motions makes it into the ten-parameter *conformal group.*

E.3.2 Infinitesimal generators

For *general* coordinate transformations (not necessarily linear), writing x for the set of n coordinates and a for the set of p parameters, we have

$$x = f(x', b), \quad x' = f(x'', a), \quad x = f(f(x'', a), b) = f(x'', c), \quad f(x, 0) = x, \tag{E.31}$$

with the multiplication law

$$c = \phi(a, b), \qquad \phi(a, 0) = \phi(0, a) = a.$$

[20] **Prove this.**
[21] **Prove this.**
[22] **Why?**

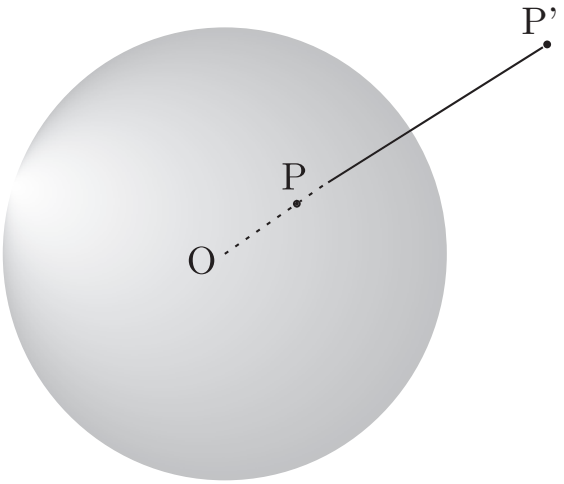

FIGURE E.1. Inversion on a sphere of radius r; the points P and P' are such that $OP \times OP' = r$.

On the assumption the group elements depend on the parameters continuously and differentiably, so that we are dealing with a Lie group, it is most convenient to look at *infinitesimal transformations.*

The transformation

$$x_i = f_i(x', a) \tag{E.32}$$

can be varied infinitesimally in two different ways:

$$x_i + dx_i = f_i(x', a + da) = f_i(x, \delta a).$$

The change da is a variation of the parameters anywhere in the group, while δa is a variation in the vicinity of $a = 0$, that is, in the neighborhood of the identity transformation. If the second way is adopted, we have

$$dx_i = \sum_k \left.\frac{\partial f_i(x,a)}{\partial a_k}\right|_{a=0} \delta a_k = \sum_k u_{ik}(x)\delta a_k, \tag{E.33}$$

where

$$u_{ik}(x) \stackrel{\text{def}}{=} \left.\frac{\partial f_i}{\partial a_k}\right|_{a=0} \qquad i = 1, \ldots, n, \;\; k = 1, \ldots, p. \tag{E.34}$$

At the same time, the multiplication law $c_i = \phi_i(a, b)$, together with $a_i = \phi_i(a, 0) = \phi_i(0, a)$, leads to

$$a_i + da_i = \phi_i(a, \delta a),$$

and therefore,

$$da_i = \sum_k \mathcal{A}_{ik}\delta a_k, \tag{E.35}$$

where

$$\mathcal{A}_{ik} \stackrel{\text{def}}{=} \left.\frac{\partial \phi_i(a,b)}{\partial b_k}\right|_{b=0}. \tag{E.36}$$

The $p \times p$ matrix $\{\mathcal{A}_{ik}\}$ must be nonsingular; this is because the set of equations $c_i = \phi_i(a,b)$ defining multiplication in the group has to be uniquely solvable for the a_i in terms of the b_k and the c_j as well as for the b_i in terms of the a_k and the c_j, which requires that the Jacobian of the system not vanish. But that Jacobian is equal to $\det \mathcal{A}$, so we must have

$$\det \mathcal{A} \neq 0, \tag{E.37}$$

and it is legitimate to define $\mathcal{B}$ as the inverse of $\mathcal{A}$: $\mathcal{B} \stackrel{\text{def}}{=} \mathcal{A}^{-1}$, i.e.,

$$\sum_l \mathcal{B}_{il}\mathcal{A}_{lk} = \delta_{ik}. \tag{E.38}$$

Therefore Eqs. (E.35) may be inverted to read

$$\delta a_i = \sum_j \mathcal{B}_{ij} da_j, \tag{E.39}$$

so that (E.33) becomes

$$dx_i = \sum_{kl} u_{ik}\mathcal{B}_{kl} da_l,$$

or

$$\frac{\partial x_i}{\partial a_l} = \sum_j u_{ij}(x)\mathcal{B}_{jl}(a), \tag{E.40}$$

which implies that

$$u_{ij} = \sum_l \frac{\partial x_i}{\partial a_l}\mathcal{A}_{lj}. \tag{E.41}$$

If the x_i are thought of as functions varying with the a_j, (E.40) may be regarded as a set of partial differential equations, whose solution is given by (E.32), with the x_i' serving as initial conditions.

The question now is, how does an arbitrarily given differentiable function $F(x)$, which depends on the parameters a_i via the x_j, change when the a_i are varied near zero? The answer is

$$dF(x) = \sum_i \frac{\partial F}{\partial x_i} dx_i = \sum_{il} \frac{\partial F}{\partial x_i} u_{il}\delta a_l = \sum_l \delta a_l \mathfrak{X}_l F, \tag{E.42}$$

where the $\mathfrak{X}_l$ are a set of differential operators,

$$\mathfrak{X}_l \stackrel{\text{def}}{=} \sum_i u_{il}(x)\frac{\partial}{\partial x_i}, \tag{E.43}$$

called the *infinitesimal generators* (or simply the *generators*) of the group. There are as many of them as there are parameters: a p-parameter group has p infinitesimal generators. With the x_i regarded as functions of the a_j, these infinitesimal generators may, according to (E.41) also be written in the form

$$\mathfrak{X}_l = \sum_{ij} \frac{\partial x_i}{\partial a_j} \mathcal{A}_{jl} \frac{\partial}{\partial x_i} = \sum_j \mathcal{A}_{jl}(a) \frac{\partial}{\partial a_j}. \tag{E.44}$$

As a simple example, take $F = x_s$. It then follows from (E.43) that $\mathfrak{X}_l x_s = u_{sl}(x)$.

We next calculate the commutation relations between the infinitesimal generators, and we find (see below)

$$[\mathfrak{X}_k, \mathfrak{X}_l] = \sum_r \mathfrak{c}_{kl}^r \mathfrak{X}_r, \tag{E.45}$$

where the $\mathfrak{c}_{kl}^r$ are given as functions of the group parameters by

$$\mathfrak{c}_{kl}^r = \sum_{ij} \left[\mathcal{A}_{ik} \frac{\partial \mathcal{A}_{jl}}{\partial a_i} - \mathcal{A}_{il} \frac{\partial \mathcal{A}_{jk}}{\partial a_i} \right] \mathcal{B}_{rj}, \tag{E.46}$$

and in terms of the x_i

$$\sum_i \left(u_{ik} \frac{\partial u_{jl}}{\partial x_i} - u_{il} \frac{\partial u_{jk}}{\partial x_i} \right) = \sum_s u_{js} \mathfrak{c}_{kl}^s. \tag{E.47}$$

From (E.46) and (E.47) we may draw the important conclusion that *the $\mathfrak{c}_{kl}^r$ are independent of the x_i and of the group parameters;* they are called the **structure constants** of the Lie group. *They are real numbers,* even though in some specific realization of the group, the x_i may be complex. Furthermore, the functions $\mathcal{B}_{ij}$ satisfy the following equation in terms of these constants, known as *Maurer's relation:*

$$\frac{\partial \mathcal{B}_{ri}}{\partial a_j} - \frac{\partial \mathcal{B}_{rj}}{\partial a_i} = \sum_{lk} \mathcal{B}_{li} \mathcal{B}_{kj} \mathfrak{c}_{lk}^r. \tag{E.48}$$

Proof of (E.45): We can proceed in two ways. First, use (E.44):

$$\begin{aligned}
[\mathfrak{X}_k, \mathfrak{X}_l] &= \sum_{ij} \left[\mathcal{A}_{ik} \frac{\partial}{\partial a_i} \mathcal{A}_{jl} \frac{\partial}{\partial a_j} - \mathcal{A}_{il} \frac{\partial}{\partial a_i} \mathcal{A}_{jk} \frac{\partial}{\partial a_j} \right] \\
&= \sum_{ij} \left[\mathcal{A}_{ik} \frac{\partial \mathcal{A}_{jl}}{\partial a_i} - \mathcal{A}_{il} \frac{\partial \mathcal{A}_{jk}}{\partial a_i} \right] \frac{\partial}{\partial a_j} \\
&= \sum_{ijs} \left[\mathcal{A}_{ik} \frac{\partial \mathcal{A}_{jl}}{\partial a_i} - \mathcal{A}_{il} \frac{\partial \mathcal{A}_{jk}}{\partial a_i} \right] \frac{\partial x_s}{\partial a_j} \frac{\partial}{\partial x_s} \\
&= \sum_{ijr} \left[\mathcal{A}_{ik} \frac{\partial \mathcal{A}_{jl}}{\partial a_i} - \mathcal{A}_{il} \frac{\partial \mathcal{A}_{jk}}{\partial a_i} \right] \mathcal{B}_{rj} \mathfrak{X}_r,
\end{aligned}$$

by the use of (E.40), which proves (E.45) and (E.46).

However, we can also proceed as follows:

$$\begin{aligned}[\mathfrak{X}_k, \mathfrak{X}_l] &= \sum_{ij} \left(u_{ik} \frac{\partial}{\partial x_i} u_{jl} \frac{\partial}{\partial x_j} - u_{il} \frac{\partial}{\partial x_i} u_{jk} \frac{\partial}{\partial x_j} \right) \\ &= \sum_{ij} \left(u_{ik} \frac{\partial u_{jl}}{\partial x_i} - u_{il} \frac{\partial u_{jk}}{\partial x_i} \right) \frac{\partial}{\partial x_j},\end{aligned}$$

from which (E.47) may be concluded because the right-hand side has to equal $\sum_s \mathfrak{c}_{kl}^s \mathfrak{X}_s$. □

Proof that the $\mathfrak{c}_{kl}^r$ are constants: Remember that (E.40) has (E.32) as a solution, with the x_i' as independent integration constants. Now, (E.46) expresses the $\mathfrak{c}_{kl}^r$ in terms of the a_i alone, so since the x_j are functions of the a_i and the independent integration constants x_j', the $\mathfrak{c}_{kl}^r$ cannot depend on the x_j.

