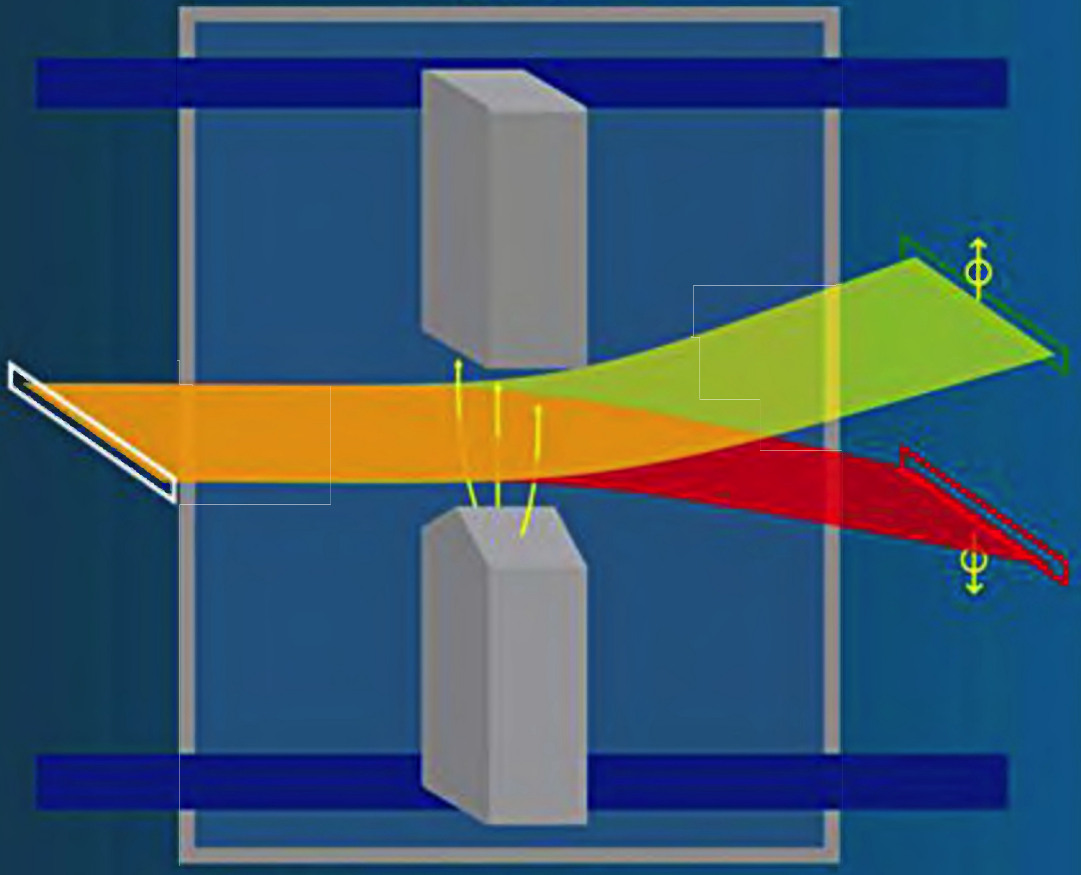
A Modern Approach to QUANTUM MECHANICS

Second Edition



John S. Townsend

University Science Books
www.uscibooks.com

Production Manager: *Paul C. Anagnostopoulos, Windfall Software*
Copyeditor: *Lee A. Young*
Proofreader: *MaryEllen N. Oliver*
Text Design: *Yvonne Tsang*
Cover Design: *Genette Itoko McGrew*
Illustrator: *Lineworks*
Compositor: *Windfall Software*
Printer & Binder: *Edwards Brothers, Inc.*

This book was set in Times Roman and Gotham and composed with ZzTEX, a macro package for Donald Knuth's TEX typesetting system.

This book is printed on acid-free paper. ∞

ISBN: 978-1-891389-78-8

Library of Congress Cataloging-in-Publication Data

Townsend, John S.
A modern approach to quantum mechanics / John S. Townsend. — 2nd ed.
p. cm.
Includes index.
ISBN 978-1-891389-78-8 (alk. paper)
1. Quantum theory—Textbooks. I. Title.
QC174.12.T69 2012
530.12—dc23 2011049655

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

Contents

Preface

There have been two revolutions in the way we view the physical world in the twentieth century: relativity and quantum mechanics. In quantum mechanics the revolution has been both profound—requiring a dramatic revision in the structure of the laws of mechanics that govern the behavior of all particles, be they electrons or photons—and far-reaching in its impact—determining the stability of matter itself, shaping the interactions of particles on the atomic, nuclear, and particle physics level, and leading to macroscopic quantum effects ranging from lasers and superconductivity to neutron stars and radiation from black holes. Moreover, in a triumph for twentieth-century physics, special relativity and quantum mechanics have been joined together in the form of quantum field theory. Field theories such as quantum electrodynamics have been tested with an extremely high precision, with agreement between theory and experiment verified to better than nine significant figures. It should be emphasized that while our understanding of the laws of physics is continually evolving, always being subjected to experimental scrutiny, so far no confirmed discrepancy between theory and experiment for quantum mechanics has been detected.

This book is intended for an upper-division course in quantum mechanics. The most likely audience for the book consists of students who have completed a course in modern physics that includes an introduction to quantum mechanics that emphasizes wave mechanics. Rather than continue with a similar approach in a second course, I have chosen to introduce the fundamentals of quantum mechanics through a detailed discussion of the physics of intrinsic spin. Such an approach has a number of significant advantages. First, students find starting a course with something "new" such as intrinsic spin both interesting and exciting, and they enjoy making the connections with what they have seen before. Second, spin systems provide us with many beautiful but straightforward illustrations of the essential structure of quantum mechanics, a structure that is not obscured by the mathematics of wave mechanics. Quantum mechanics can be presented through concrete examples. I believe that most physicists learn through specific examples and then find it easy to generalize. By

starting with spin, students are given plenty of time to assimilate this novel and striking material. I have found that they seem to learn this key introductory material easily and well—material that was often perceived to be difficult when I came to it midway through a course that began with wave mechanics. Third, when we do come to wave mechanics, students see that wave mechanics is only one aspect of quantum mechanics, not the fundamental core of the subject. They see at an early stage that wave mechanics and matrix mechanics are just different ways of calculating based on the same underlying quantum mechanics and that the approach they use depends on the particular problem they are addressing.

I have been inspired by two sources, an "introductory" treatment in Volume III of *The Feynman Lectures on Physics* and an advanced exposition in J. J. Sakurai's *Modern Quantum Mechanics*. Overall, I believe that wave mechanics is probably the best way to *introduce* students to quantum mechanics. Wave mechanics makes the largest overlap with what students know from classical mechanics and shows them the strange behavior of quantum mechanics in a familiar environment. This is probably why students find their first introduction to quantum mechanics so stimulating. However, starting a second course with wave mechanics runs the risk of diminishing much of the excitement and enthusiasm for the entirely new way of viewing nature that is demanded by quantum mechanics. It becomes sort of old hat, material the students has seen before, repeated in more depth. It is, I believe, with the second exposure to quantum mechanics that something like Feynman's approach has its best chance to be effective. But to be effective, a quantum mechanics text needs to make lots of contact with the way most physicists think and calculate in quantum mechanics using the language of kets and operators. This is Sakurai's approach in his graduate-level textbook. In a sense, the approach that I am presenting here can be viewed as a superposition of these two approaches, but at the junior-senior level.

Chapter 1 introduces the concepts of the quantum state vector, complex probability amplitudes, and the probabilistic interpretation of quantum mechanics in the context of analyzing a number of Stern–Gerlach experiments carried out with spin-$\frac{1}{2}$ particles. By introducing ket vectors at the beginning, we have the framework for thinking about states as having an existence quite apart from the way we happen to choose to represent them, whether it be with matrix mechanics, which is discussed at length in Chapter 2, or, where appropriate, with wave mechanics, which is introduced in Chapter 6. Moreover, there is a natural role for operators; in Chapter 2 they rotate spin states so that the spin "points" in a different direction. I do not follow a postulatory approach, but rather I allow the basic physics of this spin system to drive the introduction of concepts such as Hermitian operators, eigenvalues, and eigenstates.

In Chapter 3 the commutation relations of the generators of rotations are determined from the behavior of ordinary vectors under rotations. Most of the material in this chapter is fairly conventional; what is not so conventional is the introduc-

tion of operator techniques for determining the angular momentum eigenstates and eigenvalue spectrum and the derivation of the uncertainty relations from the commutation relations at such an early stage. Since so much of our initial discussion of quantum mechanics revolves around intrinsic spin, it is important for students to see how quantum mechanics can be used to determine from first principles the spin states that have been introduced in Chapters 1 and 2, without having to appeal only to experimental results.

Chapter 4 is devoted to time evolution of states. The natural operation in time development is to translate states forward in time. The Hamiltonian enters as the generator of time translations, and the states are shown to obey the Schrödinger equation. Most of the chapter is devoted to physical examples. In Chapter 5 another physical system, the spin-spin interaction of an electron and proton in the ground state of hydrogen, is used to introduce the spin states of two spin-$\frac{1}{2}$ particles. The total-spin-0 state serves as the basis for a discussion of the Einstein–Podolsky–Rosen (EPR) paradox and the Bell inequalities.

The main theme of Chapter 6 is making contact with the usual formalism of wave mechanics. The special problems in dealing with states such as position and momentum states that have a continuous eigenvalue spectrum are analyzed. The momentum operator enters naturally as the generator of translations. Sections 6.8 through 6.10 include a general discussion with examples of solutions to the Schrödinger equation that can serve as a review for students with a good background in one-dimensional wave mechanics.

Chapter 7 is devoted to the one-dimensional simple harmonic oscillator, which merits a chapter all its own. Although the material in Chapter 8 on path integrals can be skipped without affecting subsequent chapters (with the exception of Section 14.1, on the Aharonov–Bohm effect), I believe that path integrals should be discussed, if possible, since this formalism provides real insight into quantum dynamics. However, I have found it difficult to fit this material into our one-semester course, which is taken by all physics majors as well as some students majoring in other disciplines. Rather, I have chosen to postpone path integrals to a second course and then to insert the material in Chapter 8 before Chapter 14. Incidentally, the material on path integrals is the only part of the book that may require students to have had an upper-division classical mechanics course, one in which the principle of least action is discussed.

Chapters 9 through 13 cover fully three-dimensional problems, including the two-body problem, orbital angular momentum, central potentials, time-independent perturbations, identical particles, and scattering. An effort has been made to include as many physical examples as possible.

Although this is a textbook on nonrelativistic quantum mechanics, I have chosen to include a discussion of the quantized radiation field in the final chapter, Chapter 14. The use of ket and bra vectors from the beginning and the discussion of solutions

to problems such as angular momentum and the harmonic oscillator in terms of abstract raising and lowering operators should have helped to prepare the student for the exciting jump to a quantized electromagnetic field. By quantizing this field, we can really understand the properties of photons, we can calculate the lifetimes for spontaneous emission from first principles, and we can understand why a laser works. By looking at higher order processes such as photon-atom scattering, we can also see the essentials of Feynman diagrams. Although the atom is treated nonrelativistically, it is still possible to gain a sense of what quantum field theory is all about at this level without having to face the complications of the relativistic Dirac equation. For the instructor who wishes to cover time-dependent perturbation theory but does not have time for all of the chapter, Section 14.5 stands on its own.

Although SI units are the standard for undergraduate education in electricity and magnetism, I have chosen in the text to use Gaussian units, which are more commonly used to describe microscopic phenomena. However, with the possible exception of the last chapter, with its quantum treatment of the electromagnetic field, the choice of units has little impact. My own experience suggests that students who are generally at home with SI units are comfortable (as indicated in a number of footnotes through the text) replacing e^2 with $e^2/4\pi\epsilon_0$ or ignoring the factor of c in the Bohr magneton whenever they need to carry out numerical calculations. In addition, electromagnetic units are discussed in Appendix A.

In writing the second edition, I have added two sections to Chapter 5, one on entanglement and quantum teleportation and the other on the density operator. Given the importance of entanglement in quantum mechanics, it may seem strange, as it does to me now, to have written a quantum mechanics textbook without explicit use of the word entanglement. The concept of entanglement is, of course, at the heart of the discussion of the EPR paradox, which focused on the entangled state of two spin-$\frac{1}{2}$ particles in a spin-singlet state. Nonetheless, it wasn't until the early 1990s, when topics such as quantum teleportation came to the fore, that the importance of entanglement as a fundamental resource that can be utilized in novel ways was fully appreciated and the term entanglement began to be widely used. I am also somewhat embarrassed not to have included a discussion of the density operator in the first edition. Unlike a textbook author, the experimentalist does not necessarily have the luxury of being able to focus on pure states. Thus there is good reason to introduce the density operator (and the density matrix) as a systematic way to deal with mixed states as well as pure states in quantum mechanics. I have added a section on coherent states of the harmonic oscillator to Chapter 7. Coherent states were first derived by Schrödinger in his efforts to find states that satisfy the correspondence principle. The real utility of these states is most apparent in Chapter 14, where it is seen that coherent states come closest to representing classical electromagnetic waves with a well-defined phase. I have also added a section to Chapter 14 on cavity quantum electrodynamics, showing how the interaction of the quantized electromagnetic

field with atoms is modified by confinement in a reflective cavity. Like quantum teleportation, cavity quantum electrodynamics is a topic that really came to the fore in the 1990s. In addition to these new sections, I have added numerous worked example problems to the text, with the hope that these examples will help students in mastering quantum mechanics. I have also increased the end-of-chapter problems by 25 percent.

There is almost certainly enough material here for a full-year course. For a one-semester course, I have covered the material through Chapter 12, omitting Sections 6.7 through 6.10 and, as noted earlier, Chapter 8. The material in the latter half of Chapter 6 is covered thoroughly in our introductory course on quantum physics. See John S. Townsend, *Quantum Physics: A Fundamental Approach to Modern Physics*, University Science Books, 2010. In addition to Chapter 8, other sections that might be omitted in a one-semester course include parts of Chapter 5, Section 9.7, and Sections 11.5 through 11.9. Or one might choose to go as far as Chapter 10 and reserve the remaining material for a later course.

A comprehensive solutions manual for the instructor is available from the publisher, upon request of the instructor.

Finally, some grateful acknowledgments are certainly in order. Students in my quantum mechanics classes have given me useful feedback as I have taught from the book over the years. Colleagues at Harvey Mudd College who have offered valuable comments as well as encouragement include Bob Cave, Chih-Yung Chen, Tom Donnelly, Tom Helliwell, Theresa Lynn, and Peter Saeta. Art Weldon of West Virginia University suggested a number of ways to improve the accuracy and effectiveness of the first edition. This text was initially published in the McGraw-Hill International Series in Pure and Applied Physics. I have benefited from comments from the following reviewers: William Dalton, St. Cloud State University; Michael Grady, SUNY-Fredonia; Richard Hazeltine, University of Texas at Austin; Jack Mochel, University of Illinois at Urbana-Champaign; and Jae Y. Park, North Carolina State University. For the first edition, the Pew Science Program provided support for Doug Dunston and Doug Ridgway, two Harvey Mudd College students, who helped in the preparation of the text and figures, respectively, and Helen White helped in checking the galley proofs. A number of people have kindly given me feedback on the material for the second edition, including Rich Holman, Carnegie Mellon University; Randy Hulet, Rice University; Jim Napolitano, RPI; Tom Moore and David Tanenbaum, Pomona College; and John Taylor, University of Colorado.

I have been fortunate to have the production of the book carried out by a very capable group of individuals headed by Paul Anagnostopoulos, the project manager. In addition to Paul, I want to thank Lee Young for copyediting, Joe Snowden for entering the copyedits and laying out the pages, Tom Webster for the artwork, MaryEllen Oliver for her amazingly thorough job of proofreading, Yvonne Tsang for text design, and Genette Itoko McGrew for her creative cover design. I also wish

to thank Jane Ellis and Bruce Armbruster of University Science Books not only for their assistance but also for the care and attention to detail they have taken in preparing this new edition of the book. And I especially want to thank my wife, Ellen, for cheerfully letting me devote so much time to this project.

Please do not hesitate to contact me if you find errors or have suggestions that might improve the book.

John S. Townsend
Department of Physics
Harvey Mudd College
Claremont, CA 91711
townsend@hmc.edu

A Modern Approach to Quantum Mechanics

CHAPTER 1

Stern–Gerlach Experiments

We begin our discussion of quantum mechanics with a conceptually simple experiment in which we measure a component of the intrinsic spin angular momentum of an atom. This experiment was first carried out by O. Stern and W. Gerlach in 1922 using a beam of silver atoms. We will refer to the measuring apparatus as a Stern–Gerlach device. The results of experiments with a number of such devices are easy to describe but, as we shall see, nonetheless startling in their consequences.

1.1 The Original Stern–Gerlach Experiment

Before analyzing the experiment, we need to know something about the relationship between the intrinsic spin angular momentum of a particle and its corresponding magnetic moment. To the classical physicist, angular momentum is always orbital angular momentum, namely, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Although the Earth is said to have spin angular momentum $I\omega$ due to its rotation about its axis as well as orbital angular momentum due to its revolution about the Sun, both types of angular momentum are just different forms of $\mathbf{L}$. The intrinsic spin angular momentum $\mathbf{S}$ of a microscopic particle is not at all of the same sort as orbital angular momentum, but it is real angular momentum nonetheless.

To get a feeling for the relationship that exists between the angular momentum of a charged particle and its corresponding magnetic moment, we first use a classical example and then point out some of its limitations. Consider a point particle with charge q and mass m moving in a circular orbit of radius r with speed v. The magnetic moment μ is given by

$$\mu = \frac{IA}{c} = \left(\frac{q}{T}\right)\frac{\pi r^2}{c} = \frac{qvr}{2c} = \frac{q}{2mc}L \tag{1.1}$$

where A is the area of the circle formed by the orbit, the current I is the charge q divided by the period $T = (2\pi r/v)$, and $L = mvr$ is the orbital angular momentum of the particle.[1] Since the magnetic moment and the orbital angular momentum are parallel or antiparallel depending on the sign of the charge q, we may express this relationship in the vector form

$$\boldsymbol{\mu} = \frac{q}{2mc}\mathbf{L} \tag{1.2}$$

This relationship between $\mathbf{L}$ and $\boldsymbol{\mu}$ turns out to be generally true whenever the mass and charge coincide in space. One can obtain different constants of proportionality by adjusting the charge and mass distributions independently. For example, a solid spherical ball of mass m rotating about an axis through its center with the charge q distributed uniformly only on the surface of the ball has a constant of proportionality of $5q/6mc$.

When we come to intrinsic spin angular momentum of a particle, we write

$$\boldsymbol{\mu} = \frac{gq}{2mc}\mathbf{S} \tag{1.3}$$

where the value of the constant g is experimentally determined to be $g = 2.00$ for an electron, $g = 5.58$ for a proton, or even $g = -3.82$ for a neutron.[2] One might be tempted to presume that g is telling us about how the charge and mass are distributed for the different particles and that intrinsic spin angular momentum is just orbital angular momentum of the particle itself as it spins about its axis. We will see as we go along that such a simple classical picture of intrinsic spin is entirely untenable and that the intrinsic spin angular momentum we are discussing is a very different beast indeed. In fact, it appears that even a point particle in quantum mechanics may have intrinsic spin angular momentum.[3] Although there are no classical arguments that we can give to justify (1.3), we can note that such a relationship between the

---

[1] If you haven't seen them before, the Gaussian units we are using for electromagnetism may take a little getting used to. A comparison of SI and Gaussian units is given in Appendix A. In SI units the magnetic moment is just IA, so you can ignore the factor of c, the speed of light, in expressions such as (1.1) if you wish to convert to SI units.

[2] Each of these g factors has its own experimental uncertainty. Recent measurements by B. Odom, D. Hanneke, B. D'Urso, and G. Gabrielse, *Phys. Rev. Lett.* **97**, 030801 (2006), have shown that $g/2$ for an electron is 1.00115965218085(76), where the factor of 76 reflects the uncertainty in the last two places. Relativistic quantum mechanics predicts that $g = 2$ for an electron. The deviations from this value can be accounted for by quantum field theory. The much larger deviations from $g = 2$ for the proton and the (neutral) neutron are due to the fact that these particles are not fundamental but are composed of charged constituents called quarks.

[3] It is amusing to note that in 1925 S. Goudsmit and G. Uhlenbeck as graduate students "discovered" the electron's spin from an analysis of atomic spectra. They were trying to understand why the optical spectra of alkali atoms such as sodium are composed of a pair of closely spaced lines, such as the sodium doublet. Goudsmit and Uhlenbeck realized that an additional degree of freedom (an independent coordinate) was required, a degree of freedom that they could understand only if they assumed the electron was a small ball of charge that could rotate about an axis.

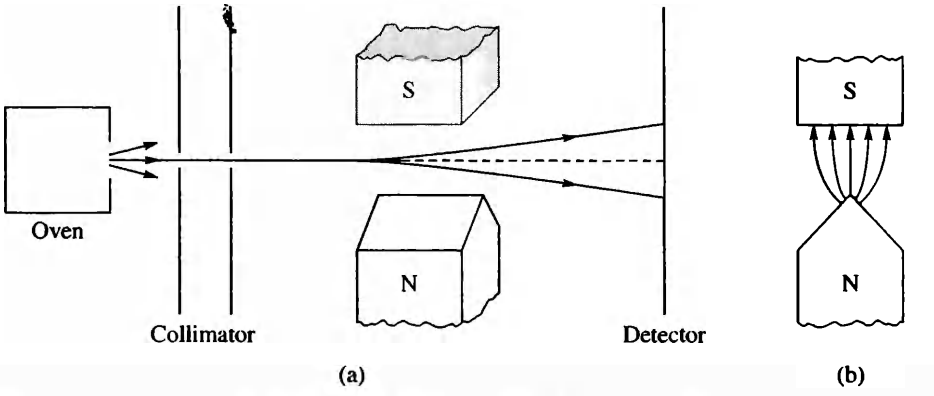


Figure 1.1 (a) A schematic diagram of the Stern–Gerlach experiment. (b) A cross-sectional view of the pole pieces of the magnet depicting the inhomogeneous magnetic field they produce.

magnetic moment and the intrinsic spin angular momentum is at least consistent with dimensional analysis. At this stage, you can think of g as a dimensionless factor that has been inserted to make the magnitudes as well as the units come out right.