Suppose now that we change the a_j and the x_i' together in such a way that the x_k remain fixed. Then differentiation of (E.47) with respect to a_r gives

$$\sum_s u_{js}(x) \frac{\partial \mathfrak{c}_{lk}^s}{\partial a_r} = 0$$

for all x and all j, l, and r. This implies that for all r

$$\frac{\partial \mathfrak{c}_{lk}^s}{\partial a_r} = 0,$$

because otherwise the functions $u_{js}(x)$ would be linearly dependent, which by (E.33) would imply that the number of parameters chosen is not minimal. Therefore, contrary to the appearance of (E.46), the $\mathfrak{c}_{kl}^r$ must be independent of the group parameters as well as of the x_j. □

Proof of (E.48): From the fact that $\mathcal{A}$ and $\mathcal{B}$ are inverses of one another it follows that

$$\begin{aligned}\mathfrak{c}_{kl}^r &= \sum_{ij} \left[\mathcal{A}_{ik} \frac{\partial \mathcal{A}_{jl}}{\partial a_i} - \mathcal{A}_{il} \frac{\partial \mathcal{A}_{jk}}{\partial a_i} \right] \mathcal{B}_{rj} \\ &= \sum_{ij} \left(-\mathcal{A}_{ik} \mathcal{A}_{jl} \frac{\partial \mathcal{B}_{rj}}{\partial a_i} + \mathcal{A}_{il} \mathcal{A}_{jk} \frac{\partial \mathcal{B}_{rj}}{\partial a_i} \right) \\ &= \sum_{ij} \left(\mathcal{A}_{il} \mathcal{A}_{jk} - \mathcal{A}_{ik} \mathcal{A}_{jl} \right) \frac{\partial \mathcal{B}_{rj}}{\partial a_i} \\ &= \sum_{ij} \mathcal{A}_{il} \mathcal{A}_{jk} \left(\frac{\partial \mathcal{B}_{rj}}{\partial a_i} - \frac{\partial \mathcal{B}_{ri}}{\partial a_j} \right),\end{aligned}$$

which implies (E.48). □

The structure constants are obviously anti-symmetric in their lower indices,

$$\mathfrak{c}_{kl}^r = -\mathfrak{c}_{lk}^r, \tag{E.49}$$

and it follows from the Jacobi identity for commutators,[23]

$$[[\mathsf{A}, \mathsf{B}], \mathsf{C}] + [[\mathsf{B}, \mathsf{C}], \mathsf{A}] + [[\mathsf{C}, \mathsf{A}], \mathsf{B}] = 0, \tag{E.50}$$

[23] **Show this as an exercise.**

that they obey the identity

$$\sum_s (\mathfrak{c}^s_{lk}\mathfrak{c}^r_{sm} + \mathfrak{c}^s_{km}\mathfrak{c}^r_{sl} + \mathfrak{c}^s_{ml}\mathfrak{c}^r_{sk}) = 0. \tag{E.51}$$

Sophus Lie proved that equations (E.49) and (E.51) are sufficient conditions for a set of $\mathcal{A}_{ij}$, and hence a set of $\mathcal{B}_{kl}$, to exist such that the differential equations (E.40) are solvable and functions f_i as in (E.32) exist that form the group. (We shall not prove this.) It is then always possible to return from infinitesimal transformations to the full Lie group of finite transformations.

E.3.3 The Lie algebra

We now form a p-dimensional linear vector space whose members are of the form $\mathsf{A} = \sum_i c_i \mathfrak{X}_i$, with *real* c_i, and make up an *algebra* by defining products of such linear combinations in terms of products of the $\mathfrak{X}_i$ by using (E.45). In this algebra *the left-hand side of (E.45) will be interpreted as the product of* $\mathfrak{X}_k$ *and* $\mathfrak{X}_l$ *rather than as the commutator.* The result is called the **Lie algebra** of the group. If the coefficients c_i are allowed to be complex, we obtain its *complex extension.*[24] Note that multiplication in a Lie algebra in general is *neither commutative nor associative:* $[\mathsf{A},\mathsf{B}] = -[\mathsf{B},\mathsf{A}]$ and in general

$$[[\mathsf{A},\mathsf{B}],\mathsf{C}] \neq [\mathsf{A},[\mathsf{B},\mathsf{C}]].$$

Examples: Take a one-parameter group defined by $x_j = f_j(x', t)$. Then $u_j(x) = \partial f_j/\partial t|_{t=0}$, $dt = \mathcal{A}(t)\delta t$, and $\mathcal{B} = 1/\mathcal{A}$, while

$$\mathfrak{X} = \sum_j u_j(x)\frac{\partial}{\partial x_j} = \mathcal{A}\frac{\partial}{\partial t},$$

and $u_j = \mathfrak{X}x_j$. So we have the differential equation

$$\frac{dx_j}{dt} = u_j\mathcal{B}(t) = \mathcal{B}(t)\mathfrak{X}x_j,$$

which is easily solved by

$$x_j(t) = e^{t'\mathfrak{X}}x_j(0), \qquad t'(t) \stackrel{\text{def}}{=} \int_0^t dt''\,\mathcal{B}(t'').$$

In terms of the *canonical parametrization* t', rather than t, we thus find that, if $x_j = e^{t''\mathfrak{X}}x''_j$ and $x''_j = e^{t'\mathfrak{X}}x'_j$, then $x_j = e^{(t'+t'')\mathfrak{X}}x'_j = e^{(t''+t')\mathfrak{X}}x'_j$. So the

[24] Two different Lie algebras may have the same complex extension. For example, two Lie algebras may differ by having several structure constants of opposite signs, but the products may be made to agree by multiplying some of the generators by i.

group is Abelian: *every one-parameter Lie group is Abelian,* and in terms of its canonical parametrization it is such that $R(t_1)R(t_2) = R(t_2)R(t_1) = R(t_1 + t_2)$, which means that $\phi(t_1, t_2) = c(t_1 + t_2)$.

For a two-parameter non-Abelian group we must have

$$[\mathfrak{X}_1, \mathfrak{X}_2] = a\mathfrak{X}_1 + b\mathfrak{X}_2.$$

It is then always possible[25] to form new generators

$$\mathfrak{X}_1' = \alpha_{11}\mathfrak{X}_1 + \alpha_{12}\mathfrak{X}_2, \quad \mathfrak{X}_2' = \alpha_{21}\mathfrak{X}_1 + \alpha_{22}\mathfrak{X}_2,$$

so that

$$[\mathfrak{X}_1', \mathfrak{X}_2'] = \mathfrak{X}_1'.$$

For a three-parameter non-Abelian group there are only the following three possibilities:

a) $[\mathfrak{X}_1, \mathfrak{X}_2] = \mathfrak{X}_1$, and all others zero. This means that $\mathfrak{X}_3$ commutes with the others and defines an invariant subgroup.

b)

$$[\mathfrak{X}_1, \mathfrak{X}_2] = \mathfrak{X}_3, \ \ [\mathfrak{X}_2, \mathfrak{X}_3] = \mathfrak{X}_1, \ \ [\mathfrak{X}_3, \mathfrak{X}_1] = \mathfrak{X}_2,$$

c)

$$[\mathfrak{X}_1, \mathfrak{X}_2] = \mathfrak{X}_3, \ \ [\mathfrak{X}_2, \mathfrak{X}_3] = -\mathfrak{X}_1, \ \ [\mathfrak{X}_3, \mathfrak{X}_1] = -\mathfrak{X}_2.$$

Case (b) is exemplified by SO(3), as is discussed in more detail further on. (Note that a simple change of $\mathfrak{X}_j \to -\mathfrak{X}_j$ for all j, changes the signs of all three commutators.) Case (c) is that of the 3-dimensional "Lorentz group," which leaves $x^2 + y^2 - z^2$ invariant.

A Lie algebra $\mathfrak{A}$ is called *simple* if it contains no ideals other than 0 and $\mathfrak{A}$ itself. (A subset $\mathfrak{B}$ of $\mathfrak{A}$ is an *ideal* of $\mathfrak{A}$ if $[X, Y] \in \mathfrak{B}$ for all $X \in \mathfrak{A}$ and all $Y \in \mathfrak{B}$.) If it contains no nonzero Abelian ideals, it is called *semisimple.* (A Lie algebra $\mathfrak{A}$ is called *Abelian* if $[X, Y] = 0$ for all $X, Y \in \mathfrak{A}$.)

E.4 Representations of Lie Groups

Just as for groups of finite order, a representation of a Lie group is a mapping of the group elements R on matrices $R \mapsto \mathcal{D}(R)$, but the $\mathcal{D}(R)$ are now allowed to be infinite-dimensional, or equivalently, bounded operators on a Hilbert space: the generally nonlinear transformations f are to be represented by *linear* operators on a finite-dimensional or infinite-dimensional linear vector space. Since each group element is uniquely labeled by the set

[25] **Prove this.**

of parameters a, so will be the representing invertible matrices or operators $\mathcal{D}(a)$, and since the multiplication law is embodied in (E.29), the same must hold for the $\mathcal{D}$s:

$$\mathcal{D}(b)\mathcal{D}(a) = \mathcal{D}(c), \qquad \text{where } c = \phi(a, b). \tag{E.52}$$

In order to ensure that as many of the results for representations of finite groups as possible also hold for Lie groups, it is necessary that wherever summations over group elements appear in the former, they can be replaced without error by integrals in the parameter space. This requires the adoption of integration measures or weight functions with special properties. If you recall the proofs of Schur's lemma or of the orthogonality theorem, the essential steps were of the following kind, in which $R = AR'$:

$$\sum_{R\in\mathcal{G}} f(R) = \sum_{R'\in A^{-1}\mathcal{G}} f(AR') = \sum_{R'\in\mathcal{G}} f(AR') = \sum_{R\in\mathcal{G}} f(AR),$$

which works because the summation is over the entire group $\mathcal{G}$, and $A^{-1}\mathcal{G} = \mathcal{G}$ for any $A \in \mathcal{G}$. Therefore, if the sums over all group elements are to be replaced by integrals

$$\int d\mu(a) \ldots = \int d^p a\, \rho(a) \ldots$$

in the parameter space and these are to be invariant under left multiplication, as above, we need

$$\int d^p h\, \rho(h)\mathcal{F}(H) = \int d^p f\, \rho(f)\mathcal{F}(CF) = \int d^p h\, \rho(h)\mathcal{F}(CH)$$

if $H = CF$. The resulting integral is called the *Hurwitz invariant integral* and the measure is called *Haar measure.* The left-invariant measure turns out to be

$$\rho_{\mathrm{l}}(a) = 1/\det[\partial\phi_i(b, a)/\partial b_j]_{b=0}, \tag{E.53}$$

while the right-invariant measure is

$$\rho_{\mathrm{r}}(a) = 1/\det[\partial\phi_i(a, b)/\partial b_j]_{b=0}. \tag{E.54}$$

Comparison with (E.36) shows that therefore

$$\rho_{\mathrm{r}} = 1/\det \mathcal{A}, \tag{E.55}$$

and it is important to note that *for compact groups, the left and right-invariant measures are equal:*

$$\rho_{\mathrm{r}} = \rho_{\mathrm{l}},$$

which we shall not prove.[26]

Proof of (E.53): Suppose that $H = BA$, $G = CB$, and $F = CBA = CH = GA$, and we denote the Jacobian by J_l; then we have

$$d^p f = J_l(c,h)d^p h = J_l(c,h)J_l(b,a)d^p a = J_l(g,a)d^p a,$$

which implies that $J_l(\phi(b,c),a) = J_l(c,\phi(a,b))J_l(b,a)$ and for $a = 0$,

$$J_l(g,0) = J_l(\phi(b,c),0) = J_l(c,b)J_l(b,0).$$

Therefore, since $d^p g = J_l(c,b)d^p b$, it follows that

$$\frac{d^p g}{J_l(g,0)} = \frac{d^p b}{J_l(b,0)}.$$

But since $g = \phi(b,c)$ for $G = CB$, the Jacobian is given by $J_l(b,c) = \det[\partial\phi_i(c,b)/\partial c_j]$ and hence

$$J_l(b,0) = \det\left[\frac{\partial\phi_i(c,b)}{\partial c_j}\right]_{c=0}.$$

Thus the left-invariant measure is given by (E.53), and similarly, the right-invariant measure is given by (E.54). □

Example. For the two-parameter group defined by the transformation $x' = (1+a_1)x + a_2$, which is not compact, we find that[27] $\rho_l = (1+a_1)^{-2}$, so that $d\mu_l(a) = da_1 da_2/(1+a_1)^2$, whereas $\rho_r = (1+a_1)^{-1}$, so that $d\mu_r(a) = da_1 da_2/(1+a_1)$. This is an example of a non-compact group for which the left and right-invariant measures are unequal.

For compact groups, all the summations over group elements that are used for groups of finite order can be replaced by integrals, and, if these integrals are over continuous functions, they necessarily converge. The order $h = \sum_{R\in\mathcal{G}}$ of the group in all the needed theorems will simply be replaced by

$$h = \int d^p a\, \rho(a) \stackrel{\text{def}}{=} \int d\mu(a),$$

which may be called the *volume* of the group. So Theorem E.1 holds for finite-dimensional matrix representations of compact groups, and it follows that if such a representation is reducible, it is fully reducible. (This statement holds also for noncompact groups that are semisimple.) Lemmas E.2 and E.3, of course, hold again. Next, look at Theorem E.4: it is valid for compact groups with the sum replaced by a Hurwitz integral. But suppose that one of the representations is infinite-dimensional; that would imply that $\int d\mu(a)\,|\mathcal{D}_{\mu\nu}(a)|^2 = 0$, for all μ and ν, which is impossible. Therefore we conclude—

Lemma E.10 *All the irreducible representations of a compact Lie group are finite-dimensional.*

---

[26] For a proof, see, e.g., [Hamermesh], pp. 316f.

[27] **Prove this as an exercise.**

The result of Theorem E.5 for compact groups reads

$$\int d\mu(a)\chi^{(i)}(a)\chi^{(j)*}(a) = \delta_{ij}h, \tag{E.56}$$

for any two primitive characters. So, if a given representation Γ is reducible to the direct sum $\Gamma = \bigoplus_i k_i\Gamma^{(i)}$, containing the i^{th} irreducible representation k_i times, then $\chi(a) = \sum_i k_i\chi^{(i)}(a)$, and the number of times the representation $\Gamma^{(i)}$ is contained in Γ is given by

$$k_i = \frac{1}{h}\int d\mu(a)\,\chi^{(i)*}(a)\chi(a);$$

as a result we have the analogue of (E.3)

$$\int d\mu(a)\,|\chi(a)|^2 = h\sum_i k_i^2. \tag{E.57}$$

The **regular representation** of a Lie group is defined as an integral operator on the parameter space (thus it is an infinite-dimensional representation):[28]

$$\int d\mu(c)\mathfrak{D}^{\text{reg}}(a;b,c)f(c) = f(\phi(a,b)), \tag{E.58}$$

which means that $\mathfrak{D}^{\text{reg}}(a;b,c)$, representing the group element $R(a)$, can be written in the form

$$\mathfrak{D}^{\text{reg}}(a;b,c) = \delta^p(c-\phi(a,b))/\rho(c),$$

where δ^p is the Dirac delta-function in the p-dimensional parameter space; its character is given by

$$\chi^{\text{reg}}(a) = h\delta^p(a).$$

To show this in a mathematically nonrigorous way, consider

$$\chi^{\text{reg}}(a) = \int d\mu(b)\,\mathfrak{D}^{\text{reg}}(a;b,b) = \int d^pb\,\delta^p(b-\phi(a,b)),$$

which vanishes unless $a=0$. But

$$\begin{aligned}\int d^pa\,\chi^{\text{reg}}(a) &= \int d^pb\,d^pa\,\delta^p\left[b-\phi(0,b)-\sum_i\left.\frac{\partial\phi}{\partial a_i}\right|_{a=0}a_i\right]\\ &= \int d^pb/\det(\partial\phi_i/\partial a_j)|_{a=0} = \int d\mu(b) = h,\end{aligned}$$

which implies that $\chi^{\text{reg}}(a) = h\delta^p(a)$.

---

[28] **Show that this is a representation.**

We therefore have

$$h\delta^p(a) = \sum_i k_i \chi^{(i)}(a),$$

where k_i is the number of times the i^{th} irreducible representation is contained in the regular representation, which implies that[29] $k_i = n_i$: just as for finite groups, *the regular representation of a compact Lie group contains its i^{th} irreducible representation $\Gamma^{(i)}$ as many times as the dimensionality of $\Gamma^{(i)}$*. Remember that, according to Lemma E.10, all the n_i are finite. But the delta-function cannot be expressed as a finite sum of functions $\chi^{(i)}$. It follows that *the number of inequivalent irreducible representations of a compact Lie group is infinite.*

The Casimir operator

Define the symmetric matrix $\{g_{ij}\}$ by

$$g_{ij} = g_{ji} \stackrel{\text{def}}{=} \sum_{sr} \mathfrak{c}_{ir}^s \mathfrak{c}_{js}^r \tag{E.59}$$

in terms of the structure constants of a given Lie algebra, and its inverse $\{g^{ik}\}$,

$$\sum_m g^{im} g_{mj} = \delta_{ij}. \tag{E.60}$$

(It is a theorem by Cartan that this inverse exists if and only if the algebra is semisimple.) The **Casimir operator** of the Lie algebra is defined by

$$\mathfrak{C} \stackrel{\text{def}}{=} -\sum_{rs} g^{rs} \mathfrak{X}_r \mathfrak{X}_s; \tag{E.61}$$

it has the important property that *it commutes with all the infinitesimal generators:*

$$\mathfrak{C}\mathfrak{X}_i = \mathfrak{X}_i \mathfrak{C}, \quad i = 1, \ldots, p. \tag{E.62}$$

Equation (E.62) is proved by calculating

$$\begin{aligned}
\mathfrak{X}_i\mathfrak{C} - \mathfrak{C}\mathfrak{X}_i &= \sum_{rs} g^{rs}[\mathfrak{X}_r\mathfrak{X}_s, \mathfrak{X}_i] \\
&= \sum_{rsl} (g^{rs}\mathfrak{c}_{si}^l \mathfrak{X}_r\mathfrak{X}_l + g^{rs}\mathfrak{c}_{ri}^l \mathfrak{X}_l\mathfrak{X}_s) = \sum_{rsl} g^{rs}\mathfrak{c}_{si}^l(\mathfrak{X}_r\mathfrak{X}_l + \mathfrak{X}_l\mathfrak{X}_r) \\
&= \sum_{rslm} g^{rs}g^{lm}\mathfrak{c}_{sim}(\mathfrak{X}_r\mathfrak{X}_l + \mathfrak{X}_l\mathfrak{X}_r) = \sum_{rslm} (g^{rs}g^{lm} + g^{ls}g^{rm})\mathfrak{c}_{sim}\mathfrak{X}_l\mathfrak{X}_r,
\end{aligned}$$

[29] **Show this.**

where

$$\mathfrak{c}_{sir} = -\mathfrak{c}_{isr} \stackrel{\text{def}}{=} \sum_l \mathfrak{c}^l_{si} g_{lr} = \sum_{lmn} \mathfrak{c}^l_{si} \mathfrak{c}^n_{lm} \mathfrak{c}^m_{rn},$$

and these constants are totally anti-symmetric in all their indices.[30] It therefore follows that

$$\sum_{sm} (g^{rs} g^{lm} + g^{ls} g^{rm}) \mathfrak{c}_{sim} = 0,$$

and hence (E.62). □

Now take a representation of the Lie algebra; this leads to a representation of $\mathfrak{C}$, and the matrix representing $\mathfrak{C}$ commutes with all the matrices in the group. Therefore, if the representation is irreducible, by Lemma E.2, the matrix representing $\mathfrak{C}$ must be a multiple of the unit matrix, i.e., *it must be simply a number,* a number that can be taken to characterize the irreducible representation. Therefore *every irreducible representation may be labeled by an eigenvalue of the Casimir operator.* For a compact, semisimple Lie algebra, it is always possible to choose a basis in which

$$g_{rs} = -\delta_{rs}$$

(which we shall not prove); in that case we have $\mathfrak{c}^s_{rl} = -\mathfrak{c}_{rls}$ and

$$\mathfrak{C} = \sum_r \mathfrak{X}_r \mathfrak{X}_r. \tag{E.63}$$

E.4.1 Multiple valuedness

If the group manifold of a Lie group $\mathcal{G}$ is not simply connected, there is the possibility that some representations of a Lie group are multi-valued functions of the parameters,[31] and there is no way to simply disconnect or disregard the redundant region of the parameters. The **universal covering group** $\mathcal{G}'$ of $\mathcal{G}$ is a group that is mapped homomorphically on $\mathcal{G}$ and whose manifold is simply connected, so that all its irreducible representations are single valued functions of the parameters.

For example, take the group $\mathcal{G}$ defined by the functions $R(\varphi) = e^{i\kappa\varphi}$, κ not an integer, for which $R(\varphi_1)R(\varphi_2) = R(\varphi_1 + \varphi_2)$, and whose manifold is the unit circle (which is doubly connected). The parameter φ cannot be restricted to $0 < \varphi < 2\pi/\kappa$ without violating the continuity of the multiplication law $\varphi_3 = \varphi_1 + \varphi_2$; hence there is no way of avoiding the multi-valuedness of functions on the group manifold, and as a result there are multi-valued representations of the group. The group $\mathcal{G}'$ formed by the points a, with $-\infty < a < \infty$, and the multiplication law defined by addition, $R'(a)R'(b) = R'(c)$ if $c = a + b$,[32] is the universal covering group

[30] **Prove this,** using (E.51).

[31] **Why is this impossible if the group manifold is simply connected?**

[32] **What is the unit element of this group?**

of $\mathcal{G}$; the mapping from $\mathcal{G}$ to $\mathcal{G}'$ via $R'(a) = R(a)$ is a homomorphism, but the irreducible representations of $\mathcal{G}'$ are single valued.

E.4.2 The group O(2)

The rotation groups in n dimensions are formed by those transformations $x_i' = \sum_{j=1}^{n} R_{ji} x_j$, $i = 1, \ldots, n$, on n-dimensional Euclidian space that leave the length invariant: $\sum_i x_i^2 = \sum_i x_i'^2$. It may also be defined as the group of real orthogonal $n \times n$ matrices.[33] We distinguish between the *pure rotation group*, which consists of those matrices M with $\det M = +1$, and the rotation-reflection group, for which $\det M = \pm 1$.[34] We now consider the specific case of two dimensions.