Let's turn to the Stern–Gerlach experiment itself. Figure 1.1a shows a schematic diagram of the apparatus. A collimated beam of silver atoms is produced by evaporating silver in a hot oven and selecting those atoms that pass through a series of narrow slits. The beam is then directed between the poles of a magnet. One of the pole pieces is flat; the other has a sharp tip. Such a magnet produces an inhomogeneous magnetic field, as shown in Fig. 1.1b. When a *neutral* atom with a magnetic moment $\boldsymbol{\mu}$ enters the magnetic field $\mathbf{B}$, it experiences a force $\mathbf{F} = \nabla(\boldsymbol{\mu} \cdot \mathbf{B})$, since $-\boldsymbol{\mu} \cdot \mathbf{B}$ is the energy of interaction of a magnetic dipole with an external magnetic field. If we call the direction in which the inhomogeneous magnetic field gradient is large the z direction, we see that

$$F_z = \boldsymbol{\mu} \cdot \frac{\partial \mathbf{B}}{\partial z} \simeq \mu_z \frac{\partial B_z}{\partial z} \tag{1.4}$$

In this way they could account for the electron's spin angular momentum and magnetic dipole moment. The splitting of the energy levels that was needed to account for the doublet could then be understood as due to the potential energy of interaction of the electron's magnetic moment in the *internal* magnetic field of the atom (see Section 11.5). Goudsmit and Uhlenbeck wrote up their results for their advisor P. Ehrenfest, who then advised them to discuss the matter with H. Lorentz. When Lorentz showed them that a classical model of the electron required that the electron must be spinning at a speed on the surface approximately ten times the speed of light, they went to Ehrenfest to tell him of their foolishness. He informed them that he had already submitted their paper for publication and that they shouldn't worry since they were "both young enough to be able to afford a stupidity." *Physics Today*, June 1976, pp. 40–48.

Notice that we have taken the magnetic field gradient $\partial B_z/\partial z$ in the figure to be negative, so that if μ_z is negative as well, then F_z is positive and the atoms are deflected in the positive z direction. Classically, $\mu_z = |\mu| \cos\theta$, where θ is the angle that the magnetic moment μ makes with the z axis. Thus μ_z should take on a continuum of values ranging from $+\mu$ to $-\mu$. Since the atoms coming from the oven are not polarized with their magnetic moments pointing in a preferred direction, we should find a corresponding continuum of deflections. In the original Stern–Gerlach experiment, the silver atoms were detected by allowing them to build up to a visible deposit on a glass plate. Figure 1.2 shows the results of this original experiment. The surprising result is that μ_z takes on only *two* values, corresponding to the values $\pm\hbar/2$ for S_z. Numerically, $\hbar = h/2\pi = 1.055 \times 10^{-27}$ erg $\cdot$ s $= 6.582 \times 10^{-16}$ eV $\cdot$ s, where h is Planck's constant.

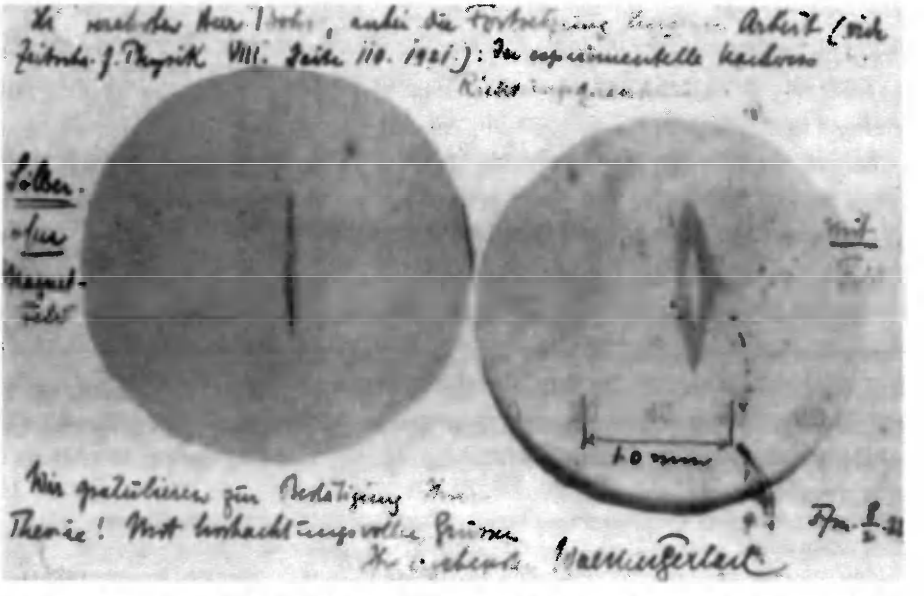


Figure 1.2 A postcard from Walther Gerlach to Niels Bohr, dated February 8, 1922. Note that the images on the postcard have been rotated by 90° relative to Fig. 1.1, where the collimating slit is horizontal. The left-hand image of the beam profile without the magnetic field shows the effect of the finite width of this collimating slit. The right-hand image shows the beam profile with the magnetic field. Only in the center of the apparatus is the magnitude of the magnetic field gradient sufficiently strong to cause splitting. The pattern is smeared because of the range of speeds of the atoms coming from the oven. Translation of the message: "My esteemed Herr Bohr, attached is the continuation of our work [vide *Zeitschr. f. Phys.* **8**, 110 (1921)]: the experimental proof of directional quantization. We congratulate you on the confirmation of your theory! With respectful greetings. Your most humble Walther Gerlach." Photograph reproduced with permission from the Niels Bohr Archive.

Silver atoms are composed of 47 electrons and a nucleus. Atomic theory tells us the total orbital and total spin angular momentum of 46 of the electrons is equal to zero, and the 47th electron has zero orbital angular momentum. Moreover, as (1.3) shows, the nucleus makes a very small contribution to the magnetic moment of the atom because the mass of the nucleus is so much larger than the mass of the electron. Therefore, the magnetic moment of the silver atom is effectively due to the magnetic moment of a single electron. Thus, in carrying out their experiment, Stern and Gerlach measured the component of the intrinsic spin angular momentum of an electron along the z axis and found it to take on only two discrete values, $+\hbar/2$ and $-\hbar/2$, commonly called "spin up" and "spin down," respectively. Later, we will see that these values are characteristic of a spin-$\frac{1}{2}$ particle. Incidentally, we chose to make the bottom N pole piece of the Stern–Gerlach (SG) device the one with the sharp tip for a simple reason. With this configuration, B_z decreases as z increases, making $\partial B_z/\partial z$ negative. As we noted earlier, atoms with a negative μ_z are deflected upward in this field. Now an electron has charge $q = -e$ and from (1.3) with $g = 2$, $\mu_z = (-e/m_e c)S_z$. Thus a silver atom with $S_z = \hbar/2$, a spin-up atom, will conveniently be deflected upward.

1.2 Four Experiments

Now that we have seen how the actual Stern–Gerlach experiment was done, let's turn our attention to four simple experiments that will tell us much about the structure of quantum mechanics. If you like, you can think of these experiments as thought experiments so that we needn't focus on any technical difficulties that might be faced in carrying them out.

EXPERIMENT 1

Let us say a particle that exits an SG**z** device, one with its inhomogeneous magnetic field parallel to the z axis, with $S_z = +\hbar/2$ is in the state $|+\mathbf{z}\rangle$. The symbol $|+\mathbf{z}\rangle$, known as a **ket vector**, is a convenient way of denoting this state. Suppose a beam of particles, *each* of which is in this state, enters another SG**z** device. We find that *all* the particles exit in the state $|+\mathbf{z}\rangle$; that is, the measurement of S_z yields the value $+\hbar/2$ for each of the particles, as indicated in Fig. 1.3a.

EXPERIMENT 2

Consider a beam of particles exiting the SG**z** device in the state $|+\mathbf{z}\rangle$, as in Experiment 1. We next send this beam into an SG**x** device, one with its inhomogeneous magnetic field oriented along the x axis. We find that 50 percent of the particles exit the second device with $S_x = \hbar/2$ and are therefore in the state $|+\mathbf{x}\rangle$, while the other 50 percent exit with $S_x = -\hbar/2$ and are therefore in the state $|-\mathbf{x}\rangle$ (see Fig. 1.3b).

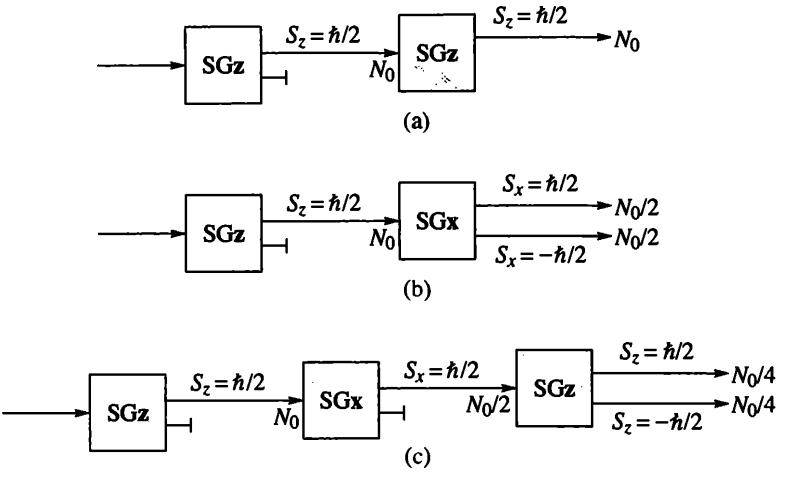


Figure 1.3 A block diagram of (a) Experiment 1, (b) Experiment 2, and (c) Experiment 3. N_0 is the number of particles in the beam exiting the first SG device with $S_z = \hbar/2$.

For completeness, we also note that if we select the beam of particles exiting the initial SGz device in the state $|-\mathbf{z}\rangle$ instead of $|+\mathbf{z}\rangle$ and send this beam through the SGx device, we also find that 50 percent of the particles yield $\hbar/2$ for a measurement of S_x and 50 percent yield $-\hbar/2$ for a measurement of S_x.

EXPERIMENT 3

Let's add a third SG device to Experiment 2, but this time with its inhomogeneous magnetic field oriented along the z axis (see Fig. 1.3c). If we send the beam of particles exiting the SGx device in the state $|+\mathbf{x}\rangle$ through the last SGz device, we find that 50 percent of the particles exit in the state $|+\mathbf{z}\rangle$ and 50 percent exit in the state $|-\mathbf{z}\rangle$. Initially, none of the particles entering the SGx device was in the state $|-\mathbf{z}\rangle$. Thus making the measurement of S_x with the second device has modified the state of the system. We cannot think of the beam entering the last SGz device as comprised of particles with $S_z = \hbar/2$ *and* $S_x = \hbar/2$, as one might expect from the results of the measurements of the first two SG devices. This cannot account for the 50 percent of the beam that exits the last SGz device with $S_z = -\hbar/2$. We will see shortly that S_z and S_x are incompatible observables; namely, we cannot know both of them simultaneously. In the macroscopic world, on the other hand, it seems to be easy to create a state with two definite nonzero components of the angular momentum, as, for example, is the case for a spinning top whose angular momentum is oriented at 45° to both the x and z axes. This is an indication that the quantum world is fundamentally different from our everyday macroscopic experience. We will see this more clearly as we go on to consider the next Stern–Gerlach experiment.

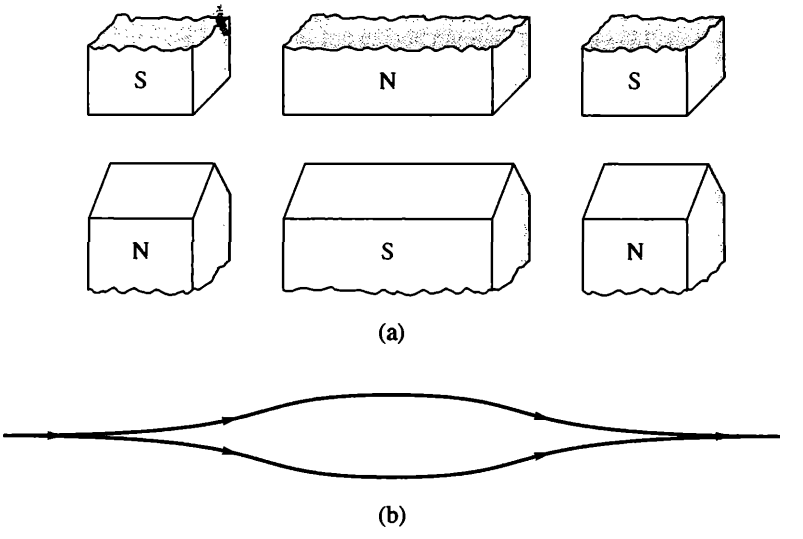


Figure 1.4 (a) The three magnets of a modified Stern–Gerlach device. (b) The paths that a spin-up and spin-down particle would follow in traversing the device.

A MODIFIED SG DEVICE

In Experiment 4 we will use a modified SG device, introduced in thought experiments by Richard Feynman.[4] This SG device, shown in Fig. 1.4a, is comprised of three high-gradient magnets, placed back to back, instead of a single magnet. The first magnet is a typical Stern–Gerlach magnet, followed by a second magnet with the same cross section as the first but twice as long and with the polarity opposite that of the first magnet. This second magnet pushes a particle with a magnetic moment in the opposite direction to the first magnet. Thus, in traversing the first half of the length of this magnet, the particle is decelerated and brought to rest in the transverse direction. In traversing the second half, the particle is accelerated back toward the axis. Although the third magnet is just like the first magnet, here it decelerates the particle so that the particle returns to the axis in the same state as it entered the first magnet. The net effect of the three magnets is to recombine the beams so that their condition upon exiting the third magnet is just like it was before entering the first magnet. Figure 1.4b indicates the paths that spin-up and spin-down particles would follow in this modified SG device.

You might think such a device serves no purpose, but we can use a modified SG device to make a measurement and select a particular spin state. For example,

[4] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison-Wesley, Reading, MA, 1965, vol. 3, Chapter 5.

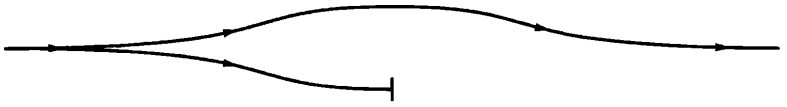


Figure 1.5 Selecting a spin-up state with a modified Stern–Gerlach device by blocking the spin-down state.

if the direction of the inhomogeneous magnetic field of the three magnets is along the x axis, we can select a particle in the $|+\mathbf{x}\rangle$ spin state by blocking the path that a particle in the $|-\mathbf{x}\rangle$ spin state would take, as indicated in Fig. 1.5. Then all the particles exiting the modified three-magnet SG**x** device would be in the state $|+\mathbf{x}\rangle$. In fact, we can repeat Experiment 3 with the SG**x** device replaced by a modified SG**x** device. If the $|-\mathbf{x}\rangle$ state is filtered out by inserting a block in the lower path, we find, of course, exactly the same results as in Experiment 3; that is, when we measure with the last SG**z** device, we find 50 percent of the particles in the state $|+\mathbf{z}\rangle$ and 50 percent in the state $|-\mathbf{z}\rangle$. Similarly, if we filter out the state $|+\mathbf{x}\rangle$ by inserting a block in the upper path, we also find 50 percent of the particles exiting the last SG**z** device in the state $|+\mathbf{z}\rangle$ and 50 percent in the state $|-\mathbf{z}\rangle$.

EXPERIMENT 4

We are now ready for Experiment 4. As in Experiment 3, a beam of particles in the state $|+\mathbf{z}\rangle$ from an initial SG**z** device enters an SG**x** device, but in this experiment it is a modified SG**x** device in which we do *not* block one of the paths and, therefore, do *not* make a measurement of S_x. We then send the beam from this modified SG**x** device into another SG**z** device. As indicated in Fig. 1.6, we find that 100 percent of the particles exit the last SG**z** device in the state $|+\mathbf{z}\rangle$, just as if the modified SG**x** device were absent from the experiment and we were repeating Experiment 1.

Before carrying out Experiment 4, it might seem obvious that 50 percent of the particles passing through the modified SG**x** device are in the state $|+\mathbf{x}\rangle$ and 50 percent are in the state $|-\mathbf{x}\rangle$. But the results of Experiment 4 contradict this assumption, since, if it were true, we would expect to find 50 percent of the particles in the state $|+\mathbf{z}\rangle$ and 50 percent of the particles in the state $|-\mathbf{z}\rangle$ when the unfiltered beam exits the last SG**z** device. Our results are completely incompatible with the hypothesis that the particles traversing the modified SG**x** device have *either* $S_x = \hbar/2$ *or* $S_x = -\hbar/2$.

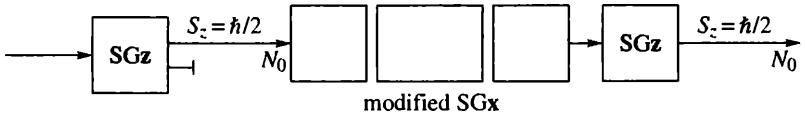


Figure 1.6 A block diagram of Experiment 4. Note that we cannot indicate the path followed through the three-magnet modified SG**x** device since no measurement is carried out to select either a $|+\mathbf{x}\rangle$ or $|-\mathbf{x}\rangle$ spin state.

Moreover, even if we carry out the experiment with a beam of such low intensity that one particle at a time is passing through the SG devices, we still find that each of the particles has $S_z = \hbar/2$ when it leaves the last SG**z** device. Thus, the issue raised by this experiment cannot be resolved by some funny business involving the interactions of the particles in the beams as they pass through the modified SG**x** device.

So far, we have been able to describe the results of these Stern–Gerlach experiments simply in terms of the percentage of particles exiting the SG devices in a particular state because the experiments have been carried out on a beam of particles, namely, on a large number of particles. For a *single* particle, it is generally *not* possible to predict with certainty the outcome of the measurement in advance. In Experiment 2, for example, *before* a measurement of S_x on a particle in the state $|+\mathbf{z}\rangle$, all we can say is that there is a 50 percent *probability* of obtaining $S_x = \hbar/2$ and a 50 percent *probability* of obtaining $S_x = -\hbar/2$. However, probabilities alone do not permit us to understand Experiment 4. We cannot explain the results of this experiment by adding the probabilities that a particle passing through the modified SG**x** device is in the state $|+\mathbf{x}\rangle$ or in the state $|-\mathbf{x}\rangle$, since this fails to account for the differences when comparing the results of Experiment 3, in which 50 percent of the particles in the state $|+\mathbf{x}\rangle$ (or $|-\mathbf{x}\rangle$) yield $S_z = -\hbar/2$, with the results of Experiment 4, in which none of the particles has $S_z = -\hbar/2$ when exiting the last SG**z** device. Somehow in Experiment 4 we must eliminate the probability that the particle is in the state $|-\mathbf{z}\rangle$ when it enters the last SG**z** device. What we need is some sort of "interference" that can cancel out the $|-\mathbf{z}\rangle$ state. Such interference is common in the physics of waves, where two waves can interfere destructively to produce minima as well as constructively to produce maxima. With electromagnetic waves, for example, it isn't the intensities that interfere but rather the electromagnetic fields themselves. For electromagnetic waves the intensity is proportional to the square of the amplitude of the wave. With this in mind, for our Stern–Gerlach experiments we introduce a **probability amplitude** that we will "square" to get the probability. If we don't observe which path is taken in the modified SG**x** device by inserting a block, or filter, we must add the amplitudes to take the two different paths corresponding to the $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ states. Even a single particle can have an amplitude to be in both states, to take both paths; when we add, or superpose, the amplitudes, we obtain an amplitude for the particle to be in the state $|+\mathbf{z}\rangle$ only.[5] In summary, when we *don't* make a measurement in the modified SG device, we must add the amplitudes, *not* the probabilities.

---

[5] In Section 2.3 we will discuss in more detail how this interference in Experiment 4 works. These results are reminiscent of the famous double-slit experiment, in which it seems logical to suppose that the particles go through one slit *or* the other, but the interference pattern on a distant screen is completely incompatible with this simple hypothesis. The double-slit experiment is discussed briefly in Section 6.7. If you are unfamiliar with this experiment from the perspective of quantum mechanics, an excellent discussion is given in *The Feynman Lectures on Physics*, vol. 3, Chapter 1.

1.3 The Quantum State Vector

In our description of the state of a particle in quantum mechanics, we have been using a new notation in which states, such as $|+\mathbf{z}\rangle$, are denoted by abstract vectors called ket vectors. Such a description includes as much information about the state of the particle as we are permitted in quantum mechanics. For example, the ket $|+\mathbf{x}\rangle$ is just a shorthand way of saying that the spin state of the particle is such that if we were to make a measurement of S_x, the intrinsic spin angular momentum in the x direction, we would obtain the value $\hbar/2$. There are clearly other attributes that are required to give a complete description of the particle, such as the particle's position or momentum. However, for the time being we are concentrating on the spin degrees of freedom of the particle.[6] Later, in Chapter 6, we will see how to introduce other degrees of freedom in the description of the state of the particle.

Classical physics uses a different type of vector in its description of nature. Some of these ordinary vectors are more abstract than others. For example, consider the electric field $\mathbf{E}$, which is a useful but somewhat abstract vector. If there is an electric field present, we know that a test charge q placed in the field will experience a force $\mathbf{F} = q\mathbf{E}$. Of course, even the force $\mathbf{F}$ will not be observed directly. We would probably allow the particle to be accelerated by the force, measure the acceleration, and then use Newton's law $\mathbf{F} = m\mathbf{a}$ to determine $\mathbf{F}$ and thence $\mathbf{E}$.

Let's suppose the electric field in the location where you are reading this book has a constant value, which you could determine in the way we have just outlined. How do you tell your friends about the value, both magnitude and direction, of $\mathbf{E}$? You might just point in the direction of $\mathbf{E}$ to show its direction. But what if your friends are not present and you want to write down $\mathbf{E}$ on a piece of paper? You would probably set up a coordinate system and choose basis vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ whose direction you could easily communicate. Using this coordinate system, you would denote the electric field as $\mathbf{E} = E_x\mathbf{i} + E_y\mathbf{j} + E_z\mathbf{k}$. In fact, we often use a shorthand notation in which we suppress the unit vectors and just say $\mathbf{E} = (E_x, E_y, E_z)$, although in the notation we will be using in our discussion of quantum mechanics, it would be better to denote this as $\mathbf{E} \rightarrow (E_x, E_y, E_z)$. How do we obtain the value for E_x, for example? We just project the electric field onto the x axis. Formally, we take the dot product to find $E_x = \mathbf{i} \cdot \mathbf{E} = |\mathbf{E}| \cos\theta$, where θ is the angle the electric field $\mathbf{E}$ makes with the x axis, as shown in Fig. 1.7.