The elements of the pure rotation group may be taken to be the 2×2 matrices

$$R(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \tag{E.64}$$

in the sense that if the coordinate system is rotated counter-clockwise by φ, the coordinates of a point are changed from (x, y) to (x', y'), where

$$x' = x\cos\varphi + y\sin\varphi, \quad y' = -x\sin\varphi + y\cos\varphi, \tag{E.65}$$

i.e.,

$$(x', y') = (x, y)R(\varphi).$$

The matrices (E.64), of course, form their own two-dimensional representation $\mathcal{D}(\varphi)$. You find easily[35] that

$$R(\varphi)R(\varphi') = R(\varphi + \varphi') = R(\varphi')R(\varphi),$$

which implies that the group is Abelian. Therefore *its irreducible representations are all one-dimensional.* Indeed, the matrix

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$$

diagonalizes all the matrices in the representation,

$$\mathcal{D}'(\varphi) = S\mathcal{D}(\varphi)S^{-1} = \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix},$$

[33] Do not confuse the dimensionality of a representation of this group with the dimensionality of the defining matrices; the latter, of course, form a representation, but there are infinitely many others.

[34] **Do the matrices with** $\det M = -1$ **form a subgroup?**

[35] **Do it.**

which means it has been reduced to the direct sum of the two one-dimensional representations $\mathcal{D}^{(1)}(\varphi) = e^{i\varphi}$ and $\mathcal{D}^{(-1)}(\varphi) = e^{-i\varphi}$.

The functions u_i are given by $u_1 = y$ and $u_2 = -x$, so that the generator, in its two forms, is given by

$$\mathfrak{X} = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} = \frac{\partial}{\partial \varphi}. \tag{E.66}$$

Since the multiplication law is given by

$$\varphi_3 = \phi(\varphi_1, \varphi_2) = \varphi_1 + \varphi_2,$$

we find that the Haar measure is $d\mu(\varphi) = d\varphi$, the Hurwitz invariant integral is simply $\int_0^{2\pi} d\varphi \ldots$, and the "volume" of the group is $h = \int_0^{2\pi} d\varphi = 2\pi$. Now, all functions of the form $e^{im\varphi}$ are representations of the group, but unless m is chosen to be 0 or a positive or negative integer, these representations are multi-valued. (The existence of such multi-valued representations was to be expected, since the group manifold if not simply connected.) If we wish to restrict ourselves to single-valued ones, we have to choose $m = 0, \pm 1, \pm 2, \ldots,$

$$\mathcal{D}^{(m)}(\varphi) = e^{im\varphi}.$$

We then have, as is required by the orthogonality theorem E.4,

$$\int_0^{2\pi} d\varphi\, \mathcal{D}^{(m)*}(\varphi)\mathcal{D}^{(m')}(\varphi) = 2\pi\delta_{mm'},$$

and if there are any other irreducible representations, they have to be orthogonal to all of these, by Theorem E.4. The Fourier theorem, however, tells us that there are no such (square-integrable) functions. Therefore, these are *all the single-valued irreducible representations of the pure rotation group in two dimensions.* From a physics point of view, it will become clear that we are interested also in the double-valued representations, and for these m can take on half-integral values as well. The coordinates on which the diagonalized matrices $\mathcal{D}'(\varphi)$ act are given by

$$(x, y)S^{-1} = \frac{1}{\sqrt{2}}(x - iy, x + iy) \stackrel{\text{def}}{=} (X, Y),$$

so that the rotations are $X' = e^{i\varphi}X$, $Y' = e^{-i\varphi}Y$.

Consider now the *rotation-reflection group* in two dimensions; the reflections here are on a line, not through the center. (The latter are equivalent to a rotation by 180^o in two dimensions.) This is a mixed continuous group, with elements $R_+(\varphi)$ of determinant $+1$ and elements $R_-(\varphi)$ of determinant -1. This means we add the matrices

$$R_-(\varphi) = \begin{pmatrix} -\cos\varphi & \sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

to the previously defined

$$R_+(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

and the canonical transformation by S turns the matrices R_- into

$$\mathcal{D}'(\varphi,-) = S\mathcal{D}(\varphi,-)S^{-1} = \begin{pmatrix} 0 & e^{-i\varphi} \\ e^{i\varphi} & 0 \end{pmatrix}.$$

This group is not Abelian, and the defining representation is not reducible. There are two one-dimensional representations, the trivial representation $\mathcal{D}^{(0)}(\varphi,+) = \mathcal{D}^{(0)}(\varphi,-) = 1$, and

$$\mathcal{D}^{(00)}(\varphi,+) = 1, \quad \mathcal{D}^{(00)}(\varphi,-) = -1,$$

and the two-dimensional representations[36]

$$\mathcal{D}^{(m)}(\varphi,+) = \begin{pmatrix} e^{im\varphi} & 0 \\ 0 & e^{-im\varphi} \end{pmatrix},$$

$$\mathcal{D}^{(m)}(\varphi,-) = \begin{pmatrix} 0 & e^{-im\varphi} \\ e^{im\varphi} & 0 \end{pmatrix},$$

with $m = 0, \pm1, \pm2, \ldots$. The rotations with reflections are all in one class;[37] that is why they all have the same character $\chi^{(m)}(\varphi,-) = 0$.

Figure E.2 shows four examples of three-dimensional objects that have the symmetry of the two-dimensional rotation group, some with, some without reflections. Cases (a) and (c) have irreducible representations of one and two dimensions, while the irreducible representations of cases (b) and (d) are all one-dimensional.

E.4.3 The group SO(3)

The three-dimensional proper rotation group is formed by the group of real, orthogonal, unimodular matrices T. As discussed in Section 5.2, it can be parametrized by an angle of rotation $0 \le \psi \le \pi$ and two polar angles θ and φ that specify the orientation of the axis of rotation, as in (5.30).

The Haar measure for the group, in terms of the three angles ψ, θ, and φ, is given by $2(1-\cos\psi)\sin\theta\, d\psi d\theta d\varphi$, so that the Hurwitz integral is of the form

$$2\int_0^\pi d\psi(1-\cos\psi)\int_0^{2\pi} d\varphi \int_0^\pi d\theta \sin\theta \ldots, \tag{E.67}$$

and the "volume" of the group is $h = 8\pi^2$.

[36] **Prove that these are irreducible.**
[37] **Prove this.**

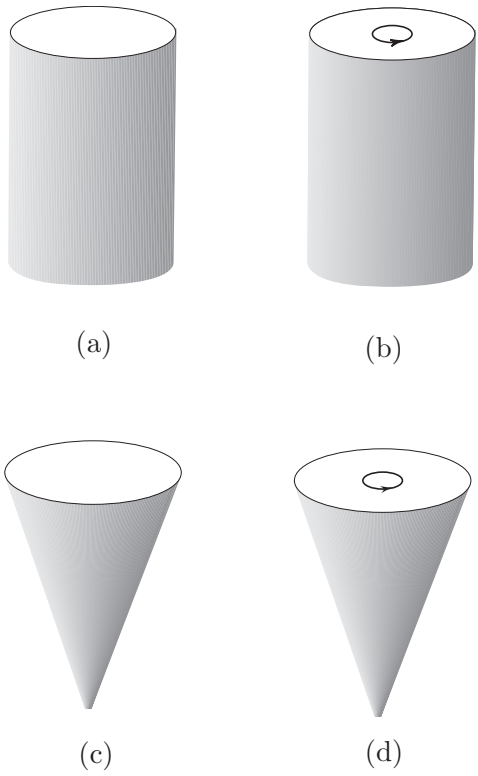

FIGURE E.2. Four three-dimensional objects with rotational symmetry about the z-axis: (a) two-dimensional rotation-reflection symmetry, plus reflections through the center; (b) two-dimensional rotation symmetry, plus reflections through the center; (c) two-dimensional rotation-reflection symmetry; (d) two-dimensional rotation symmetry only. Only object (d) has a screw sense.

To calculate the Haar measure, we begin by using the three components a_1, a_2, a_3, of the vector $\vec{a}$ defined in (5.30) as group parameters. We then have to compare the changes da_i starting from arbitrary a_i to the changes δa_i starting from zero. The matrix for an infinitesimal rotation is given by

$$T_\epsilon = \begin{pmatrix} 1 & -\delta a_3 & \delta a_2 \\ \delta a_3 & 1 & \delta a_1 \\ -\delta a_2 & \delta a_1 & 1 \end{pmatrix}$$

and an arbitrary rotation T augmented by a small increment is $T' = T_\epsilon T$. We may choose the coordinate system so that T is of the form

$$T^{(3)} = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

in which case we find that

$$T' = T + \begin{pmatrix} -\delta a_3 \sin\psi & -\delta a_3 \cos\psi & \delta a_2 \\ \delta a_3 \cos\psi & -\delta a_3 \sin\psi & -\delta a_1 \\ -\delta a_2 \cos\psi + \delta a_1 \sin\psi & \delta a_2 \sin\psi + \delta a_1 \cos\psi & 0 \end{pmatrix}.$$

The new rotation angle is obtained by taking the trace:

$$1 + 2\cos\psi' = 1 + 2\cos\psi + 2\delta a_3 \sin\psi,$$

from which we conclude that $\cos\psi' = \cos(\psi - \delta a_3)$, or $\psi' = \psi - \delta a_3$. The triple ratio

$$a_1 : a_2 : a_3 = (T_{32} - T_{23}) : (T_{13} - T_{31}) : (T_{21} - T_{12})$$

for the components of $\vec{a}$ leads to

$$a_1' = da_1 = c[\delta a_2 \sin\psi + \delta a_1(1 + \cos\psi)],$$
$$a_2' = da_2 = c[-\delta a_1 \sin\psi + \delta a_2(1 + \cos\psi)],$$
$$a_3' = a_3 + da_3 = c[2\delta a_3 \cos\psi + 2\sin\psi].$$

To find c, calculate (always keeping only linear terms in δa)

$$\psi'^2 = \sum_i a_i'^2 = c^2[4\sin^2\psi + 8\delta a_3 \sin\psi\cos\psi],$$

and from this,

$$\psi' = \psi + \delta a_3 = 2c\sin\psi(1 + \delta a_3 \cot\psi),$$

so that

$$c = \frac{\psi + \delta a_3(1 - \psi\cot\psi)}{2\sin\psi};$$

therefore, if we define $da_i = \sum_j \mathcal{A}_{ij}\delta a_j$ as in (E.35), we find that

$$\mathcal{A} = \begin{pmatrix} \psi\frac{1+\cos\psi}{2\sin\psi} & \frac{1}{2}\psi & 0 \\ -\frac{1}{2}\psi & \psi\frac{1+\cos\psi}{2\sin\psi} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

whose determinant is given by

$$\det\mathcal{A} = \frac{\psi^2}{2(1 - \cos\psi)}.$$

The weight function in the Haar measure is therefore

$$\rho = \frac{2(1 - \cos\psi)}{\psi^2}$$

and the measure is found to be $d\mu = da_1 da_2 da_3\, 2(1 - \cos\psi)/\psi^2$. Finally, the Jacobian for expressing the measure in terms of the angles ψ, θ, φ, is nothing but the familiar expression for the volume element in spherical polar coordinates, with ψ in place of r,

$$\left|\frac{\partial(a_1, a_2, a_3)}{\partial(\psi, \theta, \varphi)}\right| = \psi^2 \sin\theta,$$

and as a result we obtain (E.67).

As discussed in Section 5.2, the Cayley-Klein parameters define a homomorphism between SO(3), whose manifold is doubly connected, and its universal covering group SU(2) of unitary, unimodular 2×2 matrices S, given by (5.37) and (5.38) in terms of the Euler angles of the rotation.