Let's return to our discussion of quantum state vectors. If we send a spin-$\frac{1}{2}$ particle into an SG$\mathbf{z}$ device, we obtain only the values $\hbar/2$ and $-\hbar/2$, corresponding to the

[6] The historical development of quantum mechanics initially focused on the more obvious degrees of freedom, such as a particle's position. In fact, Goudsmit was fond of relating how, when confronted with the need to introduce a new degree of freedom for the intrinsic spin of the electron in order to explain atomic spectra, he had to ask Uhlenbeck what was meant by the expression "degree of freedom."

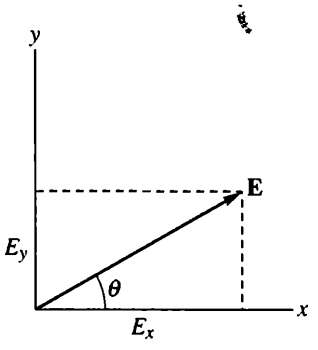


Figure 1.7 The x and y components of an electric field $\mathbf{E}$ making an angle θ with the x axis can be obtained by taking the dot product of $\mathbf{E}$ with the unit vectors $\mathbf{i}$ and $\mathbf{j}$. For a classical vector such as $\mathbf{E}$, E_x and E_y can also be obtained by projecting $\mathbf{E}$ onto the x and y axes.

particle ending up in the state $|+\mathbf{z}\rangle$ or ending down in the state $|-\mathbf{z}\rangle$, respectively. These two states can be considered as vectors that form a basis for our abstract quantum mechanical vector space. If the particle is initially in the state $|+\mathbf{z}\rangle$, we have seen in Experiment 1 that there is zero amplitude for the particle to be found in the state $|-\mathbf{z}\rangle$, which we denote by $\langle -\mathbf{z}|+\mathbf{z}\rangle = 0$. We can think of this as telling us that the vectors are orthogonal, the analogue of $\mathbf{i} \cdot \mathbf{j} = 0$ in our electric field example. Of course if we send a particle in the state $|+\mathbf{z}\rangle$ into an SG**z** device, we always find the particle in the state $|+\mathbf{z}\rangle$. In the language of quantum mechanical amplitudes this is clearly telling us that the amplitude $\langle +\mathbf{z}|+\mathbf{z}\rangle$ is nonzero. As we will see momentarily, it is convenient to require that our quantum mechanical vectors be unit vectors and therefore satisfy $\langle +\mathbf{z}|+\mathbf{z}\rangle = 1$, just as $\mathbf{i} \cdot \mathbf{i} = 1$. We similarly require that $\langle -\mathbf{z}|-\mathbf{z}\rangle = 1$ as well, just as $\mathbf{j} \cdot \mathbf{j} = 1$.

Suppose the particle is in the state $|+\mathbf{x}\rangle$. From Experiment 3 we know that the particle has nonzero amplitudes, which we can call c_+ and c_-, to be in the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$, respectively. We can express this state as $|+\mathbf{x}\rangle = c_+|+\mathbf{z}\rangle + c_-|-\mathbf{z}\rangle$, a linear combination of the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$. In fact, it is convenient at this stage to consider an arbitrary spin state $|\psi\rangle$, which could be created by sending a beam of particles with intrinsic spin-$\frac{1}{2}$ through an SG device with its inhomogeneous magnetic field oriented in some arbitrary direction and selecting those particles that are deflected, for example, upward. In general, this state, like $|+\mathbf{x}\rangle$, will have nonzero amplitudes to yield both $\hbar/2$ and $-\hbar/2$ if a measurement of S_z is made. Thus we will express this state $|\psi\rangle$ as

$$|\psi\rangle = c_+|+\mathbf{z}\rangle + c_-|-\mathbf{z}\rangle \tag{1.5}$$

where the particular values for c_+ and c_- depend on the orientation of the SG device. That an arbitrary state $|\psi\rangle$ can be expressed as a superposition of the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ means that these states form a **complete set**, just as the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ form a complete set for expressing an electric field $\mathbf{E}$ in three dimensions. Although we are describing the states of spin angular momentum of a spin-$\frac{1}{2}$ particle in, of course, three dimensions, we need only the basis states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ to span this two-dimensional vector space.

How can we formally determine the values of c_+ and c_-? In order to take the analogue of the dot product in our ordinary classical vector example, we need to introduce a new type of vector called a **bra vector**.[7] For every ket $|\psi\rangle$ there corresponds a bra $\langle\psi|$. Thus we have two different ways to denote a state with $S_z = \hbar/2$, with the ket $|+\mathbf{z}\rangle$ and the bra $\langle+\mathbf{z}|$. The fate of a bra such as $\langle\varphi|$ is to meet up with a ket $|\psi\rangle$ to form an amplitude, or inner product, $\langle\varphi|\psi\rangle$ in the form of a bracket—hence the name for bras and kets. The amplitude $\langle\varphi|\psi\rangle$ is the probability amplitude for a particle in the state $|\psi\rangle$ to be found in the state $|\varphi\rangle$. From our earlier experiments we know that $\langle-\mathbf{z}|+\mathbf{z}\rangle = 0$, and similarly $\langle+\mathbf{z}|-\mathbf{z}\rangle = 0$, since a particle in the state $|-\mathbf{z}\rangle$, with $S_z = -\hbar/2$, has zero amplitude to be found in the state $|+\mathbf{z}\rangle$, with $S_z = \hbar/2$. Thus from (1.5), we can deduce that

$$\langle+\mathbf{z}|\psi\rangle = c_+\langle+\mathbf{z}|+\mathbf{z}\rangle + c_-\langle+\mathbf{z}|-\mathbf{z}\rangle = c_+ \tag{1.6a}$$

$$\langle-\mathbf{z}|\psi\rangle = c_+\langle-\mathbf{z}|+\mathbf{z}\rangle + c_-\langle-\mathbf{z}|-\mathbf{z}\rangle = c_- \tag{1.6b}$$

or simply $c_\pm = \langle\pm\mathbf{z}|\psi\rangle$. This enables us to express (1.5) in the form

$$|\psi\rangle = \underbrace{\langle+\mathbf{z}|\psi\rangle}_{c_+}|+\mathbf{z}\rangle + \underbrace{\langle-\mathbf{z}|\psi\rangle}_{c_-}|-\mathbf{z}\rangle = |+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle \tag{1.7}$$

where in the last step we have positioned the amplitudes after the kets in a suggestive way. Note that the amplitudes $\langle+\mathbf{z}|\psi\rangle$ and $\langle-\mathbf{z}|\psi\rangle$, the brackets, are (complex) numbers, and thus the product of an amplitude times a ket vector is itself just a ket vector. It really doesn't matter whether we position the amplitude before or after the ket. Writing the ket vector $|\psi\rangle$ in the form (1.7) is analogous to expressing the electric field $\mathbf{E}$ in the form $\mathbf{E} = E_x\mathbf{i} + E_y\mathbf{j} + E_z\mathbf{k} = \mathbf{i}(\mathbf{i}\cdot\mathbf{E}) + \mathbf{j}(\mathbf{j}\cdot\mathbf{E}) + \mathbf{k}(\mathbf{k}\cdot\mathbf{E})$.

Since to each ket there corresponds a bra vector, we must be able to express $\langle\psi|$ in terms of $\langle+\mathbf{z}|$ and $\langle-\mathbf{z}|$ as

$$\langle\psi| = c'_+\langle+\mathbf{z}| + c'_-\langle-\mathbf{z}| \tag{1.8}$$

Using the same technique as before, we see that

$$\langle\psi|+\mathbf{z}\rangle = c'_+\langle+\mathbf{z}|+\mathbf{z}\rangle + c'_-\langle-\mathbf{z}|+\mathbf{z}\rangle = c'_+ \tag{1.9a}$$

$$\langle\psi|-\mathbf{z}\rangle = c'_+\langle+\mathbf{z}|-\mathbf{z}\rangle + c'_-\langle-\mathbf{z}|-\mathbf{z}\rangle = c'_- \tag{1.9b}$$

Thus the bra corresponding to the ket in (1.7) is

$$\langle\psi| = \underbrace{\langle\psi|+\mathbf{z}\rangle}_{c'_+}\langle+\mathbf{z}| + \underbrace{\langle\psi|-\mathbf{z}\rangle}_{c'_-}\langle-\mathbf{z}| \tag{1.10}$$

[7] Mathematicians call the linear vector space spanned by the bra vectors the dual space.

How are the amplitudes $\langle +\mathbf{z}|\psi\rangle$ and $\langle \psi|+\mathbf{z}\rangle$ related? Just as we require that $\langle +\mathbf{z}|+\mathbf{z}\rangle = 1$, we also require that $\langle \psi|\psi\rangle = 1$. We are demanding that all physical vectors in our abstract quantum mechanical vector space be unit vectors. As we will now see, this requirement is crucial to the probabilistic interpretation of quantum mechanics. If we use (1.7) and (1.10) to evaluate $\langle \psi|\psi\rangle$, we find

$$\langle \psi|\psi\rangle = \langle \psi|+\mathbf{z}\rangle\langle +\mathbf{z}|\psi\rangle + \langle \psi|-\mathbf{z}\rangle\langle -\mathbf{z}|\psi\rangle = 1 \tag{1.11}$$

In Section 1.5 we will examine a final Stern–Gerlach experiment that will convince you that amplitudes such as $\langle +\mathbf{z}|\psi\rangle$ and $\langle -\mathbf{z}|\psi\rangle$ are in general complex numbers. The way to guarantee that equality (1.11) is satisfied for arbitrary $|\psi\rangle$'s is to have

$$\langle \psi|+\mathbf{z}\rangle = \langle +\mathbf{z}|\psi\rangle^* \quad \text{and} \quad \langle \psi|-\mathbf{z}\rangle = \langle -\mathbf{z}|\psi\rangle^* \tag{1.12}$$

so that each of the terms in (1.11) is real. These results say that the amplitude for a particle in the state $|\psi\rangle$ to be found in the states $|\pm\mathbf{z}\rangle$ is the complex conjugate of the amplitude for a particle in the states $|\pm\mathbf{z}\rangle$ to be found in the state $|\psi\rangle$.

From (1.6) and (1.9), we see that $c'_+ = c^*_+$ and $c'_- = c^*_-$. Therefore, the bra corresponding to the ket (1.5) is

$$\langle \psi| = c^*_+\langle +\mathbf{z}| + c^*_-\langle -\mathbf{z}| \tag{1.13}$$

The bra vector is generated from the ket vector by changing all the basis kets to their corresponding bras and by changing all amplitudes (complex numbers) to their complex conjugates.

With these results, we can express (1.11) as

$$\begin{aligned}\langle \psi|\psi\rangle &= \langle +\mathbf{z}|\psi\rangle^*\langle +\mathbf{z}|\psi\rangle + \langle -\mathbf{z}|\psi\rangle^*\langle -\mathbf{z}|\psi\rangle \\ &= c^*_+c_+ + c^*_-c_- = 1\end{aligned} \tag{1.14}$$

or

$$\langle \psi|\psi\rangle = |\langle +\mathbf{z}|\psi\rangle|^2 + |\langle -\mathbf{z}|\psi\rangle|^2 = 1 \tag{1.15}$$

where $|\langle +\mathbf{z}|\psi\rangle|^2 \equiv \langle +\mathbf{z}|\psi\rangle^*\langle +\mathbf{z}|\psi\rangle$ and $|\langle -\mathbf{z}|\psi\rangle|^2 \equiv \langle -\mathbf{z}|\psi\rangle^*\langle -\mathbf{z}|\psi\rangle$. We interpret $|\langle +\mathbf{z}|\psi\rangle|^2$ as the probability that a particle in the state $|\psi\rangle$ will be found to be in the state $|+\mathbf{z}\rangle$ if a measurement of S_z is made with an SG**z** device and $|\langle -\mathbf{z}|\psi\rangle|^2$ as the probability that the particle will be found in the state $|-\mathbf{z}\rangle$. As (1.15) shows, the requirement that $\langle \psi|\psi\rangle = 1$ guarantees that the probability of finding the particle in either one state or the other sums to one, since there are only two results possible for a measurement of S_z for a spin-$\frac{1}{2}$ particle.

The striking feature of (1.7) is that when both of the probability amplitudes $\langle +\mathbf{z}|\psi\rangle$ and $\langle -\mathbf{z}|\psi\rangle$ are nonzero, then a particle in the state $|\psi\rangle$ is really in a **superposition** of the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$. There are probabilities of obtaining both $S_z = \hbar/2$ and $S_z = -\hbar/2$ if a measurement of S_z is carried out. This is to be contrasted

with classical mechanics, where for a particle *in a definite state* we do not expect measurements of, say, the orbital angular momentum of the particle at a particular time to yield two different values, such as $\mathbf{r}_1 \times \mathbf{p}_1$ and $\mathbf{r}_2 \times \mathbf{p}_2$.

EXAMPLE 1.1 A measurement of S_z is carried out on a particle in the state

$$|\psi\rangle = \frac{1}{2}|+\mathbf{z}\rangle + \frac{i\sqrt{3}}{2}|-\mathbf{z}\rangle$$

What are the possible results of this measurement and with what probability do these results occur?

SOLUTION Since

$$\langle +\mathbf{z}|\psi\rangle = \frac{1}{2}$$

and consequently

$$|\langle +\mathbf{z}|\psi\rangle|^2 = \frac{1}{4}$$

therefore there is a 25 percent probability of obtaining $S_z = \hbar/2$. Similarly,

$$\langle -\mathbf{z}|\psi\rangle = \frac{i\sqrt{3}}{2}$$

and

$$|\langle -\mathbf{z}|\psi\rangle|^2 = \left(\frac{-i\sqrt{3}}{2}\right)\left(\frac{i\sqrt{3}}{2}\right) = \frac{3}{4}$$

therefore there is a 75 percent probability of obtaining $S_z = -\hbar/2$. Since the state $|\psi\rangle$ is appropriately "normalized," namely

$$\langle\psi|\psi\rangle = |\langle +\mathbf{z}|\psi\rangle|^2 + |\langle -\mathbf{z}|\psi\rangle|^2 = \frac{1}{4} + \frac{3}{4} = 1$$

these probabilities must sum to one since the only results of a measurement of S_z for a spin-$\frac{1}{2}$ particle are $\hbar/2$ and $-\hbar/2$.

1.4 Analysis of Experiment 3

As we noted earlier, Experiment 3 is telling us that a particle in the state $|+\mathbf{x}\rangle$ is in a superposition of the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$: $|+\mathbf{x}\rangle = c_+|+\mathbf{z}\rangle + c_-|-\mathbf{z}\rangle$, since when we make measurements of S_z with the last SG**z** device in the experiment, we have

probabilities of obtaining both $\hbar/2$ and $-\hbar/2$. Because the probabilities are each 50 percent, we have

$$c_+^* c_+ = \langle +\mathbf{z}|+\mathbf{x}\rangle^* \langle +\mathbf{z}|+\mathbf{x}\rangle = |\langle +\mathbf{z}|+\mathbf{x}\rangle|^2 = \tfrac{1}{2} \tag{1.16a}$$

$$c_-^* c_- = \langle -\mathbf{z}|+\mathbf{x}\rangle^* \langle -\mathbf{z}|+\mathbf{x}\rangle = |\langle -\mathbf{z}|+\mathbf{x}\rangle|^2 = \tfrac{1}{2} \tag{1.16b}$$

One solution is to choose c_+ and c_- to be real, namely $c_+ = 1/\sqrt{2}$ and $c_- = 1/\sqrt{2}$. The more general solution for c_+ and c_- may be written as

$$c_+ = \frac{e^{i\delta_+}}{\sqrt{2}} \quad \text{and} \quad c_- = \frac{e^{i\delta_-}}{\sqrt{2}} \tag{1.17}$$

where δ_+ and δ_- are real phases that allow for the possibility that c_+ and c_- are complex.[8] The ket for the state with $S_x = \hbar/2$ is then given by

$$|+\mathbf{x}\rangle = \frac{e^{i\delta_+}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\delta_-}}{\sqrt{2}}|-\mathbf{z}\rangle \tag{1.18}$$

Notice that the probabilities (1.16) themselves do not give us any information about the values of the phases δ_+ and δ_-, since the phases cancel out when we calculate $c_+^* c_+$ and $c_-^* c_-$:

$$c_+^* c_+ = \left(\frac{e^{-i\delta_+}}{\sqrt{2}}\right)\left(\frac{e^{i\delta_+}}{\sqrt{2}}\right) = \frac{1}{2} \tag{1.19a}$$

$$c_-^* c_- = \left(\frac{e^{-i\delta_-}}{\sqrt{2}}\right)\left(\frac{e^{i\delta_-}}{\sqrt{2}}\right) = \frac{1}{2} \tag{1.19b}$$

We can use these probabilities to calculate the average value, or **expectation value**, of S_z, which is the sum of each value obtained by a measurement of S_z multiplied by the probability of obtaining that value:

$$\begin{aligned}\langle S_z\rangle &= c_+^* c_+ \left(\frac{\hbar}{2}\right) + c_-^* c_- \left(-\frac{\hbar}{2}\right) \\ &= \frac{1}{2}\left(\frac{\hbar}{2}\right) + \frac{1}{2}\left(-\frac{\hbar}{2}\right) = 0\end{aligned} \tag{1.20}$$

In this particular case, the expectation value doesn't coincide with any of the values that may be obtained by measuring S_z. An idealized set of data resulting from

---

[8] A common way to express a complex number z is in the form $z = x + iy$, where x and y—the real and imaginary parts of z, respectively—give the location of z in the complex plane. Alternatively, we can express the coordinates for z in the complex plane using the magnitude r of the complex number and its phase ϕ, where $x = r\cos\phi$ and $y = r\sin\phi$. Then $z = re^{i\phi}$, where we have taken advantage of the Euler identity $e^{i\phi} = \cos\phi + i\sin\phi$. The complex conjugate of the complex number $z = x + iy = re^{i\phi}$ is obtained by replacing i by $-i$, that is $z^* = x - iy = re^{-i\phi}$. Therefore, $z^* z = re^{-i\phi} re^{i\phi} = r^2 e^{(-i\phi + i\phi)} = r^2$.

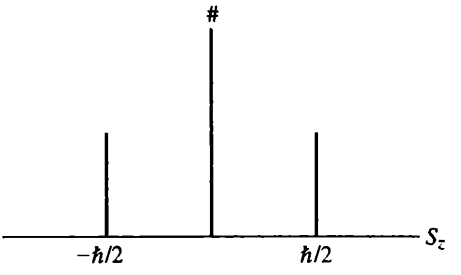


Figure 1.8 An idealized set of data resulting from measurements of S_z on a collection of particles with $S_x = \hbar/2$.

measurements of S_z on a very large collection of particles, each with $S_x = \hbar/2$, is shown in Fig. 1.8. Clearly, there is an inherent **uncertainty** in the result of the measurements, since the measurements do not all yield the same value. We calculate this uncertainty by computing the standard deviation: we determine the average value of the data, take each data point, subtract the average value from it, square and average, and finally take the square root. Thus the square of the uncertainty is given by

$$\begin{aligned}(\Delta S_z)^2 &= \langle (S_z - \langle S_z \rangle)^2 \rangle \\ &= \langle S_z^2 - 2S_z\langle S_z \rangle + \langle S_z \rangle^2 \rangle \\ &= \langle S_z^2 \rangle - 2\langle S_z \rangle \langle S_z \rangle + \langle S_z \rangle^2 \\ &= \langle S_z^2 \rangle - \langle S_z \rangle^2 \end{aligned} \tag{1.21}$$

The expectation value $\langle S_z^2 \rangle$ is the sum of each value of S_z^2 multiplied by the probability of obtaining that value:

$$\begin{aligned}\langle S_z^2 \rangle &= c_+^* c_+ \left(\frac{\hbar}{2}\right)^2 + c_-^* c_- \left(-\frac{\hbar}{2}\right)^2 \\ &= \frac{1}{2}\left(\frac{\hbar^2}{4}\right) + \frac{1}{2}\left(\frac{\hbar^2}{4}\right) = \frac{\hbar^2}{4}\end{aligned} \tag{1.22}$$

Therefore, substituting (1.20) and (1.22) into (1.21), we find $\Delta S_z = \hbar/2$ for a particle in the state $|+\mathbf{x}\rangle$. We call ΔS_z the uncertainty rather than the standard deviation since a *single* particle in the state $|+\mathbf{x}\rangle$ does not have a definite value for S_z.[9]

Of course, $\langle S_z \rangle = 0$ is not in disagreement with finding a single particle to be spin up if we make a measurement of S_z on a particle in the state $|+\mathbf{x}\rangle$. To test predictions such as (1.20) requires a statistically significant sample. Suppose we make measurements of S_z on 100 particles, each in the state $|+\mathbf{x}\rangle$, and find 55 of them to be spin up ($S_z = \hbar/2$) and 45 of them to be spin down ($S_z = -\hbar/2$). Should we be worried about a disagreement with the predictions of quantum mechanics?

[9] The experimental evidence for this assertion will be discussed in Section 5.5.

In general, if we make N measurements, we should expect fluctuations that are on the order of $\sqrt{N}$. Thus with 100 measurements, deviations from $\langle S_z \rangle = 0$ on the order of 10 percent are reasonable. However, if we were to make 10^6 measurements and find 550,000 particles spin up and 450,000 particles spin down, we should be concerned, since we should expect fluctuations of only about $\sqrt{N} = 1{,}000$, rather than the measured 50,000.

EXAMPLE 1.2 As in Example 1.1, a spin-$\frac{1}{2}$ particle is in the state

$$|\psi\rangle = \frac{1}{2}|+\mathbf{z}\rangle + \frac{i\sqrt{3}}{2}|-\mathbf{z}\rangle$$

What are the expectation value $\langle S_z \rangle$ and the uncertainty ΔS_z for this state?