To find the irreducible representations of the rotation group, we analyze the solutions of the Laplace equation (i.e., the free Schrödinger equation at zero energy)

$$\nabla^2 f(x, y, z) = 0, \tag{E.68}$$

which is invariant under rotations. Every solution of (E.68) that is a homogeneous polynomial of x, y, and z of order l can be expressed in the form

$$f = r^l \sum_{m=-l}^{l} a_m^{(l)} Y_l^m(\theta, \varphi)$$

if (E.68) is separated in spherical polar coordinates as in (4.8), where the Y_l^m are the spherical harmonics. If the coordinate system is rotated by $R_{\alpha,\beta,\gamma}$ (in terms of the Euler angles α, β, γ, so that $\theta, \varphi \mapsto \theta', \varphi'$, the homogeneous polynomial solutions of (E.68) of order l transform among themselves, and we must have

$$r^l Y_l^m(\theta', \varphi') = \mathbf{O}_{\alpha,\beta,\gamma} r^l Y_l^m(\theta, \varphi) = \sum_{m'} \mathcal{D}^{(l)}_{m'm}(\alpha, \beta, \gamma) r^l Y_l^{m'}(\theta, \varphi),$$

if $\mathbf{O}_{\alpha,\beta,\gamma}$ denotes the rotation operator. It follows that the matrices $\mathcal{D}^l_{mm'}(\alpha, \beta, \gamma)$ form a $(2l+1)$-dimensional representation of the rotation group.[38] [The same conclusion can be drawn from the fact that, according to (4.16), the Y_l^m for a given l are eigenfunctions of the rotationally invariant operator $\vec{\mathbf{L}}^2$ with the eigenvalue $l(l+1)\hbar^2$.]

Now, because the Euler angle α is a rotation of the z-axis and $Y_l^m(\theta, \varphi)$ has the form given in (D.26), we have[39]

$$\mathbf{O}_{\alpha,0,0} Y_l^m(\theta, \varphi) = e^{-im\alpha} Y_l^m(\theta, \varphi),$$

and hence

$$\mathcal{D}^{(l)}_{mm'}(\alpha, 0, 0) = e^{-im\alpha} \delta_{mm'};$$

similarly,

$$\mathcal{D}^{(l)}_{mm'}(0, 0, \gamma) = e^{-im\gamma} \delta_{mm'}.$$

We can therefore conclude that $\mathcal{D}^l_{mm'}$ must have the form

$$\mathcal{D}^{(l)}_{mm'}(\alpha, \beta, \gamma) = e^{-im\gamma} d^l_{mm'}(\beta) e^{-im'\alpha}. \tag{E.69}$$

These representations are irreducible.

---

[38] The functions $\mathcal{D}^{(l)}_{mm'}(\alpha, \beta, \gamma)$ are also the eigenfunctions of the Schrödinger equation for a symmetrical top, in which case $m\hbar$ and $m'\hbar$ are the eigenvalues of the angular-momentum projections on the body and space-fixed axes, respectively.

[39] **Where does the minus sign in $e^{-im\alpha}$ come from?**

We prove that the representation (E.69) is irreducible by showing that the only matrix M that commutes with all of the matrices $\{\mathcal{D}^{(l)}_{m'm}(\alpha,\beta,\gamma)\}$ for a fixed given value of l is a multiple of the unit matrix.

If M commutes with the matrix $\{D^l_{m',m}(\alpha,0,0)\}$ for all α, it must be diagonal. Furthermore,

$$Y_l^m(-\beta,0) = \mathbf{P}_{0,\beta,0}Y_l^m(0,0) = \sum_{m'} d^l_{m'm}(\beta)Y_l^{m'}(0,0) = i^l\sqrt{\frac{2l+1}{4\pi}}d^l_{0m}(\beta),$$

according to (D.26). Since the left-hand side of this equation does not identically vanish, $d^l_{0m}(\beta) \not\equiv 0$. Therefore, for general values of β, $0 = [M,d]_{0m} = M_{00}d_{0m} - d_{0m}M_{mm}$, and hence $M_{00} = M_{mm}$, and M is a multiple of the unit matrix. □

The $(2l+1)$-dimensional matrices $\{d^l_{mm'}(\beta)\}$ can be calculated from the fact that

$$Y_l^m(\theta-\beta,0) = \sum_{m'} d^l_{m'm}(\beta)Y_l^{m'}(\theta,0),$$

with the result that for $m' > m$

$$\begin{aligned} d^l_{mm'}(\beta) &= \sqrt{\frac{((l+m')!(l-m')!}{(l+m)!(l-m)!}}(\cos\frac{1}{2}\beta)^{m+m'} \\ &\quad \times (\sin\frac{1}{2}\beta)^{m'-m}P_{l-m'}^{(m'-m,m'+m)}(\cos\beta) \qquad \text{(E.70)} \\ &= (-1)^{m'-m}d^l_{m'm}(\beta), \end{aligned}$$

where the $P_n^{(a,b)}$ are *Jacobi polynomials*,[40] special cases of which are the Legendre polynomials, $P_n = P_n^{(0,0)}$.

The next question is, what are the primitive characters of SO(3)? Since we saw that all the elements of the group with the same total rotation angle belong to the same class, and the character $\chi(\psi)$ is a class function, we can take any convenient way of performing the rotation by ψ; take the Euler angles $(\alpha,\beta,\gamma) = (\psi,0,0)$: $\mathcal{D}^{(l)}_{mm'}(\psi,0,0) = e^{-im\psi}\delta_{mm'}$. This leads to the result

$$\begin{aligned} \chi^{(l)}(\psi) &= \sum_{m=-l}^{l} e^{-im\psi} = 1 + 2\cos\psi + 2\cos 2\psi + \ldots 2\cos l\psi \\ &= \sin[(l+\frac{1}{2})\psi]/\sin\frac{1}{2}\psi. \qquad \text{(E.71)} \end{aligned}$$

Are these representations of SO(3) we have constructed *all* the irreducible representations of the group? The answer is a qualified *yes*. If there were

[40] See [Erdélyi], vol. II, p. 168.

any others, their characters would have to be orthogonal to all the primitive characters $\chi^{(l)}$ in the sense of the Hurwitz integral (see Theorem E.5), and therefore orthogonal to all $\chi^{(l)} - \chi^{(l-1)}$:

$$\int_0^\pi d\psi\,(1-\cos\psi)\cos(l\psi)\chi(\psi) = 0, \qquad \text{for all } l.$$

By the Fourier theorem it then follows that $\chi(\psi) = 0$. Therefore, the $\{\mathcal{D}^{(l)}_{m'm}\}$ are *all* the *single-valued* irreducible representations of SO(3).

However, as we have seen, there is a homomorphism between SO(3) and SU(2), so that the representations of the rotation group form unfaithful representations of SU(2), in which the matrices representing S and $-S$ are equal; in addition, there are irreducible representations of SU(2) that form double-valued "representations" of SO(3). Let Γ be an irreducible representation of SU(2). Since the 2×2 matrix $-\mathbb{1}$ [which belongs to SU(2)] commutes with all the elements of the group, so must $\mathcal{D}(-\mathbb{1})$ with all the matrices; consequently $\mathcal{D}(-\mathbb{1})$ must be a multiple of the unit matrix. Because furthermore $\mathcal{D}^2(-\mathbb{1}) = D(\mathbb{1})$, it follows that

$$\mathcal{D}(-\mathbb{1}) = \pm\mathcal{D}(\mathbb{1});$$

and the representations of SU(2) can be divided into *odd* and *even* ones. For general S we have

$$\mathcal{D}(-S) = \mathcal{D}(-\mathbb{1})\mathcal{D}(S) = \pm\mathcal{D}(S).$$

So the even irreducible representations of SU(2) are the irreducible representations of SO(3); the others are not really representations of SO(3) but may be dubbed *quasi-representations*, for which $\mathcal{D}(R_1)\mathcal{D}(R_2) = \pm\mathcal{D}(R_1R_2)$, and the sign cannot be fixed. The matrices S themselves, of course, form an odd representation of SU(2) and hence a quasi-representation of SO(3).

The irreducible representations of SU(2) are found similarly as those of SO(3). The transformation

$$u' = au - b^*v, \qquad v' = bu + a^*v, \quad |a|^2 + |b|^2 = 1,$$

transforms the homogeneous polynomials of degree $2j$ in the two variables u and v, that is, linear combinations of the $2j+1$ products $u^{2j}, u^{2j-1}v, \ldots v^{2j}$, among themselves. Define

$$\xi_m = \frac{u^{j+m}v^{j-m}}{\sqrt{(j+m)!(j-m)!}}, \quad m = -j, \ldots, j.$$

Then, using the complex numbers a and b as the group parameters (which, because $|a|^2 + |b|^2 = 1$, means three real parameters), one finds

$$R_{a,b}\xi_m(u,v) = \xi_m(u',v') = \sum_{m'} \mathcal{D}^{(j)}_{m'm}(a,b)\xi_{m'}(u,v)$$

$$= \frac{(au - b^*v)^{j+m}(bu + a^*v)^{j-m}}{\sqrt{(j+m)!(j-m)!}}$$

$$= \sum_{\mu\nu} \frac{\sqrt{(j+m)!(j-m)!}}{(j+m-\mu)!\mu!(j-m-\nu)!\nu!} a^{j+m-\mu}a^{*\nu}(-b^*)^\mu b^{j-m-\nu}u^{2j-\mu-\nu}v^{\mu+\nu},$$

from which one obtains the representation matrices

$$\mathcal{D}^{(j)}_{m'm}(a,b) = a^{j+m}a^{*j-m'}b^{m'-m}\sum_{\mu} c^{j}_{\mu m m'}\left(-\frac{|b|^2}{|a|^2}\right)^{\mu}, \tag{E.72}$$

$$c^{j}_{\mu m m'} \stackrel{\text{def}}{=} \frac{[(j+m)!(j-m)!(j+m')!(j-m')!]^{1/2}}{(j+m-\mu)!\mu!(j-m'-\mu)!(m'-m+\mu)!}.$$

To express these matrices in terms of the Euler angles, use (5.37), with the result that again $\mathcal{D}^{j}_{m'm}(\alpha,\beta,\gamma) = e^{-i(m\alpha+m'\gamma)}d^{j}_{m'm}(\beta)$ and

$$d^{j}_{m'm}(\beta) = (\cos\frac{1}{2}\beta)^{2j}(\cot\frac{1}{2}\beta)^{m-m'}\sum_{\mu} c^{j}_{\mu m m'}(-1)^{\mu}(\tan\frac{1}{2}\beta)^{2\mu}, \tag{E.73}$$

which shows that if j is an integer, the representation is even, and if j is a half-integer, the representation is odd. The proofs that these representations are irreducible and that there are no other irreducible representations of SU(2) are the same as those for SO(3).[41]