SOLUTION

$$\langle S_z \rangle = |\langle +\mathbf{z}|\psi\rangle|^2 \left(\frac{\hbar}{2}\right) + |\langle -\mathbf{z}|\psi\rangle|^2 \left(-\frac{\hbar}{2}\right)$$

$$= \frac{1}{4}\left(\frac{\hbar}{2}\right) + \frac{3}{4}\left(-\frac{\hbar}{2}\right) = -\frac{\hbar}{4}$$

and

$$\langle S_z^2 \rangle = |\langle +\mathbf{z}|\psi\rangle|^2 \left(\frac{\hbar}{2}\right)^2 + |\langle -\mathbf{z}|\psi\rangle|^2 \left(-\frac{\hbar}{2}\right)^2$$

$$= \frac{1}{4}\left(\frac{\hbar^2}{4}\right) + \frac{3}{4}\left(\frac{\hbar^2}{4}\right) = \frac{\hbar^2}{4}$$

Consequently

$$\Delta S_z = \sqrt{\langle S_z^2 \rangle - \langle S_z \rangle^2}$$

$$= \sqrt{\frac{\hbar^2}{4} - \left(-\frac{\hbar}{4}\right)^2} = \frac{\sqrt{3}}{4}\hbar = 0.43\hbar$$

The uncertainty ΔS_z is $0.43\hbar$ for the state $|\psi\rangle$, which is smaller than the value $0.50\hbar$ for the state $|+\mathbf{x}\rangle$, reflecting the fact that there is a 75 percent probability of obtaining $\hbar/2$ if a measurement of S_z is carried out for the state $|\psi\rangle$ as compared with 50 percent probability for the state $|+\mathbf{x}\rangle$. Of course, if the state of the particle were $|+\mathbf{z}\rangle$, then there would be a 100 percent probability of obtaining $\hbar/2$ if a measurement of S_z is carried out. Correspondingly, ΔS_z vanishes in this case.

1.5 Experiment 5

We are now ready to consider the final Stern–Gerlach experiment of this chapter. In this experiment, Experiment 5, we replace the last SG**z** device of Experiment 3 with one that has its inhomogeneous magnetic field in the y direction and thus make measurements of S_y on particles exiting the SG**x** device in the state $|+\mathbf{x}\rangle$. From Experiment 3 we already know the results of this final experiment. We must find 50 percent of the particles with $S_y = \hbar/2$ and 50 percent of the particles with $S_y = -\hbar/2$. Figure 1.9 shows the last two Stern–Gerlach devices in Experiment 3 and in Experiment 5. Although we are measuring S_y instead of S_z with the last SG device in Experiment 5, the percentage of the particles that go "up" and "down" must be the same for Experiment 3 and Experiment 5, since the axis that we called the z axis in Experiment 3 could just as easily have been called the y axis, either by us or by another observer viewing the experiment. In fact, this sort of argument tells us that if we were to replace the SG**x** device in Experiment 3 with an SG**y** device, we would still find that 50 percent of the particles have $S_z = \hbar/2$ and 50 percent have $S_z = -\hbar/2$ when exiting the last SG**z** device.

These simple results have important implications. Just as we are able to express the state $|+\mathbf{x}\rangle$ by (1.18), we can express the state $|+\mathbf{y}\rangle$ as a superposition of $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ in the form

$$|+\mathbf{y}\rangle = \frac{e^{i\gamma_+}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\gamma_-}}{\sqrt{2}}|-\mathbf{z}\rangle = \frac{e^{i\gamma_+}}{\sqrt{2}}\left[|+\mathbf{z}\rangle + e^{i(\gamma_- - \gamma_+)}|-\mathbf{z}\rangle\right] \qquad (1.23)$$

where we have written the complex numbers multiplying the kets $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ in such a way as to ensure that there is a 50 percent probability of obtaining $S_z = \hbar/2$ and a 50 percent probability of obtaining $S_z = -\hbar/2$. Note that in the last step we pulled out in front an overall phase factor $e^{i\gamma_+}$ for future computational convenience. Moreover, since in Experiment 5 there is a 50 percent probability of finding a particle

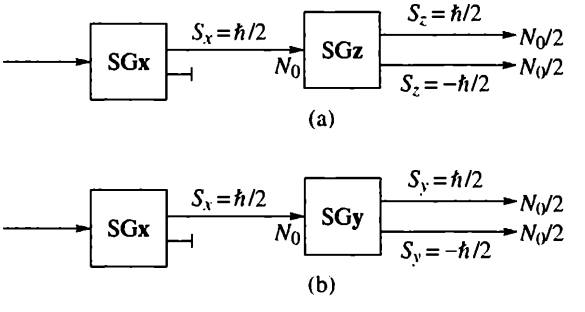


Figure 1.9 Block diagrams showing the last two SG devices in (a) Experiment 3 and in (b) Experiment 5.

with $S_y = \hbar/2$ when it exits the SGx device in the state $|+\mathbf{x}\rangle$, we must have

$$|\langle +\mathbf{y}|+\mathbf{x}\rangle|^2 = \frac{1}{2} \tag{1.24}$$

Now the bra corresponding to the ket (1.23) is

$$\langle +\mathbf{y}| = \frac{e^{-i\gamma_+}}{\sqrt{2}}\langle +\mathbf{z}| + \frac{e^{-i\gamma_-}}{\sqrt{2}}\langle -\mathbf{z}| = \frac{e^{-i\gamma_+}}{\sqrt{2}}\left[\langle +\mathbf{z}| + e^{-i(\gamma_- - \gamma_+)}\langle -\mathbf{z}|\right] \tag{1.25}$$

where we have replaced the complex numbers in (1.23) with their complex conjugates in going from (1.23) to (1.25). If we rewrite (1.18) by pulling out an overall phase factor:

$$|+\mathbf{x}\rangle = \frac{e^{i\delta_+}}{\sqrt{2}}\left[|+\mathbf{z}\rangle + e^{i(\delta_- - \delta_+)}|-\mathbf{z}\rangle\right] \tag{1.26}$$

then

$$\begin{aligned}\langle +\mathbf{y}|+\mathbf{x}\rangle &= \frac{e^{i(\delta_+ - \gamma_+)}}{2}\left(\langle +\mathbf{z}| + e^{-i\gamma}\langle -\mathbf{z}|\right)\left(|+\mathbf{z}\rangle + e^{i\delta}|-\mathbf{z}\rangle\right) \\ &= \frac{e^{i(\delta_+ - \gamma_+)}}{2}\left[1 + e^{i(\delta-\gamma)}\right]\end{aligned} \tag{1.27}$$

where $\delta = \delta_- - \delta_+$ and $\gamma = \gamma_- - \gamma_+$ are the **relative phases** between the kets $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ for these two states, and we have used $\langle +\mathbf{z}|+\mathbf{z}\rangle = \langle -\mathbf{z}|-\mathbf{z}\rangle = 1$ and $\langle +\mathbf{z}|-\mathbf{z}\rangle = \langle -\mathbf{z}|+\mathbf{z}\rangle = 0$ in evaluating the amplitude. We finally calculate the probability:

$$\begin{aligned}|\langle +\mathbf{y}|+\mathbf{x}\rangle|^2 &= \left\{\frac{e^{i(\delta_+ - \gamma_+)}}{2}\left[1 + e^{i(\delta-\gamma)}\right]\right\}\left\{\frac{e^{-i(\delta_+ - \gamma_+)}}{2}\left[1 + e^{-i(\delta-\gamma)}\right]\right\} \\ &= \frac{1}{4}\left[1 + e^{i(\delta-\gamma)}\right]\left[1 + e^{-i(\delta-\gamma)}\right] \\ &= \frac{1}{2}\left[1 + \cos(\delta - \gamma)\right]\end{aligned} \tag{1.28}$$

Agreement with (1.24) requires $\delta - \gamma = \pm\pi/2$. The common convention, which we will see in Chapter 3, is to take $\delta = 0$. If in (1.23) and (1.26) we ignore the overall phases δ_+ and γ_+, which appear in the amplitude (1.27) but do not enter into the calculation of the probability (1.28), we see that

$$|+\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle \tag{1.29}$$

and

$$|+\mathbf{y}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\pi/2}}{\sqrt{2}}|-\mathbf{z}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle \tag{1.30}$$

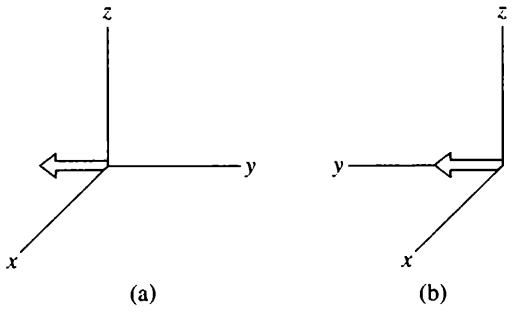


Figure 1.10 A state that is spin down along y in the right-handed coordinate system shown in (a) is spin up along y in the left-handed system shown in (b).

where we have chosen $\gamma = \pi/2$. The choice $\gamma = -\pi/2$ yields the state

$$\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle - \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle = |-\mathbf{y}\rangle \tag{1.31}$$

The reason for this ambiguity is that in discussing our series of Stern–Gerlach experiments we have not specified whether our coordinate system is right handed or left handed. The state we have called $|+\mathbf{y}\rangle$ is indeed the state with $S_y = \hbar/2$ in a right-handed coordinate system. The state we have called $|-\mathbf{y}\rangle$ is the state with $S_y = -\hbar/2$ in our right-handed coordinate system. Of course, this latter state, which is spin down along y, is spin up along y in a left-handed coordinate system, as shown in Fig. 1.10. That is why we see both solutions appearing.[10]

These complications should not detract from the main message to be learned from Experiment 5. The simple fact is that (1.24) cannot be explained without a complex amplitude. The appearance of i's such as the one in (1.30) is one of the key ingredients of a description of nature by quantum mechanics. Whereas in classical physics we often use complex numbers as an aid to do calculations, there they are not essential. The straightforward Stern–Gerlach experiments we have outlined in this chapter *demand* complex numbers for their explanation.

EXAMPLE 1.3 A spin-$\frac{1}{2}$ particle is in the state

$$|\psi\rangle = \frac{1}{2}|+\mathbf{z}\rangle + \frac{i\sqrt{3}}{2}|-\mathbf{z}\rangle$$

[10] We will see how to derive all of the results of this section from first principles in Chapter 3.

What is the probability that a measurement of S_y yields $\hbar/2$? What is $\langle S_y \rangle$ for this state?

SOLUTION From (1.30), we know that

$$|+\mathbf{y}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle$$

Thus the corresponding bra vector is

$$\langle +\mathbf{y}| = \frac{1}{\sqrt{2}}\langle +\mathbf{z}| - \frac{i}{\sqrt{2}}\langle -\mathbf{z}|$$

The probability amplitude for finding a particle in the state $|\psi\rangle$ with $S_y = \hbar/2$ if a measurement of S_y is carried out is given by

$$\langle +\mathbf{y}|\psi\rangle = \left(\frac{1}{\sqrt{2}}\langle +\mathbf{z}| - \frac{i}{\sqrt{2}}\langle -\mathbf{z}|\right)\left(\frac{1}{2}|+\mathbf{z}\rangle + \frac{i\sqrt{3}}{2}|-\mathbf{z}\rangle\right)$$

$$= \left(\frac{1}{\sqrt{2}}\right)\left(\frac{1}{2}\right) + \left(\frac{-i}{\sqrt{2}}\right)\left(\frac{i\sqrt{3}}{2}\right) = \frac{1}{2\sqrt{2}}\left(1+\sqrt{3}\right)$$

Therefore the probability is given by

$$|\langle +\mathbf{y}|\psi\rangle|^2 = \frac{1}{2} + \frac{\sqrt{3}}{4} = 0.93$$

To get a physical feel for what the spin state $|\psi\rangle$ is and why the probability of finding the particle in this state with $S_y = \hbar/2$ is as large as 0.93, take a look at Problem 1.10.

Since a measurement of S_y yields either $+\hbar/2$ or $-\hbar/2$, the probability of obtaining $S_y = -\hbar/2$ is given by

$$|\langle -\mathbf{y}|\psi\rangle|^2 = 1 - |\langle +\mathbf{y}|\psi\rangle|^2 = \frac{1}{2} - \frac{\sqrt{3}}{4} = 0.07$$

Therefore

$$\langle S_y \rangle = \left(\frac{1}{2} + \frac{\sqrt{3}}{4}\right)\left(\frac{\hbar}{2}\right) + \left(\frac{1}{2} - \frac{\sqrt{3}}{4}\right)\left(-\frac{\hbar}{2}\right) = \frac{\sqrt{3}}{4}\hbar$$

1.6 Summary

The world of quantum mechanics is both strange and wonderful, in part because it is a world filled with surprises that so often run counter to our classical expectations. Yet as we go on, we will see the remarkable insight quantum mechanics gives us

not just into microscopic phenomena but into the laws of classical mechanics as well. Since quantum mechanics subsumes classical mechanics, we cannot "derive" quantum mechanics from our classical, macroscopic experiences. Our strategy in this chapter has been to take a number of Stern–Gerlach experiments as our guide into this strange world of quantum behavior. From these experiments we can see many of the general features of quantum mechanics.

A quantum state is specified either by a ket vector $|\psi\rangle$ or a corresponding bra vector $\langle\psi|$. The complex numbers that we calculate in quantum mechanics result from a ket vector $|\psi\rangle$ meeting up with a bra vector $\langle\varphi|$, forming the bra(c)ket $\langle\varphi|\psi\rangle$, which we call the probability amplitude for a particle in the state $|\psi\rangle$ to be found in the state $|\varphi\rangle$. The amplitude $\langle\psi|\varphi\rangle$ for a particle in the state $|\varphi\rangle$ to be found in the state $|\psi\rangle$ is the complex conjugate of the amplitude for a particle in the state $|\psi\rangle$ to be found in the state $|\varphi\rangle$:

$$\langle\psi|\varphi\rangle = \langle\varphi|\psi\rangle^* \tag{1.32}$$

The probability of finding a particle to be in the state $|\varphi\rangle$ when a measurement is made on a particle in the state $|\psi\rangle$ is given by $|\langle\varphi|\psi\rangle|^2$. Notice that the probability is unchanged if the ket $|\psi\rangle$ is multiplied by an **overall phase factor** $e^{i\delta}$: $|\psi\rangle \rightarrow e^{i\delta}|\psi\rangle$.

Although we have phrased our discussion so far solely in terms of the intrinsic spin angular momentum of a spin-$\frac{1}{2}$ particle, the structure that we see emerging has a broad level of applicability. Suppose that we are considering an **observable** A for which the results of a measurement take on the discrete values $a_1, a_2, a_3, \ldots$.[11] As we will see, angular momentum and energy are good examples of observables for which the results of measurements can be grouped in a discrete (although not necessarily finite) set. A general quantum state, expressed in the form of a ket vector $|\psi\rangle$, can be written as a superposition of the states $|a_1\rangle$, $|a_2\rangle$, $|a_3\rangle$, ... that result if a measurement of A yields $a_1, a_2, a_3, \ldots$, respectively:

$$|\psi\rangle = c_1|a_1\rangle + c_2|a_2\rangle + c_3|a_3\rangle + \cdots = \sum_n c_n|a_n\rangle \tag{1.33}$$

The corresponding bra vector is given by

$$\langle\psi| = c_1^*\langle a_1| + c_2^*\langle a_2| + c_3^*\langle a_3| + \cdots = \sum_n c_n^*\langle a_n| \tag{1.34}$$

The complex number

$$c_n = \langle a_n|\psi\rangle \tag{1.35}$$

[11] The extension to observables such as position and momentum where the values form a continuum is discussed in Chapter 6.

is the amplitude to obtain a_n if a measurement of A is made for a particle in the state $|\psi\rangle$.[12]

Physically, we expect that

$$\langle a_i|a_j\rangle = 0 \quad i \neq j \tag{1.36}$$

since if the particle is in a state for which the result of a measurement is a_j, there is zero amplitude of obtaining a_i with $i \neq j$. The vectors $|a_i\rangle$ and $|a_j\rangle$ with $i \neq j$ are said to be **orthogonal**. The amplitude to obtain a_i for a particle in the state $|a_i\rangle$ is taken to be one, that is,

$$\langle a_i|a_i\rangle = 1 \tag{1.37}$$

The vector $|a_i\rangle$ is then said to be normalized. Equations (1.36) and (1.37) can be nicely summarized by

$$\langle a_i|a_j\rangle = \delta_{ij} \tag{1.38}$$

where δ_{ij} is called the **Kronecker delta** defined by the relationship

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{1.39}$$

We say that the set of vectors $|a_i\rangle$ form an **orthonormal set** of basis vectors. Equation (1.33) shows how an arbitrary vector $|\psi\rangle$ can be expressed in terms of this basis set. Thus the vectors $|a_i\rangle$ form a complete set.

Amplitudes such as (1.35) can be projected out of the ket $|\psi\rangle$ by taking the inner product of the ket $|\psi\rangle$ with the bra $\langle a_i|$:

$$\begin{aligned} \langle a_i|\psi\rangle &= \sum_n c_n \langle a_i|a_n\rangle \\ &= \sum_n c_n \delta_{in} = c_i \end{aligned} \tag{1.40}$$

Thus the ket (1.33) can be written

$$|\psi\rangle = \sum_n |a_n\rangle\langle a_n|\psi\rangle \tag{1.41}$$

which is just a sum of ket vectors $|a_i\rangle$, each multiplied by the amplitude $\langle a_i|\psi\rangle$.

---

[12] In this chapter we have used the shorthand notation $|S_z = \pm\hbar/2\rangle = |\pm\mathbf{z}\rangle$, $|S_x = \pm\hbar/2\rangle = |\pm\mathbf{x}\rangle$, and so on. Thus $\langle\pm\mathbf{z}|\psi\rangle$ are the amplitudes to obtain $S_z = \pm\hbar/2$ for a spin-$\frac{1}{2}$ particle in the state $|\psi\rangle$ if a measurement of S_z is made.

Similarly, the amplitude c_i^* can be projected out of the bra $\langle\psi|$ by taking the inner product with ket $|a_i\rangle$:

$$\begin{aligned}\langle\psi|a_i\rangle &= \sum_n c_n^*\langle a_n|a_i\rangle \\ &= \sum_n c_n^*\delta_{ni} = c_i^*\end{aligned} \tag{1.42}$$

The bra (1.34) can thus be written as

$$\langle\psi| = \sum_n \langle\psi|a_n\rangle\langle a_n| \tag{1.43}$$

which is the sum of the bra vectors $\langle a_i|$, each multiplied by the amplitude $\langle\psi|a_i\rangle$.

The normalization requirement

$$\langle\psi|\psi\rangle = 1 \tag{1.44}$$

for a physical state $|\psi\rangle$ leads to

$$\begin{aligned}1 = \langle\psi|\psi\rangle &= \left(\sum_i c_i^*\langle a_i|\right)\left(\sum_j c_j|a_j\rangle\right) \\ &= \sum_i\sum_j c_i^* c_j\langle a_i|a_j\rangle \\ &= \sum_i\sum_j c_i^* c_j\delta_{ij} = \sum_i |c_i|^2\end{aligned} \tag{1.45}$$

showing that the probabilities

$$|c_i|^2 = |\langle a_i|\psi\rangle|^2 \tag{1.46}$$

of obtaining the result a_i if a measurement of A is carried out sum to one. From these results it follows that the average value of the observable A for a particle in the state $|\psi\rangle$ is given by

$$\langle A\rangle = \sum_n |c_n|^2 a_n \tag{1.47}$$

since the average value (expectation value) is the sum of the values obtained by the measurements weighted by the probabilities of obtaining those values. The uncertainty is given by

$$\Delta A = \sqrt{\langle(A - \langle A\rangle)^2\rangle} = \sqrt{\langle A^2\rangle - \langle A\rangle^2} \tag{1.48}$$

where

$$\langle A^2\rangle = \sum_n |c_n|^2 a_n^2 \tag{1.49}$$

Equations (1.47) and (1.49) illustrate the importance of **completeness**, that is, that any state can be expressed as a superposition of basis vectors, as in (1.33). Without this completeness, we would not know how to calculate the results of measurements for the observable A for an arbitrary state.

One of the most striking features of the physical world is that if more than one of the c_n in (1.33) is nonzero, then there are amplitudes to obtain different a_n for a particle in a particular state $|\psi\rangle$. How should we interpret this result: Is the ket (1.33) telling us that the particle spends time in each of the states $|a_n\rangle$, and the probability $|\langle a_n|\psi\rangle|^2$ is just a reflection of how much time it spends in that particular state? Does this specification of the state as a superposition just reflect our lack of knowledge of which state the particle is really in? Is this why we must deal with probabilities? The answer to these questions is an emphatic no. Rather, (1.33) is to be read as a true superposition of the individual states $|a_n\rangle$, for if we parametrize the complex amplitudes in the form

$$\langle a_n|\psi\rangle = |\langle a_n|\psi\rangle|e^{i\delta_n} \tag{1.50}$$

where $|\langle a_n|\psi\rangle|$ is the magnitude, or modulus, of the amplitude and δ_n is the phase of the amplitude, the difference in phase (the relative phase) between the individual states in the superposition matters a great deal. As we have seen in our discussion of the spin-$\frac{1}{2}$ $|+\mathbf{x}\rangle$ and $|+\mathbf{y}\rangle$ kets, changing the relative phase between the kets $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ in such a superposition by $\pi/2$ changes a state with $S_x = \hbar/2$ into one with $S_y = \hbar/2$. Compare (1.29) and (1.30).[13] Thus the values of the relative phases in (1.33) dramatically affect how the states "add up," or how the amplitudes interfere with each other. Quantum mechanics is more than just a collection of probabilities. We live in a world in which the allowed states of a particle include *superpositions* of the states in which the particle possesses a definite attribute, such as the z component of the particle's spin angular momentum, and thus by superposing such states we form states for which the particle does not have definite value at all for such an attribute.

Problems

1.1. Determine the field gradient of a 50-cm-long Stern–Gerlach magnet that would produce a 1-mm separation at the detector between spin-up and spin-down silver atoms that are emitted from an oven at $T = 1500$ K. Assume the detector (see Fig. 1.1) is located 50 cm from the magnet. *Note:* While the atoms in the oven have average kinetic energy $3k_BT/2$, the more energetic atoms strike the hole in the oven more frequently. Thus the *emitted* atoms have average kinetic energy $2k_BT$, where

[13] This also shows that a spin-$\frac{1}{2}$ particle cannot have simultaneously a definite value for the x and y components of its intrinsic spin angular momentum.