The infinitesimal generators In order to obtain the generators of SO(3) it is simplest to take advantage of the fact that the rotations about the z-axis form a subgroup of SO(3), and we found in (E.66) that the generator of this two-dimensional rotation group is $\mathfrak{X}_3 = x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2}$. Similarly, the rotations about the other two axes are, separately, subgroups and their generators are formed in an analogous way. Therefore we may conclude that the three generators of SO(3) are given by

$$\mathfrak{X}_i = -\sum_{j,k=1}^{3}\epsilon_{ijk}x_j\frac{\partial}{\partial x_k}, \tag{E.74}$$

where $\epsilon_{ijk} = 1$ for $i,j,k = 1,2,3$ and ϵ_{ijk} is totally anti-symmetric in all its indices, that is,

$$\mathfrak{X}_1 = x_3\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_3}$$

and its cyclic permutations. It is then an easy calculation to obtain the commutation relations

$$[\mathfrak{X}_1,\mathfrak{X}_2] = \mathfrak{X}_3,\ \ [\mathfrak{X}_3,\mathfrak{X}_1] = \mathfrak{X}_2,\ \ [\mathfrak{X}_2,\mathfrak{X}_3] = \mathfrak{X}_1. \tag{E.75}$$

This implies that the structure constants defined by (E.45) are $\mathfrak{c}^3_{12} = \mathfrak{c}^2_{31} = \mathfrak{c}^1_{23} = 1$, and the Casimir operator, as defined by (E.61), is given by

$$\mathfrak{C} = -2\sum_{l}\mathfrak{X}_l^2. \tag{E.76}$$

---

[41] **The proof that these representations are unitary is left as an exercise.**

Clebsch-Gordan coefficients

Let $\Psi(j_1, m_1)$ and $\Psi(j_2, m_2)$ be vectors in the carrier spaces of irreducible representations of SO(3) of dimensions $2j_1 + 1$ and $2j_2 + 1$, respectively, in other words, eigenvectors of two independent angular momenta $\vec{\mathbf{J}}^{(1)2}$, $\vec{\mathbf{J}}^{(2)2}$, and $\mathbf{J}_z^{(1)}$, $\mathbf{J}_z^{(2)}$, respectively. We now want to construct the Clebsch-Gordan coefficients, connecting the eigenstates of the total angular momentum $\vec{\mathbf{J}} = \vec{\mathbf{J}}^{(1)} + \vec{\mathbf{J}}^{(2)}$ to the Kronecker product of those of the two individual angular momenta, in the same notation as in (5.64),

$$\Psi(j, m) = \sum_{m_1+m_2=m} \langle j, m | j_1, m_1; j_2, m_2 \rangle^* \Psi(j_1, j_2; m_1, m_2),$$

where $\Psi(j_1, j_2; m_1, m_2) \stackrel{\text{def}}{=} \Psi(j_1, m_1) \otimes \Psi(j_2, m_2)$. In order to do that, define $\vec{\mathbf{J}} = \hbar(\vec{\mathscr{J}}_1 + \vec{\mathscr{J}}_2)$ and $\mathbf{j}_\pm^{(i)}$ as in (5.7), and utilize (5.11) and (5.12). As we saw in Chapter 5, when $m = j$, then $\mathbf{j}_+ \Psi(j, j) = 0$, so that, if

$$\Psi(j, j) = \sum_m a_m \Psi(j_1, m) \otimes \Psi(j_2, j - m),$$

where $a_m = \langle j, j | j_1, m; j_2, j - m \rangle$, we must have

$$\begin{aligned}
0 &= (\mathbf{j}_+^{(1)} + \mathbf{j}_+^{(2)}) \Psi(j, j) \\
&= \sum_m a_m [\sqrt{(j_1 - m)(j_1 + m + 1)} \Psi(j_1, m + 1) \otimes \Psi(j_2, j - m) \\
&\quad + \sqrt{(j_2 - j + m)(j_2 + j - m + 1)} \Psi(j_1, m) \otimes \Psi(j_2, j - m + 1)] \\
&= \sum_m [\sqrt{(j_1 - m)(j_1 + m + 1)} a_m \\
&\quad + \sqrt{(j_2 - j + m + 1)(j_2 + j - m)} a_{m+1}] \Psi(j_2, m + 1) \otimes \Psi(j_2, j - m),
\end{aligned}$$

which implies that

$$a_{m+1} = -\sqrt{\frac{(j_1 - m)(j_1 + m + 1)}{j_2 - j + m + 1)(j_2 + j - m)}}\, a_m. \tag{E.77}$$

This determines all the a_m in terms of he lowest, a_{-j_1}, and the latter can then be fixed (to within a phase, which is arbitrary) by normalization. In order to construct all the other Clebsch-Gordan coefficients, apply $\mathbf{j}_-$ to $\Psi(j, j)$ and use (5.12). Explicit expressions for all of them can be found in [Wigner].

Since $\langle jm | j_1, m_1; j_2, m_2 \rangle = (\Psi(j, m), \Psi(j_1, m_1) \otimes \Psi(j_2, m_2))$ and all these angular-momentum eigenstates are orthonormal, the Clebsch-Gordan coef-

ficients form a unitary matrix:[42]

$$\sum_{j,m}\langle j,m|j_1,m_1;j_2,m_2\rangle\langle j,m|j_1,m_1';j_2,m_2'\rangle^* = \delta_{m_1'm_1}\delta_{m_2'm_2} \tag{E.78}$$

as well as

$$\sum_{m_1,m_2}\langle j',m'|j_1,m_1;j_2,m_2\rangle\langle j,m|j_1,m_1;j_2,m_2\rangle^* = \delta_{jj'}\delta_{mm'}. \tag{E.79}$$

The group O(3)

The group of all real orthogonal 3×3 matrices T includes those with $\det T = -1$; these are the improper rotations, which combine a proper rotation with an *inversion* of the coordinate system. This group can be regarded as the direct product of SO(3) with the Abelian group of order two whose two elements are E and I, with $I^2 = E$. The group O(3) therefore has two irreducible representations, $D^{(l)}_{\pm}$ for each irreducible representation $D^{(l)}$ of SO(3), those for which

$$D^{(l)}_{+}(IR) = D^{(l)}_{+}(R),$$

and those for which

$$D^{(l)}_{-}(IR) = -D^{(l)}_{-}(R).$$

[42]Note that j_1 and j_2 are always fixed and never summed over.

E.5 Problems and Exercises

1. Prove that a group whose order is a prime number must be Abelian.

2. Consider the group S_3, whose multiplication table is given earlier in section E.1. Here are three representations:

   1) $D^{(1)} = 1$ for all elements.

   2) $D^{(2)}(E) = D^{(2)}(D) = D^{(2)}(F) = 1$, and $D^{(2)}(A) = D^{(2)}(B) = D^{(2)}(C) = -1$.

   3)

$$
\begin{aligned}
D^{(3)}(E) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad D^{(3)}(A) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
D^{(3)}(B) &= \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} \qquad D^{(3)}(C) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix} \\
D^{(3)}(D) &= \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix} \qquad D^{(3)}(F) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}
\end{aligned}
$$

   Verify that these are, indeed, representations. Write down the characters. Is the third representation irreducible? If not, reduce it; if it is, are there any other irreducible representations?

3. Prove that the representations of $SU(2)$ given in terms of (E.73) are unitary.

4. Prove that all the irreducible representations of a group $\mathcal{G}$ are one-dimensional if and only if $\mathcal{G}$ is Abelian.

5. Consider the *tetrahedral group T,* which consists of the 12 rotations that leave a regular tretrahedron invariant: there are 3 rotations C_2 by 180^o, four rotations C_3 by 120^o and four rotations C_3^2 by 240^o. Show that they form four classes and find the dimensions of all the irreducible representations. Also construct the table of primitive characters, and, finally, all the irreducible representations. (The orthogonality theorems are the essential tools for this.)

6. Let $\mathsf{A} \stackrel{\text{def}}{=} \sum_i c_i \mathfrak{X}_i$ and $\mathsf{B} \stackrel{\text{def}}{=} \sum_i c'_i \mathfrak{X}_i$ be two elements of a Lie algebra with the structure constants $\mathfrak{c}_{ij}^s$. Show that $[\mathsf{A}, \mathsf{B}] = \mathsf{C}$ with $\mathsf{C} = \sum_s c''_s \mathfrak{X}_s$, where $c''_s = \sum_{ij} c_i c'_j \mathfrak{c}_{ij}^s$.

This page intentionally left blank

References

[Bander] H. Bander and C. Itzykson. Group theory and the hydrogen atom, *Reviews of Modern Physics,* **38**, 330–345 and 346–358 (1966).

[Bohm, A.] A. Bohm and M. Gadella. *Dirac Kets, Gamow Vectors and Gel'fand Triplets,* New York: Springer-Verlag, 1989.

[Bohm, D.] D. Bohm. *Quantum Theory*, New York: Prentice Hall, 1951.

[Canright] G.S. Canright and S.M. Girvin. Fractional statistics: quantum possibilities in two dimensions, *Science* **247**, 1197–1205 (1990).

[Chadan] K. Chadan and P.C. Sabatier. *Inverse Problems in Quantum Scattering Theory,* 2nd ed., New York: Springer-Verlag, 1989.

[Crichton] J.H. Crichton. Phase-shift ambiguities for spin-independent scattering, *Nouvo Cimento* **45**, 256 (1966).

[Duck] I. Duck and E.C.G. Sudarshan. Toward an understanding of the spin-statistics theorem, *American Journal of Physics,* **66**, 284–303 (1998).

[EPR] A. Einstein, B. Podolsky, and N. Rosen. Can quantum mechanical description of physical reality be considered complete?, *Physical Review* **47**, 777 (1935).

[Erdélyi] A. Erdélyi et al., eds. *Higher Transcendental Functions,* New York: McGraw-Hill, 1953.

[Feynman] R.P. Feynman and A.R. Hibbs. *Quantum Mechanics and Path Integrals,* New York: McGraw-Hill, 1965.

[Ford 59a] K.W. Ford et al. Quantum effects near a barrier maximum, *Annals of Physics (NY)*, **7**, 239–258 (1959).

[Ford 59b] K.W. Ford and J.A. Wheeler. Semiclassical description of scattering, *Annals of Physics (NY)*, **7**, 259–288, 287 (1959); Application of semiclassical scattering analysis, *ibid*, 287–322.

[Fröman] N. Fröman and P.O. Fröman. *JWKB Approximation*, Amsterdam: North-Holland Publ. Co., 1965.

[Goldstein] H. Goldstein. *Classical Mechanics,* Reading, MA: Addison-Wesley, 1980.

[Hamermesh] M. Hamermesh. *Group Theory*, Reading, MA: Addison-Wesley, 1962.

[Jackson] J.D. Jackson. *Classical Electrodynamics,* 2nd ed., New York: John Wiley & Sons, 1975.

[Klauder] J.R. Klauder and E.C.G. Sudarshan. *Fundamentals of Quantum Optics,* New York: W.A. Benjamin, 1968.

[Lieb] E.H. Lieb and M. Loss. *Analysis,* 2nd ed., Providence, RI: American Mathematical Society, 2001.

[Newton 61] R.G. Newton. The exponential decay law of unstable systems, *Annals of Physics* **14**, 333–345 (1961).

[Newton 76] R.G. Newton. Optical theorem and beyond, *American Journal of Physics,* **44**, 639–642 (1976).

[Newton 82] R.G. Newton. *Scattering Theory of Waves and Particles*, 2nd ed., New York: Springer-Verlag, 1982.

[Newton 00] R.G. Newton. *Thinking about Physics*, Princeton: Princeton University Press, 2000.