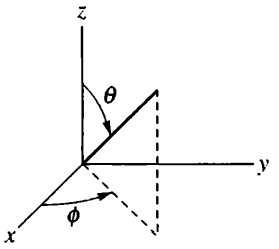


Figure 1.11 The angles θ and ϕ specifying the orientation of an SGn device.

k_B is the Boltzmann constant. The magnetic dipole moment of the silver atom is due to the intrinsic spin of the single electron. Appendix F gives the numerical value of the Bohr magneton, $e\hbar/2m_e c$, in a convenient form.

1.2. Show for a solid spherical ball of mass m rotating about an axis through its center with a charge q uniformly distributed on the surface of the ball that the magnetic moment $\boldsymbol{\mu}$ is related to the angular momentum $\mathbf{L}$ by the relation

$$\boldsymbol{\mu} = \frac{5q}{6mc}\mathbf{L}$$

Reminder: The factor of c is a consequence of our using Gaussian units. If you work in SI units, just add the c in by hand to compare with this result.

1.3. In Problem 3.2 we will see that the state of a spin-$\frac{1}{2}$ particle that is spin up along the axis whose direction is specified by the unit vector

$$\mathbf{n} = \sin\theta\cos\phi\mathbf{i} + \sin\theta\sin\phi\mathbf{j} + \cos\theta\mathbf{k}$$

with θ and ϕ shown in Fig. 1.11, is given by

$$|+\mathbf{n}\rangle = \cos\frac{\theta}{2}|+\mathbf{z}\rangle + e^{i\phi}\sin\frac{\theta}{2}|-\mathbf{z}\rangle$$

(a) Verify that the state $|+\mathbf{n}\rangle$ reduces to the states $|+\mathbf{x}\rangle$ and $|+\mathbf{y}\rangle$ given in this chapter for the appropriate choice of the angles θ and ϕ.

(b) Suppose that a measurement of S_z is carried out on a particle in the state $|+\mathbf{n}\rangle$. What is the probability that the measurement yields (i) $\hbar/2$? (ii) $-\hbar/2$?

(c) Determine the uncertainty ΔS_z of your measurements.

1.4. Repeat the calculations of Problem 1.3 (b) and (c) for measurements of S_x. *Hint:* Infer what the probability of obtaining $-\hbar/2$ for S_x is from the probability of obtaining $\hbar/2$.

1.5.

(a) What is the amplitude to find a particle that is in the state $|+\mathbf{n}\rangle$ (from Problem 1.3) with $S_y = \hbar/2$? What is the probability? Check your result by evaluating the probability for an appropriate choice of the angles θ and ϕ.

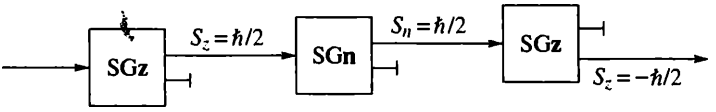


Figure 1.12 A Stern–Gerlach experiment with spin-$\frac{1}{2}$ particles.

(b) What is the amplitude to find a particle that is in the state $|+\mathbf{y}\rangle$ with $S_n = \hbar/2$? What is the probability?

1.6. Show that the state

$$|-\mathbf{n}\rangle = \sin\frac{\theta}{2}|+\mathbf{z}\rangle - e^{i\phi}\cos\frac{\theta}{2}|-\mathbf{z}\rangle$$

satisfies $\langle+\mathbf{n}|-\mathbf{n}\rangle = 0$, where the state $|+\mathbf{n}\rangle$ is given in Problem 1.3. Verify that $\langle-\mathbf{n}|-\mathbf{n}\rangle = 1$.

1.7. A beam of spin-$\frac{1}{2}$ particles is sent through a series of three Stern–Gerlach measuring devices, as illustrated in Fig. 1.12. The first SG**z** device transmits particles with $S_z = \hbar/2$ and filters out particles with $S_z = -\hbar/2$. The second device, an SG**n** device, transmits particles with $S_n = \hbar/2$ and filters out particles with $S_n = -\hbar/2$, where the axis **n** makes an angle θ in the x-z plane with respect to the z axis. Thus particles after passage through this SG**n** device are in the state $|+\mathbf{n}\rangle$ given in Problem 1.3 with the angle $\phi = 0$. A last SG**z** device transmits particles with $S_z = -\hbar/2$ and filters out particles with $S_z = \hbar/2$.

(a) What fraction of the particles transmitted by the first SG**z** device will survive the third measurement?

(b) How must the angle θ of the SG**n** device be oriented so as to maximize the number of particles that are transmitted by the final SG**z** device? What fraction of the particles survive the third measurement for this value of θ?

(c) What fraction of the particles survive the last measurement if the SG**n** device is simply removed from the experiment?

1.8. The state of a spin-$\frac{1}{2}$ particle is given by

$$|\psi\rangle = \frac{i}{\sqrt{3}}|+\mathbf{z}\rangle + \sqrt{\frac{2}{3}}|-\mathbf{z}\rangle$$

What are $\langle S_z\rangle$ and ΔS_z for this state? Suppose that an experiment is carried out on 100 particles, each of which is in this state. Make up a reasonable set of data for S_z that could result from such an experiment. What if the measurements were carried out on 1,000 particles? What about 10,000?

1.9. Verify that $\Delta S_x = \sqrt{\langle S_x^2\rangle - \langle S_x\rangle^2} = 0$ for the state $|+\mathbf{x}\rangle$.

1.10. The state

$$|\psi\rangle = \frac{1}{2}|+\mathbf{z}\rangle + \frac{i\sqrt{3}}{2}|-\mathbf{z}\rangle$$

is a state with $S_n = \hbar/2$ along a particular axis $\mathbf{n}$. Compare the state $|\psi\rangle$ with the state $|+\mathbf{n}\rangle$ in Problem 1.3 to find $\mathbf{n}$. Determine $\langle S_x\rangle$, $\langle S_y\rangle$, and $\langle S_z\rangle$ for this state. *Note:* $\langle S_z\rangle$ and $\langle S_y\rangle$ for this state are given in Example 1.2 and Example 1.3, respectively.

1.11. Calculate $\langle S_x\rangle$, $\langle S_y\rangle$, and $\langle S_z\rangle$ for the state

$$|\psi\rangle = -\frac{i}{2}|+\mathbf{z}\rangle + \frac{\sqrt{3}}{2}|-\mathbf{z}\rangle$$

Compare your results with those from Problem 1.10. What can you conclude about these two states?

1.12. The state

$$|\psi\rangle = \frac{1}{2}|+\mathbf{z}\rangle + \frac{\sqrt{3}}{2}|-\mathbf{z}\rangle$$

is similar to the one given in Problem 1.10. It is just "missing" the i. By comparing the state with the state $|+\mathbf{n}\rangle$ given in Problem 1.3, determine along which direction $\mathbf{n}$ the state is spin up. Calculate $\langle S_x\rangle$, $\langle S_y\rangle$, and $\langle S_z\rangle$ for the state $|\psi\rangle$. Compare your results with those of Problem 1.10.

1.13. Show that neither the probability of obtaining the result a_i nor the expectation value $\langle A\rangle$ is affected by $|\psi\rangle \rightarrow e^{i\delta}|\psi\rangle$, that is, by an overall phase change for the state $|\psi\rangle$.

1.14. It is known that there is a 36% probability of obtaining $S_z = \hbar/2$ and therefore a 64% chance of obtaining $S_z = -\hbar/2$ if a measurement of S_z is carried out on a spin-$\frac{1}{2}$ particle. In addition, it is known that the probability of finding the particle with $S_x = \hbar/2$, that is in the state $|+\mathbf{x}\rangle$, is 50%. Determine the state of the particle as completely as possible from this information.

1.15. It is known that there is a 90% probability of obtaining $S_z = \hbar/2$ if a measurement of S_z is carried out on a spin-$\frac{1}{2}$ particle. In addition, it is known that there is a 20% probability of obtaining $S_y = \hbar/2$ if a measurement of S_y is carried out. Determine the spin state of the particle as completely as possible from this information. What is the probability of obtaining $S_x = \hbar/2$ if a measurement of S_x is carried out?

CHAPTER 2

Rotation of Basis States and Matrix Mechanics

In this chapter we will see that transforming a vector into a different vector in our quantum mechanical vector space requires an operator. We will also introduce a convenient shorthand notation in which we represent ket vectors by column vectors, bra vectors by row vectors, and operators by matrices. Our discussion will be primarily phrased in terms of the two-state spin-$\frac{1}{2}$ system introduced in Chapter 1, but we will also analyze another two-state system, the polarization of the electromagnetic field.

2.1 The Beginnings of Matrix Mechanics

REPRESENTING KETS AND BRAS

We have seen that we can express an arbitrary spin state $|\psi\rangle$ of a spin-$\frac{1}{2}$ particle as

$$|\psi\rangle = |+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle = c_+|+\mathbf{z}\rangle + c_-|-\mathbf{z}\rangle \tag{2.1}$$

Such a spin state may, for example, be created by sending spin-$\frac{1}{2}$ particles through a Stern–Gerlach device with its magnetic field gradient oriented in some arbitrary direction. The complex numbers $c_\pm = \langle\pm\mathbf{z}|\psi\rangle$ tell us how our state $|\psi\rangle$ is oriented in our quantum mechanical vector space, that is, how much of $|\psi\rangle$ is projected onto each of the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$.

One convenient way of representing $|\psi\rangle$ is just to keep track of these complex numbers. Just as we can avoid unit vectors in writing the classical electric field

$$\mathbf{E} = E_x\mathbf{i} + E_y\mathbf{j} + E_z\mathbf{k} \tag{2.2a}$$

by using the notation

$$\mathbf{E} \rightarrow (E_x, E_y, E_z) \tag{2.2b}$$

we can represent the ket (2.1) by the column vector

$$|\psi\rangle \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle +\mathbf{z}|\psi\rangle \\ \langle -\mathbf{z}|\psi\rangle \end{pmatrix} = \begin{pmatrix} c_+ \\ c_- \end{pmatrix} \tag{2.3}$$

In this basis, the ket $|+\mathbf{z}\rangle$ is represented by the column vector

$$|+\mathbf{z}\rangle \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle +\mathbf{z}|+\mathbf{z}\rangle \\ \langle -\mathbf{z}|+\mathbf{z}\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.4}$$

and the ket $|-\mathbf{z}\rangle$ is represented by the column vector

$$|-\mathbf{z}\rangle \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle +\mathbf{z}|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|-\mathbf{z}\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.5}$$

although the label under the arrow is really superfluous in (2.4) and (2.5) given the form of the column vectors on the right. Using (1.29), we can also write, for example,

$$|+\mathbf{x}\rangle \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle +\mathbf{z}|+\mathbf{x}\rangle \\ \langle -\mathbf{z}|+\mathbf{x}\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{2.6}$$

How do we represent bra vectors? We know that the bra vector corresponding to the ket vector (2.1) is

$$\langle\psi| = \langle\psi|+\mathbf{z}\rangle\langle +\mathbf{z}| + \langle\psi|-\mathbf{z}\rangle\langle -\mathbf{z}| = c_+^*\langle +\mathbf{z}| + c_-^*\langle -\mathbf{z}| \tag{2.7}$$

We can express

$$\langle\psi|\psi\rangle = \langle\psi|+\mathbf{z}\rangle\langle +\mathbf{z}|\psi\rangle + \langle\psi|-\mathbf{z}\rangle\langle -\mathbf{z}|\psi\rangle = 1 \tag{2.8}$$

conveniently as

$$\langle\psi|\psi\rangle = \underbrace{(\langle\psi|+\mathbf{z}\rangle,\ \langle\psi|-\mathbf{z}\rangle)}_{\text{bra vector}} \underbrace{\begin{pmatrix} \langle +\mathbf{z}|\psi\rangle \\ \langle -\mathbf{z}|\psi\rangle \end{pmatrix}}_{\text{ket vector}} = 1 \tag{2.9}$$

where we are using the usual rules of matrix multiplication for row and column vectors. This suggests that we represent the bra $\langle\psi|$ by the row vector

$$\langle\psi| \xrightarrow[S_z \text{ basis}]{} (\langle\psi|+\mathbf{z}\rangle,\ \langle\psi|-\mathbf{z}\rangle) \tag{2.10}$$

Since $\langle\psi|+\mathbf{z}\rangle = \langle +\mathbf{z}|\psi\rangle^*$ and $\langle\psi|-\mathbf{z}\rangle = \langle -\mathbf{z}|\psi\rangle^*$, (2.10) can also be expressed as

$$\langle\psi| \xrightarrow[S_z \text{ basis}]{} (\langle +\mathbf{z}|\psi\rangle^*,\ \langle -\mathbf{z}|\psi\rangle^*) = (c_+^*,\ c_-^*) \tag{2.11}$$

Comparing (2.11) with (2.3), we see that the row vector that *represents* the bra is the complex conjugate and transpose of the column vector that *represents* the corresponding ket. In this representation, an inner product such as (2.9) is carried out using the usual rules of matrix multiplication.

As an example, we may determine the representation for the ket $|-\mathbf{x}\rangle$ in the S_z basis. We know from the Stern–Gerlach experiments that there is zero amplitude to obtain $S_x = -\hbar/2$ for a state with $S_x = \hbar/2$, that is, $\langle -\mathbf{x}|+\mathbf{x}\rangle = 0$. Making the amplitude $\langle -\mathbf{x}|+\mathbf{x}\rangle$ vanish requires that

$$|-\mathbf{x}\rangle \xrightarrow[S_z \text{ basis}]{} \frac{e^{i\delta}}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{2.12}$$

since then

$$\langle -\mathbf{x}|+\mathbf{x}\rangle = \frac{e^{-i\delta}}{\sqrt{2}}(1, -1)\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0 \tag{2.13}$$

Note that the $1/\sqrt{2}$ in front of the column vector in (2.12) has been chosen so that the ket $|-\mathbf{x}\rangle$ is properly normalized:

$$\langle -\mathbf{x}|-\mathbf{x}\rangle = \frac{e^{-i\delta}}{\sqrt{2}}(1, -1)\frac{e^{i\delta}}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 \tag{2.14}$$

The common convention, and the one that we will generally follow, is to choose the overall phase $\delta = 0$ so that

$$|-\mathbf{x}\rangle \xrightarrow[S_z \text{ basis}]{} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{2.15}$$

However, in Section 2.5 we will see that an interesting case can be made for choosing $\delta = \pi$.

As another example, (1.30) indicates that the state with $S_y = \hbar/2$ is

$$|+\mathbf{y}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle \tag{2.16}$$

which may be represented in the S_z basis by

$$|+\mathbf{y}\rangle \rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \tag{2.17a}$$

The bra corresponding to this ket is represented in the same basis by

$$\langle +\mathbf{y}| \rightarrow \frac{1}{\sqrt{2}} (1, -i) \tag{2.17b}$$

Note the appearance of the $-i$ in this representation for the bra vector. Using these representations, we can check that

$$\langle +\mathbf{y}|+\mathbf{y}\rangle = \frac{1}{\sqrt{2}}(1, -i)\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = 1 \tag{2.18}$$

If we had used the row vector

$$\frac{1}{\sqrt{2}}(1, +i)$$

in evaluating the inner product, we would have obtained zero instead of one. Since $\langle -\mathbf{y}|+\mathbf{y}\rangle = 0$, this tells us that in the S_z basis

$$\langle -\mathbf{y}| \rightarrow \frac{1}{\sqrt{2}}(1, +i) \tag{2.19a}$$

and thus

$$|-\mathbf{y}\rangle \rightarrow \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} \tag{2.19b}$$

Putting these pieces together, we can use these matrix representations to calculate the probability that a spin-$\frac{1}{2}$ particle with $S_x = \hbar/2$ is found to have $S_y = \hbar/2$ when a measurement is carried out:

$$\begin{aligned} |\langle +\mathbf{y}|+\mathbf{x}\rangle|^2 &= \left|\frac{1}{\sqrt{2}}(1, -i)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right|^2 \\ &= \left|\frac{1-i}{2}\right|^2 = \frac{(1-i)}{2}\frac{(1+i)}{2} = \frac{1}{2} \end{aligned} \tag{2.20}$$

EXAMPLE 2.1 Use matrix mechanics to determine the probability that a measurement of S_y yields $\hbar/2$ for a spin-$\frac{1}{2}$ particle in the state

$$|\psi\rangle = \frac{1}{2}|+\mathbf{z}\rangle + \frac{i\sqrt{3}}{2}|-\mathbf{z}\rangle$$

SOLUTION

$$\begin{aligned} |\langle +\mathbf{y}|\psi\rangle|^2 &= \left|\frac{1}{\sqrt{2}}(1, -i)\frac{1}{2}\begin{pmatrix} 1 \\ i\sqrt{3} \end{pmatrix}\right|^2 \\ &= \left|\frac{1}{2\sqrt{2}}(1+\sqrt{3})\right|^2 = \frac{1}{8}(4+2\sqrt{3}) = \frac{1}{2} + \frac{\sqrt{3}}{4} \end{aligned}$$

Compare this relatively compact derivation with the use of kets and bras in Example 1.3.

FREEDOM OF REPRESENTATION

It is often convenient to use a number of different basis sets to express a particular state $|\psi\rangle$. Just as we can write the electric field in a particular coordinate system as

(2.2), we could use a different coordinate system with unit vectors $\mathbf{i}'$, $\mathbf{j}'$, and $\mathbf{k}'$ to write the same electric field as

$$\mathbf{E} = E_{x'}\mathbf{i}' + E_{y'}\mathbf{j}' + E_{z'}\mathbf{k}' \tag{2.21a}$$

or

$$\mathbf{E} \rightarrow (E_{x'}, E_{y'}, E_{z'}) \tag{2.21b}$$

Of course, the electric field $\mathbf{E}$ hasn't changed. It still has the same magnitude and direction, but we have chosen a different set of unit vectors, or basis vectors, to express it. Similarly, we can take the quantum state $|\psi\rangle$ in (2.1) and write it in terms of the basis states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ as

$$|\psi\rangle = |+\mathbf{x}\rangle\langle+\mathbf{x}|\psi\rangle + |-\mathbf{x}\rangle\langle-\mathbf{x}|\psi\rangle \tag{2.22}$$

which expresses the state as a superposition of the states with $S_x = \pm\hbar/2$ multiplied by the amplitudes for the particle to be found in these states. We can then construct a column vector representing $|\psi\rangle$ in this basis using these amplitudes:

$$|\psi\rangle \xrightarrow[S_x \text{ basis}]{} \begin{pmatrix} \langle+\mathbf{x}|\psi\rangle \\ \langle-\mathbf{x}|\psi\rangle \end{pmatrix} \tag{2.23}$$

Thus the column vector representing the ket $|+\mathbf{x}\rangle$ is

$$|+\mathbf{x}\rangle \xrightarrow[S_x \text{ basis}]{} \begin{pmatrix} \langle+\mathbf{x}|+\mathbf{x}\rangle \\ \langle-\mathbf{x}|+\mathbf{x}\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.24}$$

which is to be compared with the column vector (2.6). The ket $|+\mathbf{x}\rangle$ is the same state in the two cases; we have just written it out using the S_z basis in the first case and the S_x basis in the second case. Which basis we use is determined by what is convenient, such as what measurements we are going to perform on the state $|+\mathbf{x}\rangle$.

2.2 Rotation Operators

There is a nice physical way to transform the kets themselves from one basis set to another.[1] Recall that within classical physics a magnetic moment placed in a uniform magnetic field precesses about the direction of the field. When we discuss time evolution in Chapter 4, we will see that the interaction of the magnetic moment of a spin-$\frac{1}{2}$ particle with the magnetic field also causes the quantum spin state of the particle to rotate about the direction of the field as time progresses. In particular, if

[1] You may object to calling anything dealing directly with kets *physical* since ket vectors are abstract vectors specifying the quantum state of the system and involve, as we have seen, complex numbers.

the magnetic field points in the y direction and the particle is initially in the state $|+\mathbf{z}\rangle$, the spin will rotate in the x-z plane. At some later time the particle will be in the state $|+\mathbf{x}\rangle$. With this example in mind, it is useful at this stage to introduce a **rotation operator** $\hat{R}(\frac{\pi}{2}\mathbf{j})$ that acts on the ket $|+\mathbf{z}\rangle$, a state that is spin up along the z axis, and transforms it into the ket $|+\mathbf{x}\rangle$, a state that is spin up along the x axis:

$$|+\mathbf{x}\rangle = \hat{R}(\tfrac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle \tag{2.25}$$

Changing or transforming a ket in our vector space into a different ket requires an operator. To distinguish operators from ordinary numbers, we denote all operators with a hat.

What is the nature of the transformation effected by the operator $\hat{R}(\frac{\pi}{2}\mathbf{j})$? This operator just rotates the ket $|+\mathbf{z}\rangle$ by $\pi/2$ radians, or 90°, about the y axis (indicated by the unit vector $\mathbf{j}$) in a counterclockwise direction as viewed from the positive y axis, turning, or rotating, it into the ket $|+\mathbf{x}\rangle$, as indicated in Fig. 2.1a. The same rotation operator should rotate $|-\mathbf{z}\rangle$ into $|-\mathbf{x}\rangle$. In fact, since the most general state of a spin-$\frac{1}{2}$ particle may be expressed in the form of (2.1), the operator rotates this ket as well:

$$\begin{aligned}\hat{R}(\tfrac{\pi}{2}\mathbf{j})|\psi\rangle &= \hat{R}(\tfrac{\pi}{2}\mathbf{j})\left(c_+|+\mathbf{z}\rangle + c_-|-\mathbf{z}\rangle\right)\\ &= c_+\hat{R}(\tfrac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle + c_-\hat{R}(\tfrac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle\\ &= c_+|+\mathbf{x}\rangle + c_-|-\mathbf{x}\rangle\end{aligned} \tag{2.26}$$

Note that the operator acts on kets, not on the complex numbers.[2]

THE ADJOINT OPERATOR

What is the bra equation corresponding to the ket equation (2.25)? You may be tempted to guess that $\langle+\mathbf{x}| = \langle+\mathbf{z}|\hat{R}(\frac{\pi}{2}\mathbf{j})$, but we can quickly see that this *cannot* be correct, for if it were, we could calculate[3]

$$\langle+\mathbf{x}|+\mathbf{x}\rangle = \left[\langle+\mathbf{z}|\hat{R}(\tfrac{\pi}{2}\mathbf{j})\right]\left[\hat{R}(\tfrac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle\right] = \langle+\mathbf{z}|\hat{R}(\tfrac{\pi}{2}\mathbf{j})\hat{R}(\tfrac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle$$

We know that $\langle+\mathbf{x}|+\mathbf{x}\rangle = 1$, but since $\hat{R}(\frac{\pi}{2}\mathbf{j})$ rotates by 90° around the y axis, $\hat{R}(\frac{\pi}{2}\mathbf{j})\hat{R}(\frac{\pi}{2}\mathbf{j}) = \hat{R}(\pi\mathbf{j})$ performs a rotation of 180° about the y axis. But as indicated

---

[2] An operator $\hat{A}$ satisfying

$$\hat{A}(a|\psi\rangle + b|\varphi\rangle) = a\hat{A}|\psi\rangle + b\hat{A}|\varphi\rangle$$

where a and b are complex numbers, is referred to as a **linear operator**.