[Reed] M. Reed and B. Simon. *Methods of Modern Mathematical Physics,* New York: Academic Press, 1978.

[Rose] M.E. Rose. *Relativistic Electron Theory,* New York: John Wiley & Sons, 1961.

[Shapere] A. Shapere and F. Wilczek, eds. *Geometric Phases in Physics,* Singapore: World Scientific, 1989.

[Wigner] E.P. Wigner. *Group Theory,* New York: Academic Press, 1959.

Index

δ-normalized, 7
π-light, 251
σ-light, 251

Abelian group, 355
absorption, 133
  of photons, 244
  probability, 248
acausal, 3
accidental degeneracy, *see* degeneracy, accidental
action integral, 35
action-at-a-distance, 10
adiabatic approximation, 235, 255
  change, 255
Airy integral, 185
ammonia molecule, 80
angular momentum, 43, 93, 94, 161
  conservation in the Dirac equation, 302
  precession, 218
  spectrum, 146, 148
annihilation operators, 270
anti-commutation relations, 26, 55–56
anti-commutator, 26, 55
antilinear, 316
antilinear operator, 143
antiunitary operator, 173
antiunitary transformation, 143
anyons, 266
associated Laguerre polynomials, 114, 352
Auger effect, 259
auto-ionization, 257

back-propagator, 129
barrier penetration, 73
barycentric coordinates, 98
  frame, 110
basis, 6, 317, 318
beams, 31, 69
Berry's phase, 255
Berry, M.V., 255
Bessel's equation, 347
Bloch functions, 82
Bohm theory, 11
Bohm, David, 10, 165
Bohr magneton, 162, 218, 222, 229, 301
Bohr radius, 206
Bohr, Niels, 112
Born approximation, 135, 242
Born series, 135, 337

radius of convergence, 338
Bose, Satyendranath, 269
Bose-Einstein condensation, 269
Bose-Einstein statistics, 26, 268, 277
boson, 268–270, 274
boson attraction, 274
bound state, 75, 339
embedded in the continuum, 101, 259
of zero binding energy, 109
bounded operator, 320
Brillouin, L., 182

c-number, 145
canonical commutation relations, 19, 39, 53, 59, 60
canonical momenta, 37
canonical transformation, 29, 35, 63, 358
carrier space of a representation, 155, 284, 357, 358
Cartan, E., 384
Casimir operator, 155, 384
of SO(3), 394
Cauchy principal value, 311
Cauchy sequence, 318
Cayley's theorem, 355
Cayley-Klein parameters, 153, 390
ccrs, *see* canonical commutation relations
center-of-mass coodinates, *see* barycentric coordinates
center-of-mass motion, 98
center-of-mass reference frame, *see* barycentric frame
centrifugal barrier, 110
centrifugal potential energy, 94
channels, 261, 262
closed, 262
open, 262
character of a representation, 361
charge conjugate, 291, 294, 301
circular polarization, 62
class, 152, 356
Clebsch-Gordan
coefficients, 163, 395–396
for SO(3), 395
series, 367
coarse grain, 13, 42, 241
coherence, 16
degree of, 41
collapse of the wave function, 21
collective vibrations, 56
commutator, 11
commuting observables, 5, 33
complete set of, 5, 17
commuting operators, 330
compact Lie group, 373
compact operator, 334, 336
complete linear vector space, 318
complete set of commuting Hermitian operators, 330
completeness, 6, 17
completeness relation, 130, 324
complex extension of a Lie algebra, 379
complex potential, 175
composition law, 30
Compton wavelength, 297
conduction bands, 82
configuration
representation, 30
space, 19
wave function, 43, 68
confluent hypergeometric
equation, 119
function, 351
conformal group, 374
conjugate group element, 356
connecting formulas, 184
conservation laws, 40
conservation of flux, 71, 134
conservation of probability, 134
conserved quantities, 40
constant of the motion, 146
continuity equation, 45
continuous spectrum, 18, 323
coordinate representation, 21
correct linear combinations for degenerate perturbation theory, 200–204
correlation, 2, 3, 10
Coulomb barrier, 120
Coulomb gauge, 59
Coulomb potential, 112
creation operators, 270
cross section, 111
crossing of energy levels, 198–199

current density, 45
current operator four-vector, 302
cyclotron frequency, 220, 223
cyclotron radius, 219

d-wave, 95
deflection function, 193
degeneracy, 6, 12, 68, 78, 91, 95, 116, 323
  accidental, 204, 371
  normal, 175, 371
delta function, 311
  three-dimensional, 313
dense set of vectors, 318
density matrix, 21, 160
density of final states, 238, 242
  for photons, 250
density of particles, 271
density operator, 13, 15, 41, 160
diagonal, 29
diagonalize, 30, 329
differential operator, 22
differential scattering cross section, 104
dimension of a vector space, 317
dipole approximation, 247, 250
dipole transition, 248
Dirac equation, 293, 298, 301
Dirac notation, viii
Dirac picture, 236
Dirac, P.A.M., 112
direct product, 366
  of groups, 356
direct sum, 317
discrete spectrum, 323
dispersion, 5, 11, 48
dispersion-free, 11
distribution, 311
disturbance, 4, 8, 9
domain of an operator, 319, 321
double-valued representations
  of O(2), 387
  of SO(3), 393
dynamical variables, 3, 19

Ehrenfest's theorem, 44
eigenspace, 323, 330
eigenstate, 4, 330
eigenvalue, 4, 323
  embedded in the continuum, 323
eigenvector, 323
Einstein A coefficient, 251
Einstein B coefficient, 248
Einstein, Albert, 269
electric charge density, 302
electric current density, 302
electric dipole moment, 208, 248
  effective, 223
  induced, 208
electric dipole radiation, 251
electric field, 206
electrical conduction, 82
electromagnetic field, 156, 244, 274, 280, 296, 299
  angular momentum, 61
  momentum, 61
  spin operator, 61
electromagnetic radiation, 244
emission of photons, 244
energy, 40, 46
  not conserved, 235
  positive, 296
energy bands, 81
energy density, 58
energy splitting, 79
energy-momentum relation, relativistic, 296
ensemble, 14
entanglement, 2, 8, 10, 165
EPR, 8
equations of motion, 34–36
equivalent representations, 358
essential spectrum, 323
Euler angles, 153, 391, 394
Euler's theorem, 143, 151
Euler-Lagrange equations, 37
even representations of SU(2), 393
exchange degeneracy, 267
exchange energy, 275
exchange reaction, 259
exclusion principle, 27, 268
expectation value, 5, 11, 320
  relativistic, 285
expected value, *see* expectation value
exponential decay law, 240–241

factorizing an operator, 150
faithful representations, 357

Fermi surface, 269
Fermi's golden rule, 239
Fermi-Dirac statistics, 268, 277
fermion, 268–270, 274
Feynman, Richard, 112, 302
Feynman-Hellmann formula, *see* Hellmann-Feynman formula
field operator, 23
field quanta, 55
fine structure, 225, 280, 307
fine-structure constant, 206
first Fredholm minor, 335
Floquet's theorem, 82
flux, 71
  density, 69
  ratio, 104
Fock space, 270, 271
Fock, V., 270
forbidden transitions, 171, 172, 249
Fredholm
  alternative, 136, 334
  determinant, 135, 335
    modified, 336
  equation
    of the first kind, 333
    of the second kind, 333
  method, 135, 334
Frobenius *ansatz*, 113, 115, 305, 342, 348
fully reducible representation, 358

g-factor, 162, 226
  effective for atom, 228
Gamow factor, 120
gauge invariance, 65
gauge transformation, 221
  non-abelian, 65
  of the first kind, 63, 64
  of the second kind, 64, 65
Gaunt coefficients, 163
Gaussian integral, 49
Gaussian wave packets, 87
generating function
  for associated Laguerre polynomials, 352
  for Hermite polynomials, 348
  for Laguerre polynomials, 351
  for Legendre polynomials, 342
generator, 35, 39, 43, 144
  of a Lie group, 377
  of rotations, 143, 151, 155, 159, 281
  of SO(3), 394
  of translations, 20, 52
geometric phase, 255
geometrical optics, 181, 194
GL(n), 373
golden rule, 239
good quantum number, 146
Gordon, W., 280
graviton, 170
Green's function
  complete, 128–136
  free, 127, 128
ground state, 101
group of rigid motions, 374
group velocity, 50
gyromagnetic ratio, 162, 218, 301

Haar measure, 153, 381
  for O(2), 387
  for SO(3), 388
half-bound state, 109
half-integer representation, 364
half-life, 240
Hamilton's characteristic function, 182
Hamilton's principal function, 182
Hamilton's principle, 36
Hamilton-Jacobi equation, 182
Hamiltonian, 39, 44, 52, 280, 293, 295, 297
  effective, 263
  in Fock space, 273
  of the field, 53
  of the helium atom, 210
  relativistic, 205
Hamiltonian equations of motion, 37, 39
Hamiltonian flow, 42, 241
harmonic oscillator, 54, 60, 82, 148
  in two dimensions, 219
  wave functions, 83
Heaviside function, 312
Heisenberg picture, 33, 40, 42
Heisenberg's equation of motion, 39
Heisenberg's uncertainty relation, 12
helicity, 62, 294, 296
helium atom, 259

Hellmann-Feynman formula, 101, 206
Helmholtz equation, 92
Hermite polynomials, 83, 220, 221
Hermitian, 320
    adjoint, 320
    conjugate, 320
    operator, 3, 323
hidden variables, 11
high energy, 135, 181
Hilbert space, 1, 318
    rigged, 326
    separable, 318
Hilbert-Schmidt class, 336, 337
hole theory, 301
homomorphic, 355
Hurwitz integral, 153, 381
hydrogen atom, 112
    Dirac equation, 305
hydrogen spectrum, 21cm line, 232
hyperfine structure, 231–232
hypergeometric equation, 118
    series, 350

idempotent, 320
ignored outcome, 14
improper rotations, 396
incoherent, 164
indistinguishable, 26, 27, 265
inelastic scattering, 104, 133, 243, 257, 258
infinitesimal generators, *see* generator
infinitesimal transformations, 375
infinitesimal variations, 34
initial condition, 70
initial-value problem, 47
inner product, 316
integrable functions, 319
integral kernel, 330
integral operator, 18
interaction picture, 236
interference, 9, 10
intermediate state, 244
invariant subgroup, 356
invariant subspace, 358
inverse of an operator, 321
inversion on spheres, 374
irreducible representation, 358
isolated state, 2
isometric, 328
isomorphic, 355

Jacobi identity, 378
Jacobi polynomials, 392
Jeffreys, J., 182
Jost solution, 123

kernel of an operator, 322
kinetic-energy operator, 67
Klein, O., 280
Klein-Gordon equation, 280
Kramers's theorem, 176
Kramers, H.A., 182
Kronecker symbol, 317
Kush, Polykarp, 232

l-degeneracy, 113, 116–117, 306
laboratory frame, 110
Lagrange's theorem, 355
Lagrangian, 35, 44, 279
Lagrangian density, 52, 53, 280, 293
Lagrangian equations of motion, 37, 52, 53
Laguerre polynomials, 207
Lamb shift, 226, 307
Landé g-factor, 169
Laplace's equation, 352
Laplace-Runge-Lenz vector, 116
large components of a relativistic spinor, 299, 303
Larmor frequency, 218
Legendre polynomials, 342
Legendre transformation, 37
Legendre's differential equation, 341
Levinson's theorem, 110
Lie algebra, 155
Lie, Sophus, 379
lifetime, 108, 241, 253, 264
line breadth, natural, 253
linear group, 373
linear operator, 319
linear vector space, 315
    real, 315
linearly independent, 317
linebreadth, 252
Lippmann-Schwinger equation, 129, 134, 338
    homogeneous, 136

local wave number, 183
local wavelength, 183
logarithmic derivative, 75
Lorentz four-vector, 288
Lorentz group, 286
Lorentz transformation, 281, 286–288, 292, 301
low energies, 106
lowering operator, 85, 90
LS coupling scheme, 169

m-degeneracy, 148
magnetic field, 218, 226
magnetic moment, 162
  anomalous of the electron, 232
  effective, 218, 228
  induced, 229
  intrinsic, 162
  of a nucleus, 229
  of an atom, 169
  of the neutron, 162
  of the proton, 162, 229
magnetic quantum number, 162
Marchenko equation, 122, 125
matrix element, 319
Maurer's relation, 377
Maxwell equations, 58, 65, 302
mean-square deviation, 5
measurement, 3, 27, 41
mechanical momentum, 64
minimal coupling, 65
minimal wave packet, 85
mixed state, 15, 160
mixing, degree of, 16
momentum operator, 20
momentum representation, 30, 68
momentum space, 22
momentum spectrum
  discretized, 23
momentum transfer, 135
momentum-space wave function, 43
motion reversal, 172
multiplication law for Lie groups, 373
multiply connected group manifold, 385

Neumann series, 135, 337
nodal surfaces, 100
nodes, 77, 78
Noether's theorem, 40, 146
nonlocal potential, 137, 140, 174
nonrelativistic limit of the Dirac equation, 301
norm
  of a vector, 316
  of an operator, 320
normal degeneracy, 371
normal divisor, 356
normal subgroup, 356
normalized, 316
nuclear magneton, 162, 229
nullspace, 322
number operator, 25, 27

O(2), 374, 386
O(3), 374, 396
O(4), 117, 281
O(5), 281
O(n), 282, 373
occupation-number representation, 26, 269, 270
odd representations of SU(2), 393
one-particle state, 271
optical theorem, 106, 132
  generalized, 106, 125, 134
orbiting, 108, 260
order of a group, 355
orthogonal, 6, 316
orthogonal complement, 317
orthogonal group, 373
orthonormal basis, 27, 317
  quasi-basis, 7
outgoing waves, 71
outgoing-wave boundary condition, 127, 128

p-wave, 95
parabolic coordinates, 114–116, 119
parity transformation, 68
Parseval's equality, 317, 324, 326, 328
partial-wave expansion, 105
particle-wave duality, 2
particles from quantum field, 55
Paschen-Back effect, 226
path integrals, 33
Pauli exclusion principle, 56, 268, 270, 301
Pauli matrices, 150, 157

Pauli term, 223, 280
Pauli, Wolfgang, 222
periodic potentials, 82
phase information, 2
phase shift, 103
  in the *WKB* approximation, 191
  variational calculation, 213
phase velocity, 50
phonons, 56, 57
photons, 62, 156, 166
Planck distribution, 252, 277
Planck's formula, 251
plane polarization, 62, 161
point spectrum, 209, 323
Poisson brackets, 39
polarization
  degree of, 160
  of radiation, 246
polarization vector, 160
positive definite, 320
  semi-definite, 320
positrons, 301
predissociation, 259
primitive character, 361
principal quantum number, 113
principal value, 311
principle of stationary action, 36
probabilistic predictions, 3
probabilistic theories, 4, 241
probabilities, 6
  addition of, 10
probability amplitude, 7
probability conservation, 45
probability density, 7, 22, 45, 48
probability-current density, 21
  four-vector, 298
probability-density amplitude, 21, 23
projection operator, 13, 320
projection postulate, 4
propagator, 42, 96, 128
  retarded, 42, 127
pure rotation group, 386

QED, *see* quantum electrodynamics
quantum electrodynamics, 112, 114, 162, 223, 226, 302, 307
quantum field, 5, 27, 162, 245
quasi-basis, 33, 330
quasi-eigenfunction, 325
quasi-eigenvalue, 4, 323
quasi-eigenvectors, 4, 29
quasi-vector, 326

radial equation, 99
radial Hamiltonian, 101
radial quantum number, 113
radiation condition, 103
radiation field, 62
radiation gauge, 59, 244
radiationless transition, 259
radioactive decay, 3, 15
raising operator, 90
Ramsauer-Townsend effect, 106
range of an operator, 319
ray in Hilbert space, 2, 6
Rayleigh-Jeans law, 252
reciprocity, 132, 240
reduced mass, 99, 112
reduced matrix element, 172
reducible representation, 358
reflection coefficient, 71, 77
reflection operator, 281–283
reflections, 68, 159
regular representation, 362
  of a Lie group, 383
regular solution of the Schrödinger equation, 95, 122
relative coordinates, 98
relative motion, 98
relativistic field equation, 279
relativistic Schrödinger equation, 280
representations of groups
  double-valued, 154, 158
  even, 154
  integer, 364
  irreducible, 154, 204
  odd, 154
  of type 1, 364
  of type 2, 364
  of type 3, 365
representations of vectors, 16, 27
repulsion of energy levels, 199
resolution of the identity, 324, 328
resolvent operator, 195
resonance, 73, 107, 108, 110, 260, 264
  Breit-Wigner curve, 107
  zero-energy, 109
resonance scattering, 257

resonance shape, 260
reversibility, 240
Riccati-Bessel function, 95, 100, 347
Riccati-Hankel function, 347
Riccati-Neumann function, 102, 347
Riemann-Lebesgue lemma, 50
rigid translation, 43
Rodrigues's formula, 344
Rollnik class, 338
root-mean-square deviation, 11
rotation, 43
  active, 145
  generator, 282, 283
  infinitesimal, 285
  invariance, 43, 105, 145
  matrix, 151
  of the coordinate system, 144
  of the physical system, 144
  operator, 281, 283, 286
  passive, 145
rotation-reflection group, 386
  in two dimensions, 387
Rutherford cross section, 120, 135
Rydberg, 225

S matrix, 108, 134
  unitarity, 134
s-wave, 95
scalar potential, 58
scattering amplitude, 104
  WKB approximation, 193
  phase near resonance, 264
scattering cross sections, 242
scattering length, 106, 109
Schläfli's integral, 343
Schrödinger equation, 41, 45, 53, 55, 70
  time-independent, 46, 67
Schrödinger picture, 41, 42
Schur's lemma, 167, 204, 359
Schwarz's inequality, 6, 12, 316
  for the operator norm, 320
Schwinger, Julian, 36, 112, 148, 302
second quantization, 53, 302
selection rule, 167, 169, 171, 239, 248, 251
self-adjoint operator, 3, 321
self-conjugate subgroup, 356
semiclassical approximation, 182
semisimple Lie algebra, 380
shadow scattering, 133
similitude group, 374
simple Lie Algebra, 380
simply reducible, 372
singlet states, 164
SL(n), 373
Slater coefficients, 163
Slater determinant, 267
slow changes, 235
small components of a relativistic spinor, 299, 303
SO(3), 151, 388
SO(n), 374
Sommerfeld quantization rule, 188
Sommerfeld, Arnold, 112
spatial translation, 43
special linear group, 373
special unitary group, 374
spectral gaps, 81
spectral intensity, 239
spectral line shape, 253
spectral projection, 328
spectral theorem, 324, 327
spectrum, 5, 6
  continuous, 323, 326
  of a Hermitian operator, 322
spherical Bessel function, 95, 347
spherical Hankel function, 347
spherical harmonics, 93, 94
spherical Neumann function, 347
spherical polar coordinates, 92, 99
spherical vector, 156
spin
  density operator, 164
  direction, 159
  exchange operator, 165, 275, 288
  of the photon, 61
  operator, 160
  precession, 223
spin-1, 156, 160
spin-1/2, 222
spin-orbit
  coupling, 224
  energy shift, 225
spin-statistics connection, 56, 268, 295, 296, 301
spinor, 157, 164
  adjoint, 285, 287

relativistic, 284, 286, 293
splitting of energy levels, 199
spontaneous emission of radiation, 249
spontaneous symmetry breaking, 68, 78, 96, 208
square-integrable functions, 319, 321
square-summable sequences, 319
state, 1
anti-symmetric, 56
isolated, 2
preparation of, 4
pure, 1
two-particle, 26
unstable, 260
stationary phase, 50, 71, 103, 193
statistical mechanics, 42
steady state, 75
Stern-Gerlach experiment, 162
strictly forbidden transitions, 249
structure constants
of a Lie group, 377
of SO(3), 394
Sturm-Liouville problem, 77
SU(2), 154
SU(n), 374
subgroup, 355
subspace, 317
subsystem, 2, 7, 14
sudden approximation, 235, 240, 243
summation convention, 292
superposition principle, 2, 10
superselection rules, 1
symmetric operator, 3, 321
symmetrical top, 391

tensor product, 2, 317, 318
tensors, 145
Cartesian, 170
spherical, 170–172
tetrahedral group, 397
Thomas factor, 224
threshold, 258
time delay, 51, 73, 74, 107, 260
classical, 73
time reversal, 172
invariance, 132, 173, 174
effect on degeneracy, 175–176
operator, 173
time-development operator, 39
time-translation invariance, 40
Tomonaga, Sin-Itiro, 112, 302
total momentum, 43, 52
total scattering cross section, 105
trace of an operator, 13, 329
transformation function, 30
transformations
active, 144
passive, 144
transition
energy nonconserving, 243
induced, 239
probability, 171, 238, 274
per unit time, 238, 239
virtual, 239
translation, 74
invariance, 52
transmission coefficient, 71, 77
trial functions, 210
triangle inequality, 316
for the operator norm, 320
triplet states, 164
tunneling, 73, 80, 188
turning point, 184
two-boson wave function, 272
two-particle spectrum, 258
two-slit experiment, 2, 14

U(n), 374
unfaithful representations, 357
unit vector, 316
unitary
group, 374
operator, 329
representation, 358
transformation, 28, 143
universal covering group, 154, 385
unpolarized, 160
unstable state, 108, 207, 241, 252, 260, 264

vacuum state, 25
variance, 5, 11, 48
vector potential, 58, 60, 219, 245
detectable effects, 222
virial of Clausius, 44
virial theorem, 44, 85
virtual state, 244

virtual transition, 238, 239
volume of a Lie group, 382

wave function, 26, 45
    anti-symmetric, 27, 265, 273
    collapse of, 41
    normalization, 103
    spreading, 50, 51
    symmetric, 27, 265, 273
    time reversal, 174
    two-particle, 27
Wenzel, G., 182
width of a state, 253
Wigner coefficients, 163
Wigner's theorem, 143, 173
Wigner, E. P., 143
Wigner-Eckart theorem, 172
WKB approximation, 182
Wronskian, 71

Zeeman effect
    intermediate, 228
    strong, 226
    weak-field, 227
zero-point energy, 83
Zitterbewegung, 297
zonal harmonics, 100