[3] You can see why we position the operator to the right of the bra vector when we go to calculate an amplitude. Otherwise we would evaluate the inner product and the operator would be left alone with no vector to act on.

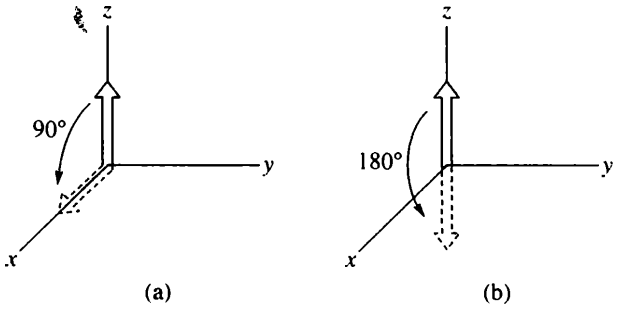


Figure 2.1 Rotating $|+\mathbf{z}\rangle$ counterclockwise about the y axis (a) by $\pi/2$ radians transforms the state into $|+\mathbf{x}\rangle$ and (b) by π radians transforms the state into $|-\mathbf{z}\rangle$. The spin state of a spin-$\frac{1}{2}$ particle with a magnetic moment would rotate in the x-z plane if the particle were placed in a magnetic field in the y direction.

in Fig. 2.1b, $\hat{R}(\pi\mathbf{j})|+\mathbf{z}\rangle = |-\mathbf{z}\rangle$, and since $\langle+\mathbf{z}|\hat{R}(\pi\mathbf{j})|+\mathbf{z}\rangle = \langle+\mathbf{z}|-\mathbf{z}\rangle = 0$, we are left with a contradiction.

For the ket vector $|\psi\rangle = c_+|+\mathbf{z}\rangle + c_-|-\mathbf{z}\rangle$, the corresponding bra vector is $\langle\psi| = c_+^*\langle+\mathbf{z}| + c_-^*\langle-\mathbf{z}|$, with the complex numbers in the ket turning into their complex conjugates in the bra. Since we are dealing here with operators and not just complex numbers, we need an additional rule for determining the bra equation corresponding to a ket equation like (2.25) that involves an operator. We introduce a new operator $\hat{R}^\dagger$, called the **adjoint operator** of the operator $\hat{R}$, so that the bra equation corresponding to (2.25) is

$$\langle+\mathbf{x}| = \langle+\mathbf{z}|\hat{R}^\dagger(\tfrac{\pi}{2}\mathbf{j}) \tag{2.27}$$

We can then satisfy

$$1 = \langle+\mathbf{x}|+\mathbf{x}\rangle = \langle+\mathbf{z}|\hat{R}^\dagger(\tfrac{\pi}{2}\mathbf{j})\hat{R}(\tfrac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle = \langle+\mathbf{z}|+\mathbf{z}\rangle \tag{2.28}$$

if the adjoint operator $\hat{R}^\dagger$ is inverse of the operator $\hat{R}$. In particular, the adjoint operator $\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})$ is a rotation operator that can be viewed as operating to the right on the ket $\hat{R}(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle$. If $\hat{R}(\frac{\pi}{2}\mathbf{j})$ rotates by 90° *counterclockwise*, then $\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})$ rotates by 90° *clockwise* so that $\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})\hat{R}(\frac{\pi}{2}\mathbf{j}) = 1$, and we are left with $\langle+\mathbf{z}|+\mathbf{z}\rangle = 1$.[4]

In general, an operator $\hat{U}$ satisfying $\hat{U}^\dagger\hat{U} = 1$ is called a **unitary operator**. Thus the rotation operator must be unitary in order that the amplitude for a state to be itself—that is, so that $\langle\psi|\psi\rangle = 1$—doesn't change under rotation. Otherwise, probability would not be conserved under rotation.

[4] As this example illustrates, the adjoint operator can act to the right on ket vectors as well to the left on bra vectors.

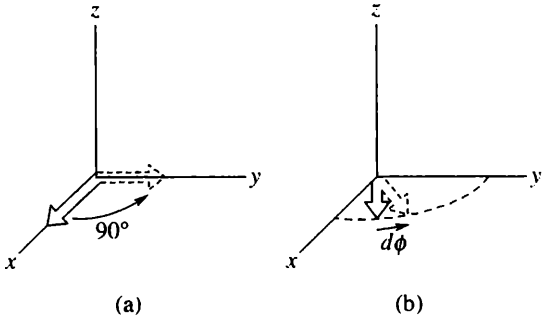


Figure 2.2 (a) Rotating $|+\mathbf{x}\rangle$ by $\pi/2$ radians counterclockwise about the z axis transforms the state into $|+\mathbf{y}\rangle$. (b) Rotation of a state by an infinitesimal angle $d\phi$ about the z axis.

THE GENERATOR OF ROTATIONS

Instead of performing rotations about the y axis, let's rotate about the z axis. If we rotate by 90° counterclockwise about the z axis, we will, for example, turn $|+\mathbf{x}\rangle$ into $|+\mathbf{y}\rangle$, as indicated in Fig. 2.2a. Instead of carrying out this whole rotation initially, let us first focus on an infinitesimal rotation by an angle $d\phi$ about the z axis, as shown in Fig. 2.2b. A useful way to express this infinitesimal rotation operator is in the form

$$\hat{R}(d\phi\mathbf{k}) = 1 - \frac{i}{\hbar}\hat{J}_z\, d\phi \tag{2.29}$$

where we have introduced an operator $\hat{J}_z$ that "generates" rotations about the z axis and moves us away from the identity element. Our form for $\hat{R}(d\phi\mathbf{k})$ clearly satisfies the requirement that $\hat{R}(d\phi\mathbf{k}) \to 1$ as $d\phi \to 0$. As we will see, the factor of i and the factor of $\hbar$ have been introduced to bring out the physical significance of the operator $\hat{J}_z$. In particular, because the factor of $\hbar$ occurs in the denominator of the second term in (2.29), the operator $\hat{J}_z$ must have the dimensions of $\hbar$, namely, the dimensions of angular momentum. We will see that a convincing case can be made that we should identify this operator $\hat{J}_z$, the generator of rotations about the z axis, with the z component of the intrinsic spin angular momentum of the particle.

We first establish that $\hat{J}_z$ belongs to a special class of operators known as Hermitian operators. Physically, the operator $\hat{R}^\dagger(d\phi\mathbf{k})$ is the inverse of the rotation operator $\hat{R}(d\phi\mathbf{k})$. By taking the adjoint of (2.29), we can write this operator in the form

$$\hat{R}^\dagger(d\phi\mathbf{k}) = 1 + \frac{i}{\hbar}\hat{J}_z^\dagger\, d\phi \tag{2.30}$$

where $\hat{J}_z^\dagger$ is the adjoint of the operator $\hat{J}_z$. Note that since the bra corresponding to the ket $c|\psi\rangle$ is $\langle\psi|c^*$, complex numbers get replaced by their complex conjugates when

forming the adjoint operator. Thus $i \to -i$ in going from (2.29) to (2.30), which has the same effect as changing $d\phi$ to $-d\phi$, and therefore $\hat{R}^\dagger(d\phi\mathbf{k}) = \hat{R}(-d\phi\mathbf{k})$, provided $\hat{J}_z^\dagger = \hat{J}_z$. More formally, since the rotation operator $\hat{R}^\dagger(d\phi\mathbf{k})$ is the inverse of the rotation operator $\hat{R}(d\phi\mathbf{k})$, these operators must satisfy the condition

$$\begin{aligned}\hat{R}^\dagger(d\phi\mathbf{k})\hat{R}(d\phi\mathbf{k}) &= \left(1 + \frac{i}{\hbar}\hat{J}_z^\dagger\, d\phi\right)\left(1 - \frac{i}{\hbar}\hat{J}_z\, d\phi\right) \\ &= 1 + \frac{i}{\hbar}\left(\hat{J}_z^\dagger - \hat{J}_z\right) d\phi + O(d\phi^2) = 1 \qquad (2.31)\end{aligned}$$

Since the angle $d\phi$ is infinitesimal, we can neglect the second-order terms in $d\phi$ and (2.31) will be satisfied only if $\hat{J}_z = \hat{J}_z^\dagger$. In general, an operator that is equal to its adjoint is called **self-adjoint**, or **Hermitian**. Thus $\hat{J}_z$ must be a Hermitian operator. Hermitian operators have a number of nice properties that permit them to play major roles in quantum mechanics. After some specific examples, we will discuss some of these general properties in Section 2.8.[5]

One of the reasons that infinitesimal rotations are useful is that once we know how to perform an infinitesimal rotation about the z axis by an angle $d\phi$, we can carry out a rotation by any finite angle ϕ by compounding an infinite number of infinitesimal rotations with

$$d\phi = \lim_{N\to\infty} \frac{\phi}{N}$$

The rotation operator $\hat{R}(\phi\mathbf{k})$ is then given by

$$\hat{R}(\phi\mathbf{k}) = \lim_{N\to\infty}\left[1 - \frac{i}{\hbar}\hat{J}_z\left(\frac{\phi}{N}\right)\right]^N = e^{-i\hat{J}_z\phi/\hbar} \qquad (2.32)$$

The last identity in (2.32) can be established by expanding both sides in a Taylor series and showing that they agree term by term (see Problem 2.1). In fact, a series expansion is really the only way to make sense of an expression such as an exponential of an operator.

EIGENSTATES AND EIGENVALUES

What happens to a ket $|+\mathbf{z}\rangle$ if we rotate it about the z axis—that is, what is $\hat{R}(\phi\mathbf{k})|+\mathbf{z}\rangle$? If you were to rotate a classical spinning top about its axis of rotation, it would still be in the same state with its angular momentum pointing in the same direction. Similarly, rotating a state of a spin-$\frac{1}{2}$ particle that is spin up along z about the z axis should still yield a state that is spin up along z, as illustrated in

[5] Now you can see one reason for introducing the i in the defining relation (2.29) for an infinitesimal rotation operator. Without it, the generator $\hat{J}_z$ would not have turned out to be Hermitian.

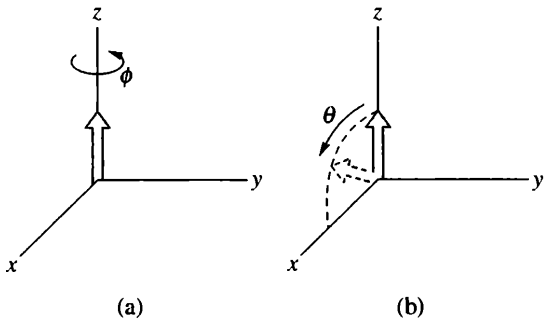


Figure 2.3 (a) Rotating $|+\mathbf{z}\rangle$ by angle ϕ about the z axis with the operator $\hat{R}(\phi\mathbf{k})$ does not change the state, in contrast to the action of the operator $\hat{R}(\theta\mathbf{j})$, which rotates $|+\mathbf{z}\rangle$ by angle θ about the y axis, producing a different state, as indicated in (b).

Fig. 2.3. In Chapter 1 we saw that the overall phase of a state does not enter into the calculation of probabilities, such as in (1.24). This turns out to be quite a general feature: two states that differ only by an *overall* phase are really the same state. We will now show that in order for $\hat{R}(\phi\mathbf{k})|+\mathbf{z}\rangle$ to differ from $|+\mathbf{z}\rangle$ only by an overall phase, it is necessary that

$$\hat{J}_z|+\mathbf{z}\rangle = (\text{constant})|+\mathbf{z}\rangle \tag{2.33}$$

In general, when an operator acting on a state yields a constant times the state, we call the state an **eigenstate** of the operator and the constant the corresponding **eigenvalue**.

First we will establish the eigenstate condition (2.33). If we expand the exponential in the rotation operator (2.32) in a Taylor series, we have

$$\hat{R}(\phi\mathbf{k})|+\mathbf{z}\rangle = \left[1 - \frac{i\phi\hat{J}_z}{\hbar} + \frac{1}{2!}\left(-\frac{i\phi\hat{J}_z}{\hbar}\right)^2 + \cdots\right]|+\mathbf{z}\rangle \tag{2.34}$$

If (2.33) is not satisfied and $\hat{J}_z|+\mathbf{z}\rangle$ is something other than a constant times $|+\mathbf{z}\rangle$, such as $|+\mathbf{x}\rangle$, the first two terms in the series will yield $|+\mathbf{z}\rangle$ plus a term involving $|+\mathbf{x}\rangle$, which would mean that $\hat{R}(\phi\mathbf{k})|+\mathbf{z}\rangle$ differs from $|+\mathbf{z}\rangle$ by other than a multiplicative constant. Note that other terms in the series cannot cancel this unwanted $|+\mathbf{x}\rangle$ term, since each term involving a different power of ϕ is linearly independent from the rest. Thus we deduce that the ket $|+\mathbf{z}\rangle$ must be an eigenstate, or **eigenket**, of the operator $\hat{J}_z$.

Let's now turn our attention to the value of the constant, the eigenvalue, in (2.33). We will give a self-consistency argument to show that we will have agreement with

the analysis of the Stern–Gerlach experiments in Chapter 1 provided

$$\hat{J}_z|\pm\mathbf{z}\rangle = \pm\frac{\hbar}{2}|\pm\mathbf{z}\rangle \tag{2.35}$$

This equation asserts that the eigenvalues for the spin-up and spin-down states are the values of S_z that these states are observed to have in the Stern–Gerlach experiments.[6] First consider the spin-up state. If

$$\hat{J}_z|+\mathbf{z}\rangle = \frac{\hbar}{2}|+\mathbf{z}\rangle \tag{2.36a}$$

then

$$\hat{J}_z^2|+\mathbf{z}\rangle = \hat{J}_z\frac{\hbar}{2}|+\mathbf{z}\rangle = \frac{\hbar}{2}\hat{J}_z|+\mathbf{z}\rangle = \left(\frac{\hbar}{2}\right)^2|+\mathbf{z}\rangle \tag{2.36b}$$

and so on. From (2.34), we obtain

$$\hat{R}(\phi\mathbf{k})|+\mathbf{z}\rangle = \left[1 - \frac{i\phi}{2} + \frac{1}{2!}\left(-\frac{i\phi}{2}\right)^2 + \cdots\right]|+\mathbf{z}\rangle = e^{-i\phi/2}|+\mathbf{z}\rangle \tag{2.37}$$

The state has picked up an overall phase, just as we would hope if the state is not to change. The value of the phase is determined by the eigenvalue in (2.36a).

In order to see why the eigenvalue should be $\hbar/2$, let's consider what happens if we rotate a spin-down state $|-\mathbf{z}\rangle$ about the z axis, that is, if we evaluate $\hat{R}(\phi\mathbf{k})|-\mathbf{z}\rangle$. Just as before, we can argue that $|-\mathbf{z}\rangle$ must be an eigenstate of $\hat{J}_z$. We can also argue that the eigenvalue for $|-\mathbf{z}\rangle$ must be different from that for $|+\mathbf{z}\rangle$. After all, if the eigenvalues were the same, applying the rotation operator $\hat{R}(\phi\mathbf{k})$ to the state

$$|+\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle \tag{2.38}$$

would not rotate the state, since $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ would each pick up the same phase factor, and the state in (2.38) would itself pick up just an overall phase. Therefore, it would still be the same state. But if we rotate the state $|+\mathbf{x}\rangle$ by an angle ϕ in the x-y plane, we expect the state to change. If we try

$$\hat{J}_z|-\mathbf{z}\rangle = -\frac{\hbar}{2}|-\mathbf{z}\rangle \tag{2.39}$$

for the eigenvalue equation for the spin-down state, we find

$$\hat{R}(\phi\mathbf{k})|-\mathbf{z}\rangle = \left[1 + \frac{i\phi}{2} + \frac{1}{2!}\left(\frac{i\phi}{2}\right)^2 + \cdots\right]|-\mathbf{z}\rangle = e^{i\phi/2}|-\mathbf{z}\rangle \tag{2.40}$$

---

[6] You can start to see why we introduced a factor of $1/\hbar$ in the defining relation (2.29) between the infinitesimal rotation operator and the generator of rotations.

Using (2.37) and (2.40), we see that

$$\begin{aligned}\hat{R}(\phi\mathbf{k})|+\mathbf{x}\rangle &= \frac{e^{-i\phi/2}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\phi/2}}{\sqrt{2}}|-\mathbf{z}\rangle \\ &= e^{-i\phi/2}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\phi}}{\sqrt{2}}|-\mathbf{z}\rangle\right) \end{aligned} \tag{2.41}$$

which is clearly a different state from (2.38) for $\phi \neq 0$. In particular, with the choice $\phi = \pi/2$, we obtain

$$\begin{aligned}\hat{R}(\tfrac{\pi}{2}\mathbf{k})|+\mathbf{x}\rangle &= e^{-i\pi/4}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\pi/2}}{\sqrt{2}}|-\mathbf{z}\rangle\right) \\ &= e^{-i\pi/4}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle\right) = e^{-i\pi/4}|+\mathbf{y}\rangle \end{aligned} \tag{2.42}$$

where we have replaced the term in the brackets by the state $|+\mathbf{y}\rangle$ that we determined in (1.30). Since two states that differ only by an overall phase are the same state, we see that rotating the state $|+\mathbf{x}\rangle$ by 90° counterclockwise about the z axis does generate the state $|+\mathbf{y}\rangle$ *when (2.35) holds.* Thus we are led to a striking conclusion: When the operator that generates rotations about the z axis acts on the spin-up-along-z and spin-down-along-z states, it throws out a constant (the eigenvalue) times the state (the eigenstate); the eigenvalues for the two states are just the values of the z component of the intrinsic spin angular momentum that characterize these states.

Finally, let us note something really perplexing about the effects of rotations on spin-$\frac{1}{2}$ particles: namely,

$$\hat{R}(2\pi\mathbf{k})|+\mathbf{z}\rangle = e^{-i\pi}|+\mathbf{z}\rangle = -|+\mathbf{z}\rangle \tag{2.43a}$$

and

$$\hat{R}(2\pi\mathbf{k})|+\mathbf{z}\rangle = e^{i\pi}|-\mathbf{z}\rangle = -|-\mathbf{z}\rangle \tag{2.43b}$$

Thus, if we rotate a spin-$\frac{1}{2}$ state by 360° and *end up right where we started*, we find that the state picks up an overall minus sign. Earlier we remarked that we could actually perform these rotations on our spin systems by inserting them in a magnetic field. When we come to time evolution in Chapter 4, we will see how this strange prediction (2.43) for spin-$\frac{1}{2}$ particles may be verified experimentally.

EXAMPLE 2.2 Show that rotating the spin-up-along-x state $|+\mathbf{x}\rangle$ by 180° about the z axis yields the spin-down-along-x state.

SOLUTION

$$\begin{aligned}\hat{R}(\pi\mathbf{k})|+\mathbf{x}\rangle &= \hat{R}(\pi\mathbf{k})\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle\right)\\ &= \frac{e^{-i\pi/2}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\pi/2}}{\sqrt{2}}|-\mathbf{z}\rangle\\ &= e^{-i\pi/2}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\pi}}{\sqrt{2}}|-\mathbf{z}\rangle\right)\\ &= e^{-i\pi/2}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle - \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle\right)\\ &= e^{-i\pi/2}|-\mathbf{x}\rangle\end{aligned}$$

where in the last line we have used the phase convention for the state $|-\mathbf{x}\rangle$ given in (2.15).

2.3 The Identity and Projection Operators

In general, the operator $\hat{R}(\theta\mathbf{n})$ changes a ket into a different ket by rotating it by an angle θ around the axis specified by the unit vector $\mathbf{n}$. Most operators tend to *do* something when they act on ket vectors, but it is convenient to introduce an operator that acts on a ket vector and does nothing: the **identity operator**. Surprisingly, we will see that this operator is a powerful operator that will be very useful to us.

We have expressed the spin state $|\psi\rangle$ of a spin-$\frac{1}{2}$ particle in the S_z basis as $|\psi\rangle = |+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle$. We can think of the rather strange-looking object

$$|+\mathbf{z}\rangle\langle+\mathbf{z}| + |-\mathbf{z}\rangle\langle-\mathbf{z}| \tag{2.44}$$

as the identity operator. It is an operator because when it is applied to a ket, it yields another ket. Moreover, if we apply it to the ket $|\psi\rangle$, we obtain

$$(|+\mathbf{z}\rangle\langle+\mathbf{z}| + |-\mathbf{z}\rangle\langle-\mathbf{z}|)|\psi\rangle = |+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle = |\psi\rangle \tag{2.45}$$

We earlier discussed a nice physical mechanism for inserting such an identity operator when we analyzed the effect of introducing a modified Stern–Gerlach device in Experiment 4 in Chapter 1. Here, since we are expressing an arbitrary state $|\psi\rangle$ in terms of the amplitudes to be in the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$, we use a modified SG device with its magnetic field gradient oriented along the z direction, as shown in Fig. 2.4a. The important point that we made in our discussion of the modified SG device was that because we do not make a measurement with such a device, the amplitudes to be in the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ combine together to yield

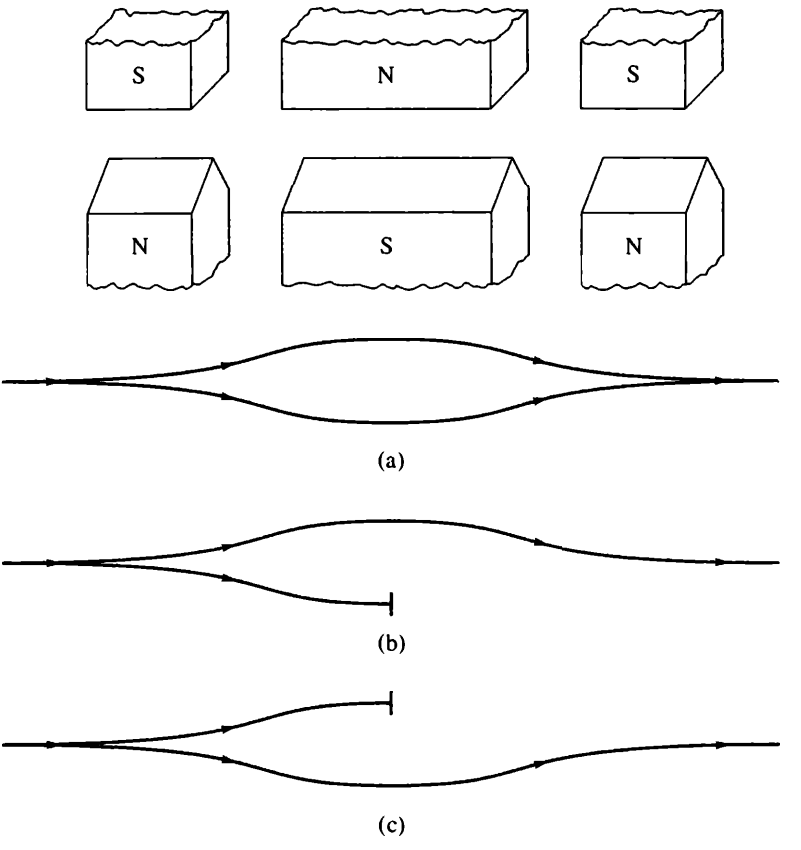


Figure 2.4 (a) A modified Stern–Gerlach device serves as the identity operator. (b) Blocking the path that a spin-down particle follows produces the projection operator $\hat{P}_+$. (c) Blocking the path that a spin-up particle follows produces the projection operator $\hat{P}_-$.

the same state exiting as entering the device, just as if the device were absent. Hence, it is indeed an identity operator.

The identity operator (2.44) may be viewed as being composed of two operators called **projection operators**:

$$\hat{P}_+ = |+\mathbf{z}\rangle\langle+\mathbf{z}| \tag{2.46a}$$

and

$$\hat{P}_- = |-\mathbf{z}\rangle\langle-\mathbf{z}| \tag{2.46b}$$

They are called projection operators because

$$\hat{P}_+|\psi\rangle = |+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle \tag{2.47a}$$

projects out the component of the ket $|\psi\rangle$ along $|+\mathbf{z}\rangle$ and

$$\hat{P}_-|\psi\rangle = |-\mathbf{z}\rangle\langle -\mathbf{z}|\psi\rangle \tag{2.47b}$$

projects out the component of the ket $|\psi\rangle$ along $|-\mathbf{z}\rangle$.[7] That (2.44) is the identity operator may be expressed in terms of the projection operators as

$$\hat{P}_+ + \hat{P}_- = 1 \tag{2.48}$$

This relation is often referred to as a **completeness relation**. Projecting onto the two vectors corresponding to spin up and spin down are the only possibilities for a spin-$\frac{1}{2}$ particle. As (2.45) shows, (2.48) is equivalent to saying that an arbitrary state $|\psi\rangle$ can be expressed as a superposition of the two basis states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$.

Notice that if we apply the projection operator $\hat{P}_+$ to the basis states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$, we obtain

$$\hat{P}_+|+\mathbf{z}\rangle = |+\mathbf{z}\rangle\langle +\mathbf{z}|+\mathbf{z}\rangle = |+\mathbf{z}\rangle \tag{2.49a}$$

and

$$\hat{P}_+|-\mathbf{z}\rangle = |+\mathbf{z}\rangle\langle +\mathbf{z}|-\mathbf{z}\rangle = 0 \tag{2.49b}$$

Thus $|+\mathbf{z}\rangle$ is an eigenstate of the projection operator $\hat{P}_+$ with eigenvalue 1, and $|-\mathbf{z}\rangle$ is an eigenstate of the projection operator $\hat{P}_+$ with eigenvalue 0. We can obtain a physical realization of the projection operator $\hat{P}_+$ from the modified SG device by blocking the path that would be taken by a particle in the state $|-\mathbf{z}\rangle$, that is, by blocking the lower path, as shown in Fig. 2.4b. Each particle in the state $|+\mathbf{z}\rangle$ entering the device exits the device. We can then say we have obtained the eigenvalue 1. Since none of the particles in the state $|-\mathbf{z}\rangle$ that enters the device also exits the device, we can say we have obtained the eigenvalue 0 in this case.

Similarly, we can create a physical realization of the projection operator $\hat{P}_-$ by blocking the upper path in the modified SG device, as shown in Fig. 2.4c. Then each particle in the state $|-\mathbf{z}\rangle$ that enters the device also exits the device:

$$\hat{P}_-|-\mathbf{z}\rangle = |-\mathbf{z}\rangle\langle -\mathbf{z}|-\mathbf{z}\rangle = |-\mathbf{z}\rangle \tag{2.50a}$$

while none of the particles in the state $|+\mathbf{z}\rangle$ exits the device:

$$\hat{P}_-|+\mathbf{z}\rangle = |-\mathbf{z}\rangle\langle -\mathbf{z}|+\mathbf{z}\rangle = 0 \tag{2.50b}$$

Hence the eigenvalues of $\hat{P}_-$ are 1 and 0 for the states $|-\mathbf{z}\rangle$ and $|+\mathbf{z}\rangle$, respectively.

---

[7] Notice that the projection operator may be applied to a bra vector as well:

$$\langle\psi|\hat{P}_+ = \langle\psi|+\mathbf{z}\rangle\langle +\mathbf{z}| \qquad \langle\psi|\hat{P}_- = \langle\psi|-\mathbf{z}\rangle\langle -\mathbf{z}|$$

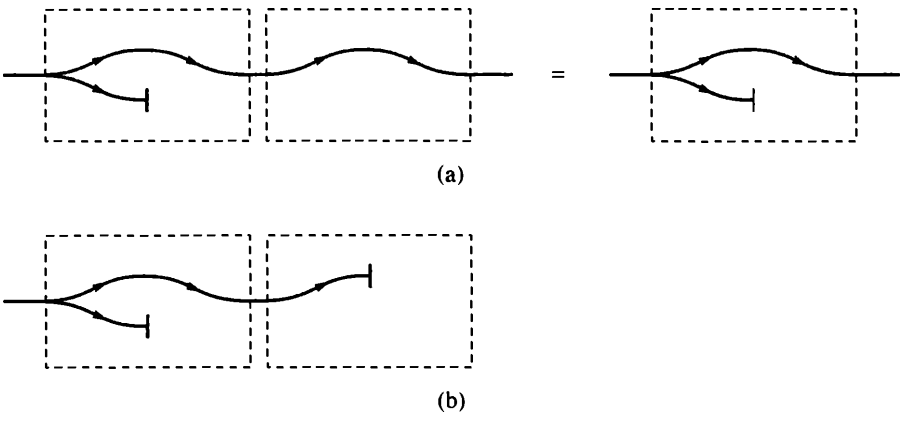


Figure 2.5 Physical realizations of (a) $\hat{P}_+^2 = \hat{P}_+$ and (b) $\hat{P}_- \hat{P}_+ = 0$.

Notice that each of the particles that has traversed one of the projection devices is certain to pass through a subsequent projection device of the same type:

$$\begin{aligned}\hat{P}_+^2 &= (|+\mathbf{z}\rangle\langle+\mathbf{z}|)(|+\mathbf{z}\rangle\langle+\mathbf{z}|) \\ &= |+\mathbf{z}\rangle\langle+\mathbf{z}|+\mathbf{z}\rangle\langle+\mathbf{z}| = |+\mathbf{z}\rangle\langle+\mathbf{z}| = \hat{P}_+ \qquad (2.51a)\end{aligned}$$

$$\begin{aligned}\hat{P}_-^2 &= (|-\mathbf{z}\rangle\langle-\mathbf{z}|)(|-\mathbf{z}\rangle\langle-\mathbf{z}|) \\ &= |-\mathbf{z}\rangle\langle-\mathbf{z}|-\mathbf{z}\rangle\langle-\mathbf{z}| = |-\mathbf{z}\rangle\langle-\mathbf{z}| = \hat{P}_- \qquad (2.51b)\end{aligned}$$

while a particle that passes a first projection device will surely fail to pass a subsequent projection device of the opposite type:

$$\begin{aligned}\hat{P}_+ \hat{P}_- &= (|+\mathbf{z}\rangle\langle+\mathbf{z}|)(|-\mathbf{z}\rangle\langle-\mathbf{z}|) \\ &= |+\mathbf{z}\rangle\langle+\mathbf{z}|-\mathbf{z}\rangle\langle-\mathbf{z}| = 0 \qquad (2.52a)\end{aligned}$$

$$\begin{aligned}\hat{P}_- \hat{P}_+ &= (|-\mathbf{z}\rangle\langle-\mathbf{z}|)(|+\mathbf{z}\rangle\langle+\mathbf{z}|) \\ &= |-\mathbf{z}\rangle\langle-\mathbf{z}|+\mathbf{z}\rangle\langle+\mathbf{z}| = 0 \qquad (2.52b)\end{aligned}$$

These results are illustrated in Fig. 2.5.

Our discussion of the identity operator and the projection operators has arbitrarily been phrased in terms of the S_z basis. We could as easily have expressed the same state $|\psi\rangle$ in terms of the S_x basis as $|\psi\rangle = |+\mathbf{x}\rangle\langle+\mathbf{x}|\psi\rangle + |-\mathbf{x}\rangle\langle-\mathbf{x}|\psi\rangle$. Thus we can also express the identity operator as

$$|+\mathbf{x}\rangle\langle+\mathbf{x}| + |-\mathbf{x}\rangle\langle-\mathbf{x}| = 1 \qquad (2.53)$$

and view it as being composed of projection operators onto the states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$.

Let's use this formalism to reexamine Experiment 4 of Chapter 1. In this experiment a particle in the state $|+\mathbf{z}\rangle$ passes through a modified SG**x** device and then

enters an SGz device. Since the modified SGx device acts as an identity operator, the particle entering the last SGz device is still in the state $|+\mathbf{z}\rangle$ and thus the amplitude to find the particle in the state $|-\mathbf{z}\rangle$ vanishes: $\langle -\mathbf{z}|+\mathbf{z}\rangle = 0$. There is, however, another way to express this amplitude. We use the identity operator (2.53) to express the initial ket in terms of the amplitudes to be the states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$:

$$|+\mathbf{z}\rangle = |+\mathbf{x}\rangle\langle +\mathbf{x}|+\mathbf{z}\rangle + |-\mathbf{x}\rangle\langle -\mathbf{x}|+\mathbf{z}\rangle \tag{2.54}$$

Then we have

$$\langle -\mathbf{z}|+\mathbf{z}\rangle = \langle -\mathbf{z}|+\mathbf{x}\rangle\langle +\mathbf{x}|+\mathbf{z}\rangle + \langle -\mathbf{z}|-\mathbf{x}\rangle\langle -\mathbf{x}|+\mathbf{z}\rangle \tag{2.55}$$

Thus the amplitude for a particle with $S_z = \hbar/2$ to have $S_z = -\hbar/2$ has now been written as the sum of two amplitudes. We read each of these amplitudes from right to left. The first amplitude on the right-hand side is the amplitude for a particle with $S_z = \hbar/2$ to have $S_x = \hbar/2$ times the amplitude for a particle with $S_x = \hbar/2$ to have $S_z = -\hbar/2$. The second amplitude is the amplitude for a particle with $S_z = \hbar/2$ to have $S_x = -\hbar/2$ times the amplitude for a particle with $S_x = -\hbar/2$ to have $S_z = -\hbar/2$. Notice that we multiply the individual amplitudes together and then add the resulting two amplitudes with the $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ intermediate states together to determine the total amplitude.

We now calculate the probability:

$$\begin{aligned} |\langle -\mathbf{z}|+\mathbf{z}\rangle|^2 = {} & |\langle -\mathbf{z}|+\mathbf{x}\rangle|^2 |\langle +\mathbf{x}|+\mathbf{z}\rangle|^2 + |\langle -\mathbf{z}|-\mathbf{x}\rangle|^2 |\langle -\mathbf{x}|+\mathbf{z}\rangle|^2 \\ & + \langle -\mathbf{z}|+\mathbf{x}\rangle\langle +\mathbf{x}|+\mathbf{z}\rangle\langle -\mathbf{z}|-\mathbf{x}\rangle^*\langle -\mathbf{x}|+\mathbf{z}\rangle^* \\ & + \langle -\mathbf{z}|+\mathbf{x}\rangle^*\langle +\mathbf{x}|+\mathbf{z}\rangle^*\langle -\mathbf{z}|-\mathbf{x}\rangle\langle -\mathbf{x}|+\mathbf{z}\rangle \end{aligned} \tag{2.56}$$

This looks like a pretty complicated way to calculate zero, but it is interesting to examine the significance of the four terms on the right-hand side. The first term is just the probability that a measurement of S_x on the initial state yields $\hbar/2$ times the probability that a measurement of S_z on a state with $S_x = \hbar/2$ yields $-\hbar/2$. The second term is the probability that a measurement of S_x on the initial state yields $-\hbar/2$ times the probability that a measurement of S_z on a state with $S_x = -\hbar/2$ yields $-\hbar/2$. These two terms, which sum to $\frac{1}{2}$, are just the terms we would have expected *if* we had made a measurement of S_x with the modified SGx device. But we did not make a measurement and actually distinguish which path the particle followed in the modified SGx device.[8] Thus there are two additional terms in (2.56), *interference* terms, that arise because we added the amplitudes on the right-hand side together *before* squaring to get the probability. You can verify that these two

[8] It should be emphasized that a measurement here means any physical interaction that would have permitted us *in principle* to distinguish which path is taken (such as arranging for the particle to leave a track in passing through the modified SG device), whether or not we actually choose to record this data.

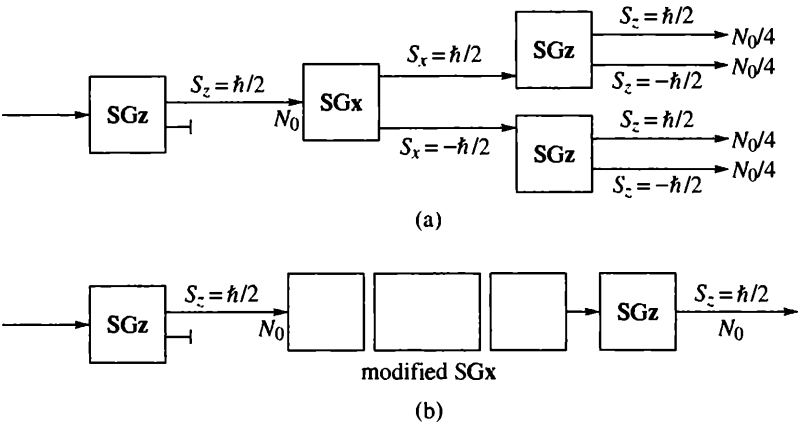


Figure 2.6 Block diagrams of experiments with SG devices in which (a) a measurement of S_x is carried out, illustrating $|\langle -\mathbf{z}|-\mathbf{x}\rangle|^2|\langle -\mathbf{x}|+\mathbf{z}\rangle|^2 + |\langle -\mathbf{z}|+\mathbf{x}\rangle|^2|\langle +\mathbf{x}|+\mathbf{z}\rangle|^2 = \frac{1}{2}$; and (b) no measurement of S_x is made, either by inserting a modified SGx device between the two SGz devices or by simply eliminating the SGx device pictured in (a), illustrating $|\langle -\mathbf{z}|-\mathbf{x}\rangle\langle -\mathbf{x}|+\mathbf{z}\rangle + \langle -\mathbf{z}|+\mathbf{x}\rangle\langle +\mathbf{x}|+\mathbf{z}\rangle|^2 = |\langle -\mathbf{z}|+\mathbf{z}\rangle|^2 = 0$.

interference terms do cancel the first two probabilities. These results are summarized in Fig. 2.6. In more general terms, if you do not make a measurement, you add the amplitudes to be in the different (indistinguishable) intermediate states, whereas if you do make a measurement that would permit you to distinguish among these states, you add the probabilities.

Finally, it is convenient to introduce the following shorthand notation. For a given two-dimensional basis, we can label our basis states by $|1\rangle$ and $|2\rangle$. We can then express the identity operator as

$$\sum_i |i\rangle\langle i| = 1 \tag{2.57}$$

where the sum is from $i = 1$ to $i = 2$. The straightforward generalization of this relationship to larger dimensional bases will be very useful to us later.

2.4 Matrix Representations of Operators

In order to change, or transform, kets, operators are required. Although one can discuss concepts such as the adjoint operator abstractly in terms of its action on the bra vectors, it is helpful to construct matrix representations for operators, making concepts such as adjoint and Hermitian operators more concrete, as well as providing the framework for matrix mechanics. Equation (2.25) is a typical equation of the form

$$\hat{A}|\psi\rangle = |\varphi\rangle \tag{2.58}$$

where $\hat{A}$ is an operator and $|\psi\rangle$ and $|\varphi\rangle$ are, in general, different kets. We can also think of the eigenvalue equation (2.35) as being of this form with $|\varphi\rangle$ just a constant times $|\psi\rangle$. Just as we can express a quantum spin state $|\psi\rangle$ using the S_z basis states by

$$|\psi\rangle = |+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle \tag{2.59}$$

we can write a comparable expression for $|\varphi\rangle$:

$$|\varphi\rangle = |+\mathbf{z}\rangle\langle+\mathbf{z}|\varphi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\varphi\rangle \tag{2.60}$$

Thus (2.58) becomes

$$\hat{A}\,(|+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle) = |+\mathbf{z}\rangle\langle+\mathbf{z}|\varphi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\varphi\rangle \tag{2.61}$$

In ordinary three-dimensional space, a vector equation such as $\mathbf{F} = m\mathbf{a}$ is really the three equations: $F_x = ma_x$, $F_y = ma_y$, and $F_z = ma_z$. We can formally obtain these three equations by taking the dot product of the vector equation with the basis vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$; for example, $\mathbf{i} \cdot \mathbf{F} = \mathbf{i} \cdot m\mathbf{a}$ yields $F_x = ma_x$. Similarly, we can think of (2.61) as two equations that we obtain by projecting (2.61) onto our two basis states, that is, by taking the inner product of this equation with the bras $\langle+\mathbf{z}|$ and $\langle-\mathbf{z}|$:

$$\langle+\mathbf{z}|\hat{A}|+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + \langle+\mathbf{z}|\hat{A}|-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle = \langle+\mathbf{z}|\varphi\rangle \tag{2.62a}$$

and

$$\langle-\mathbf{z}|\hat{A}|+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + \langle-\mathbf{z}|\hat{A}|-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle = \langle-\mathbf{z}|\varphi\rangle \tag{2.62b}$$

These two equations can be conveniently cast in matrix form:

$$\begin{pmatrix} \langle+\mathbf{z}|\hat{A}|+\mathbf{z}\rangle & \langle+\mathbf{z}|\hat{A}|-\mathbf{z}\rangle \\ \langle-\mathbf{z}|\hat{A}|+\mathbf{z}\rangle & \langle-\mathbf{z}|\hat{A}|-\mathbf{z}\rangle \end{pmatrix} \begin{pmatrix} \langle+\mathbf{z}|\psi\rangle \\ \langle-\mathbf{z}|\psi\rangle \end{pmatrix} = \begin{pmatrix} \langle+\mathbf{z}|\varphi\rangle \\ \langle-\mathbf{z}|\varphi\rangle \end{pmatrix} \tag{2.63}$$

In the same way that we can *represent* a ket $|\psi\rangle$ in the S_z basis by the column vector

$$|\psi\rangle \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle+\mathbf{z}|\psi\rangle \\ \langle-\mathbf{z}|\psi\rangle \end{pmatrix} \tag{2.64}$$

we can also *represent* the operator $\hat{A}$ in the S_z basis by the 2×2 matrix in (2.63). Just as for states, we indicate a representation of an operator with an arrow:

$$\hat{A} \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle+\mathbf{z}|\hat{A}|+\mathbf{z}\rangle & \langle+\mathbf{z}|\hat{A}|-\mathbf{z}\rangle \\ \langle-\mathbf{z}|\hat{A}|+\mathbf{z}\rangle & \langle-\mathbf{z}|\hat{A}|-\mathbf{z}\rangle \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{2.65}$$

If we label our basis vectors by $|1\rangle$ and $|2\rangle$ for the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$, respectively, we can express the **matrix elements** A_{ij} in the convenient form

$$A_{ij} = \langle i|\hat{A}|j\rangle \tag{2.66}$$

where i labels the rows and j labels the columns of the matrix. Note that knowing the four matrix elements in (2.63) allows us to determine the action of the operator $\hat{A}$ on any state $|\psi\rangle$.

MATRIX REPRESENTATIONS OF THE PROJECTION OPERATORS

As an example, the matrix representation of the projection operator $\hat{P}_+$ is given by

$$\hat{P}_+ \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle +\mathbf{z}|\hat{P}_+|+\mathbf{z}\rangle & \langle +\mathbf{z}|\hat{P}_+|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|\hat{P}_+|+\mathbf{z}\rangle & \langle -\mathbf{z}|\hat{P}_+|-\mathbf{z}\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.67a}$$

where we have taken advantage of (2.49) in evaluating the matrix elements. Similarly, the matrix representation of the projection operator $\hat{P}_-$ is given by

$$\hat{P}_- \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.67b}$$

Thus, the completeness relation $\hat{P}_+ + \hat{P}_- = 1$ in matrix form becomes

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{I} \tag{2.68}$$

where $\mathbb{I}$ is the identity matrix. The action of the projection operator $\hat{P}_+$ on the basis states is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.69a}$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.69b}$$

in agreement with equations (2.49a) and (2.49b), respectively.

MATRIX REPRESENTATION OF $\hat{J}_z$

As another example, consider the operator $\hat{J}_z$, the generator of rotations about the z axis. With the aid of (2.35), we can evaluate the matrix elements:

$$\begin{aligned} \hat{J}_z \xrightarrow[S_z \text{ basis}]{} & \begin{pmatrix} \langle +\mathbf{z}|\hat{J}_z|+\mathbf{z}\rangle & \langle +\mathbf{z}|\hat{J}_z|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|\hat{J}_z|+\mathbf{z}\rangle & \langle -\mathbf{z}|\hat{J}_z|-\mathbf{z}\rangle \end{pmatrix} \\ & = \begin{pmatrix} (\hbar/2)\langle +\mathbf{z}|+\mathbf{z}\rangle & (-\hbar/2)\langle +\mathbf{z}|-\mathbf{z}\rangle \\ (\hbar/2)\langle -\mathbf{z}|+\mathbf{z}\rangle & (-\hbar/2)\langle -\mathbf{z}|-\mathbf{z}\rangle \end{pmatrix} \\ & = \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \end{aligned} \tag{2.70}$$

The matrix is diagonal with the eigenvalues as the diagonal matrix elements because we are using the eigenstates of the operator as a basis and these eigenstates are orthogonal to each other. The eigenvalue equations $\hat{J}_z|+\mathbf{z}\rangle = (\hbar/2)|+\mathbf{z}\rangle$ and $\hat{J}_z|-\mathbf{z}\rangle = (-\hbar/2)|-\mathbf{z}\rangle$ may be expressed in matrix mechanics as

$$\begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.71}$$

and

$$\begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.72}$$

respectively. Incidentally, we can write the matrix representation (2.70) in the form

$$\hat{J}_z \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \frac{\hbar}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.73a}$$

which indicates that

$$\hat{J}_z = \frac{\hbar}{2}\hat{P}_+ - \frac{\hbar}{2}\hat{P}_- = \frac{\hbar}{2}|+\mathbf{z}\rangle\langle+\mathbf{z}| - \frac{\hbar}{2}|-\mathbf{z}\rangle\langle-\mathbf{z}| \tag{2.73b}$$

We could have also obtained this result directly in terms of bra and ket vectors by applying $\hat{J}_z$ to the identity operator (2.48).

EXAMPLE 2.3 Obtain the matrix representation of the rotation operator $\hat{R}(\phi\mathbf{k})$ in the S_z basis.

SOLUTION Since $\hat{R}(\phi\mathbf{k}) = e^{-i\hat{J}_z\phi/\hbar}$ and $e^{-i\hat{J}_z\phi/\hbar}|\pm\mathbf{z}\rangle = e^{\mp i\phi/2}|\pm\mathbf{z}\rangle$

$$\hat{R}(\phi\mathbf{k}) \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}$$

This matrix is diagonal because we are using the eigenstates of $\hat{J}_z$ as a basis.

MATRIX ELEMENTS OF THE ADJOINT OPERATOR

We next form the matrix representing the adjoint operator $\hat{A}^\dagger$. If an operator $\hat{A}$ acting on a ket $|\psi\rangle$ satisfies

$$\hat{A}|\psi\rangle = |\varphi\rangle \tag{2.74}$$

then, by definition,

$$\langle\psi|\hat{A}^\dagger = \langle\varphi| \tag{2.75}$$

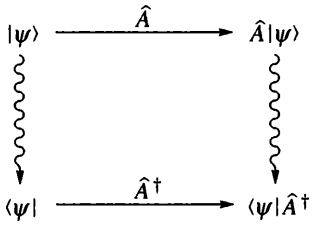


Figure 2.7 The adjoint operator $\hat{A}^\dagger$ of an operator $\hat{A}$ is defined by the correspondence between bras and kets.

(See Fig. 2.7.) If we take the inner product of (2.74) with the bra $\langle\chi|$, we have

$$\langle\chi|\hat{A}|\psi\rangle = \langle\chi|\varphi\rangle \tag{2.76}$$

while taking the inner product of (2.75) with the ket $|\chi\rangle$, we obtain

$$\langle\psi|\hat{A}^\dagger|\chi\rangle = \langle\varphi|\chi\rangle \tag{2.77}$$

Since $\langle\chi|\varphi\rangle = \langle\varphi|\chi\rangle^*$, we see that

$$\langle\psi|\hat{A}^\dagger|\chi\rangle = \langle\chi|\hat{A}|\psi\rangle^* \tag{2.78}$$

This straightforward but important result follows directly from our definition (2.75) of the adjoint operator. It can be used to tell us how the matrix representations of an operator and its adjoint are related. If we replace $|\psi\rangle$ and $|\chi\rangle$ with basis states such as $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$, we obtain

$$\langle i|\hat{A}^\dagger|j\rangle = \langle j|\hat{A}|i\rangle^* \tag{2.79}$$

We denote this as

$$A^\dagger_{ij} = A^*_{ji} \tag{2.80}$$

which tells us that the matrix representing the operator $\hat{A}^\dagger$ is the transpose conjugate of the matrix representing $\hat{A}$. We can define the **adjoint matrix** $\mathbb{A}^\dagger$ as the transpose conjugate of the matrix $\mathbb{A}$.

We also find another important result. Since by definition a Hermitian operator $\hat{A}$ satisfies $\hat{A} = \hat{A}^\dagger$, then $\langle i|\hat{A}|j\rangle = \langle j|\hat{A}|i\rangle^*$, showing that *the matrix representation of a Hermitian operator equals its transpose conjugate matrix*. Our terminology for adjoint and Hermitian operators is consistent with the terminology used in linear algebra for their matrix representations. We can now see from the explicit matrix representations of the operators $\hat{P}_+$ in (2.67) and $\hat{J}_z$ in (2.70) that these are Hermitian operators, since the matrices are diagonal with real elements (the eigenvalues) on the diagonal. In Chapter 3 we will see examples of Hermitian operators with off-diagonal elements when we examine the matrix representations for $\hat{J}_x$ and $\hat{J}_y$ for spin-$\frac{1}{2}$ and spin-1 particles.

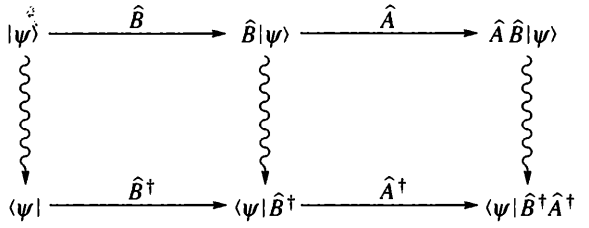


Figure 2.8 The adjoint of the product of operators is determined by the correspondence between bras and kets.

THE PRODUCT OF OPERATORS

We often must deal with situations where we have a product of operators, such as (2.51), which involves the product of two projection operators. Another way such a product of operators might arise is to perform two successive rotations on a state. To obtain the matrix representation of the product $\hat{A}\hat{B}$ of two operators, we first form the matrix element

$$\langle i|\hat{A}\hat{B}|j\rangle$$

If we insert the identity operator (2.57), we obtain

$$\langle i|\hat{A}\hat{B}|j\rangle = \langle i|\hat{A}\left(\sum_k |k\rangle\langle k|\right)\hat{B}|j\rangle = \sum_k \langle i|\hat{A}|k\rangle\langle k|\hat{B}|j\rangle = \sum_k A_{ik}B_{kj} \quad (2.81)$$

which is the usual rule for the multiplication of the matrices representing $\hat{A}$ and $\hat{B}$.

What is the adjoint operator for the product $\hat{A}\hat{B}$ of two operators? As Fig. 2.8 shows,

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \quad (2.82)$$

EXAMPLE 2.4 Use matrix mechanics to show that $\hat{P}_+^2 = \hat{P}_+$, $\hat{P}_-^2 = \hat{P}_-$, and $\hat{P}_+\hat{P}_- = 0$.

SOLUTION

$$\hat{P}_+\hat{P}_+ \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix} = \begin{pmatrix}1&0\\0&0\end{pmatrix} \xleftarrow[S_z \text{ basis}]{} \hat{P}_+$$

$$\hat{P}_-\hat{P}_- \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix}0&0\\0&1\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix} = \begin{pmatrix}0&0\\0&1\end{pmatrix} \xleftarrow[S_z \text{ basis}]{} \hat{P}_-$$

$$\hat{P}_+\hat{P}_- \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix} = \begin{pmatrix}0&0\\0&0\end{pmatrix}$$

2.5 Changing Representations

The rotation operator $\hat{R}^\dagger$ can be used to rotate a ket $|\psi\rangle$ into a new ket $|\psi'\rangle$ in an active transformation:

$$|\psi'\rangle = \hat{R}^\dagger|\psi\rangle \tag{2.83}$$

Recall that the rotation operator $\hat{R}^\dagger$ is just the inverse of the rotation operator $\hat{R}$, so if $\hat{R}$ rotates the state counterclockwise about the axis $\mathbf{n}$ by some angle θ, then $\hat{R}^\dagger$ rotates the state clockwise about the axis $\mathbf{n}$ by the same angle θ:

$$\hat{R}^\dagger(\theta\mathbf{n}) = \hat{R}(-\theta\mathbf{n}) \tag{2.84}$$

We can form a representation for the ket $|\psi'\rangle$ in the S_z basis, for example, in the usual way:

$$|\psi'\rangle \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle +\mathbf{z}|\psi'\rangle \\ \langle -\mathbf{z}|\psi'\rangle \end{pmatrix} = \begin{pmatrix} \langle +\mathbf{z}|\hat{R}^\dagger|\psi\rangle \\ \langle -\mathbf{z}|\hat{R}^\dagger|\psi\rangle \end{pmatrix} \tag{2.85}$$

There is, however, another way to view this transformation. Instead of the operator $\hat{R}^\dagger$ acting to the right on the ket, we can consider it as acting to the left on the bras. From our earlier discussion of the adjoint operator, we know that kets corresponding to the bras $\langle \pm\mathbf{z}|\hat{R}^\dagger$ are $\hat{R}|\pm\mathbf{z}\rangle$. Since $\hat{R}$ is the inverse of the operator $\hat{R}^\dagger$, we see that instead of $\hat{R}^\dagger$ rotating the state $|\psi\rangle$ into a new state $|\psi'\rangle$ as in (2.83), we may consider the operator $\hat{R}^\dagger$ in (2.85) to be performing the inverse rotation on the basis states that are used to form the representation.

Let's take some specific examples to illustrate. In Problem 3.5 it is shown that

$$|+\mathbf{x}\rangle = \hat{R}(\tfrac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle \tag{2.86}$$

where

$$|+\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle \tag{2.87}$$

From (2.42) we see that

$$\hat{R}(\tfrac{\pi}{2}\mathbf{k})|+\mathbf{x}\rangle = e^{-i\pi/4}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle\right) \tag{2.88}$$

which as we noted differs from the state we have defined as $|+\mathbf{y}\rangle$ by the overall phase factor of $e^{-i\pi/4}$. An alternative would be to define $|+\mathbf{y}\rangle = \hat{R}(\frac{\pi}{2}\mathbf{k})|+\mathbf{x}\rangle$ including this phase factor. Similarly, we would define the state $|-\mathbf{x}\rangle$ as one that is obtained by

$$|-\mathbf{x}\rangle = \hat{R}(\tfrac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle \tag{2.89}$$

that is by a rotation of the state $|-\mathbf{z}\rangle$ by 90° around the y axis. Following this procedure, as Problem 3.5 shows, we find that

$$|-\mathbf{x}\rangle = -\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle \tag{2.90}$$

which differs from (2.15) by an overall minus sign.

We will use the states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ shown in (2.87) and (2.90) for the remainder of this section since it is convenient to focus our discussion on basis states that are related to the states by $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ by application of a rotation operator, specifically

$$|\pm\mathbf{x}\rangle = \hat{R}(\tfrac{\pi}{2}\mathbf{j})|\pm\mathbf{z}\rangle \tag{2.91a}$$

and therefore

$$\langle\pm\mathbf{x}| = \langle\pm\mathbf{z}|\hat{R}^{\dagger}(\tfrac{\pi}{2}\mathbf{j}) \tag{2.91b}$$

If we take the operator $\hat{R}^{\dagger}$ in (2.85) to be the specific rotation operator $\hat{R}^{\dagger}(\frac{\pi}{2}\mathbf{j})$, then when this operator acts to the left on the bra vectors it transforms the S_z basis to the S_x basis according to (2.91b). But if $\hat{R}^{\dagger}(\frac{\pi}{2}\mathbf{j})$ acts to the right, it generates a new state

$$|\psi'\rangle = \hat{R}^{\dagger}(\tfrac{\pi}{2}\mathbf{j})|\psi\rangle \tag{2.92}$$

We can summarize our discussion in the following equation:

$$|\psi'\rangle \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \langle+\mathbf{z}|\psi'\rangle \\ \langle-\mathbf{z}|\psi'\rangle \end{pmatrix} = \begin{pmatrix} \langle+\mathbf{z}|\hat{R}^{\dagger}(\frac{\pi}{2}\mathbf{j})|\psi\rangle \\ \langle-\mathbf{z}|\hat{R}^{\dagger}(\frac{\pi}{2}\mathbf{j})|\psi\rangle \end{pmatrix} = \begin{pmatrix} \langle+\mathbf{x}|\psi\rangle \\ \langle-\mathbf{x}|\psi\rangle \end{pmatrix} \xleftarrow[S_x \text{ basis}]{} |\psi\rangle \tag{2.93}$$

Read from the left, this equation gives the representation in the S_z basis of the state $|\psi'\rangle$ that has been rotated by 90° clockwise around the y axis, whereas read from the right, it shows the state $|\psi\rangle$ as being unaffected but the basis vectors being rotated in the opposite direction, by 90° counterclockwise around the y axis. Both of these transformations lead to the same amplitudes, which we have combined into the column vector in (2.93). This alternative of rotating the basis states used to form a representation is often referred to as a **passive transformation** to distinguish it from an **active transformation** in which the state itself is rotated. A passive transformation is really just a rotation of our coordinate axes in our quantum mechanical vector space, as illustrated in Fig. 2.9.[9]

[9] If (2.43) did not seem sufficiently strange to you, try considering it from the perspective of a passive transformation. If we rotate our coordinate axes by 360° and end up with the same configuration of coordinate axes that we had originally, we find the state of a spin-$\frac{1}{2}$ particle has turned into the negative of itself.

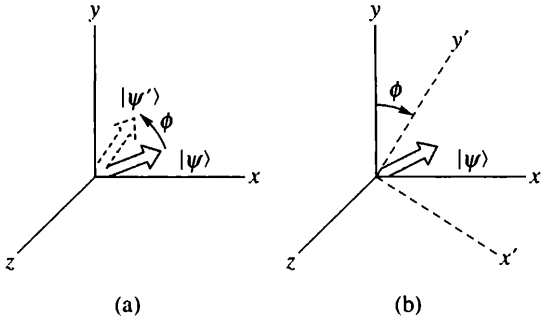


Figure 2.9 (a) Rotating a state by angle ϕ counterclockwise about an axis is equivalent to (b) rotating the coordinate axes by the same angle in the opposite direction, keeping the state fixed.

Equation (2.93) suggests a way to relate the column vector representing the ket $|\psi\rangle$ in one basis to the column vector representing the same ket in another basis. If we start with the representation of the ket $|\psi\rangle$ in the S_x basis and insert the identity operator, expressed in terms of S_z basis states, between the bra and the ket vectors, we obtain

$$\begin{pmatrix} \langle +\mathbf{x}|\psi\rangle \\ \langle -\mathbf{x}|\psi\rangle \end{pmatrix} = \begin{pmatrix} \langle +\mathbf{x}|+\mathbf{z}\rangle & \langle +\mathbf{x}|-\mathbf{z}\rangle \\ \langle -\mathbf{x}|+\mathbf{z}\rangle & \langle -\mathbf{x}|-\mathbf{z}\rangle \end{pmatrix} \begin{pmatrix} \langle +\mathbf{z}|\psi\rangle \\ \langle -\mathbf{z}|\psi\rangle \end{pmatrix}$$
$$= \begin{pmatrix} \langle +\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle & \langle +\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle & \langle -\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle \end{pmatrix} \begin{pmatrix} \langle +\mathbf{z}|\psi\rangle \\ \langle -\mathbf{z}|\psi\rangle \end{pmatrix} \quad (2.94)$$

where the second line follows from (2.91b). We call the 2×2 matrix in (2.94) $\mathbb{S}^\dagger$, or more precisely in this specific example $\mathbb{S}^\dagger(\frac{\pi}{2}\mathbf{j})$, since it is really the matrix representation in the S_z basis of the operator $\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})$ that rotates kets by 90° clockwise about the y axis. Equation (2.94) transforms a given ket $|\psi\rangle$ in the S_z basis into the S_x basis.

We can transform from the S_x basis to the S_z basis in analogous fashion:

$$\begin{pmatrix} \langle +\mathbf{z}|\psi\rangle \\ \langle -\mathbf{z}|\psi\rangle \end{pmatrix} = \begin{pmatrix} \langle +\mathbf{z}|+\mathbf{x}\rangle & \langle +\mathbf{z}|-\mathbf{x}\rangle \\ \langle -\mathbf{z}|+\mathbf{x}\rangle & \langle -\mathbf{z}|-\mathbf{x}\rangle \end{pmatrix} \begin{pmatrix} \langle +\mathbf{x}|\psi\rangle \\ \langle -\mathbf{x}|\psi\rangle \end{pmatrix}$$
$$= \begin{pmatrix} \langle +\mathbf{z}|\hat{R}(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle & \langle +\mathbf{z}|\hat{R}(\frac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|\hat{R}(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle & \langle -\mathbf{z}|\hat{R}(\frac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle \end{pmatrix} \begin{pmatrix} \langle +\mathbf{x}|\psi\rangle \\ \langle -\mathbf{x}|\psi\rangle \end{pmatrix} \quad (2.95)$$

where in the first line we have inserted the identity operator, this time expressed in terms of the S_x basis states. Also we have used (2.91a) to express the 2×2 matrix in the second line of the equation in terms of the matrix representation of the operator $\hat{R}(\frac{\pi}{2}\mathbf{j})$. Comparing the first lines of (2.94) and (2.95) reveals that the 2×2 matrix in (2.95) is the matrix $\mathbb{S}$, the adjoint matrix of the matrix $\mathbb{S}^\dagger$, since the matrix elements

of $\mathbb{S}$ are simply obtained from the matrix elements of $\mathbb{S}^\dagger$ by taking the transpose conjugate. Also, a comparison of the second lines of (2.94) and (2.95) shows that the 2×2 matrix in (2.95) is the matrix representation of $\hat{R}(\frac{\pi}{2}\mathbf{j})$, while the 2×2 matrix in (2.94) is the matrix representation of $\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})$. Since the rotation operators are unitary, the matrices must satisfy

$$\mathbb{S}^\dagger\mathbb{S} = \mathbb{I} \tag{2.96}$$

which can also be verified by substituting equation (2.95) directly into equation (2.94).

We can now determine how the matrix representation of an operator in one basis is related to the matrix representation in some other basis. For example, the matrix representing an operator $\hat{A}$ in the S_x basis is given by

$$\hat{A} \xrightarrow[S_x \text{ basis}]{} \begin{pmatrix} \langle +\mathbf{x}|\hat{A}|+\mathbf{x}\rangle & \langle +\mathbf{x}|\hat{A}|-\mathbf{x}\rangle \\ \langle -\mathbf{x}|\hat{A}|+\mathbf{x}\rangle & \langle -\mathbf{x}|\hat{A}|-\mathbf{x}\rangle \end{pmatrix} \tag{2.97}$$

A typical matrix element can be expressed as

$$\langle +\mathbf{x}|\hat{A}|-\mathbf{x}\rangle = \langle +\mathbf{z}|\hat{R}^\dagger(\tfrac{\pi}{2}\mathbf{j})\hat{A}\hat{R}(\tfrac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle$$

Inserting the identity operator (2.44) before and after the operator $\hat{A}$ on the left-hand side or between each of the operators on the right-hand side [or using result (2.81) for the matrix representation of the product of operators] permits us to write

$$\hat{A} \xrightarrow[S_x \text{ basis}]{} \mathbb{S}^\dagger\mathbb{A}\mathbb{S} \tag{2.98}$$

where $\mathbb{A}$ is the matrix representation of $\hat{A}$ in the S_z basis.[10]

Let's take the example of evaluating the matrix representation of $\hat{J}_z$ in the S_x basis. Using (2.87) and (2.90) to evaluate the matrix $\mathbb{S}$ in (2.95), we find

$$\mathbb{S} = \begin{pmatrix} \langle +\mathbf{z}|+\mathbf{x}\rangle & \langle +\mathbf{z}|-\mathbf{x}\rangle \\ \langle -\mathbf{z}|+\mathbf{x}\rangle & \langle -\mathbf{z}|-\mathbf{x}\rangle \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{2.99}$$

---

[10] The first lines of (2.94) and (2.95) form a good advertisement for the power of the identity operator. Rather than trying to remember such equations, it is probably easier and safer to derive them whenever needed by starting with the matrix elements (or amplitudes) that you are trying to find and inserting the identity operator from the appropriate basis set in the appropriate place(s). In this way we can work out the matrices in (2.98):

$$\begin{pmatrix} \langle +\mathbf{x}|\hat{A}|+\mathbf{x}\rangle & \langle +\mathbf{x}|\hat{A}|-\mathbf{x}\rangle \\ \langle -\mathbf{x}|\hat{A}|+\mathbf{x}\rangle & \langle -\mathbf{x}|\hat{A}|-\mathbf{x}\rangle \end{pmatrix}$$

$$= \begin{pmatrix} \langle +\mathbf{x}|+\mathbf{z}\rangle & \langle +\mathbf{x}|-\mathbf{z}\rangle \\ \langle -\mathbf{x}|+\mathbf{z}\rangle & \langle -\mathbf{x}|-\mathbf{z}\rangle \end{pmatrix} \begin{pmatrix} \langle +\mathbf{z}|\hat{A}|+\mathbf{z}\rangle & \langle +\mathbf{z}|\hat{A}|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|\hat{A}|+\mathbf{z}\rangle & \langle -\mathbf{z}|\hat{A}|-\mathbf{z}\rangle \end{pmatrix} \begin{pmatrix} \langle +\mathbf{z}|+\mathbf{x}\rangle & \langle +\mathbf{z}|-\mathbf{x}\rangle \\ \langle -\mathbf{z}|+\mathbf{x}\rangle & \langle -\mathbf{z}|-\mathbf{x}\rangle \end{pmatrix}$$

Carrying out the matrix multiplication (2.98) using the matrix representation of $\hat{J}_z$ in the S_z basis from (2.70), we obtain

$$\hat{J}_z \xrightarrow[S_x \text{ basis}]{} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \tag{2.100}$$

Comparing (2.100) with (2.70), we see that the matrix representation of the operator is no longer diagonal, since we are not using the eigenstates of the operator as the basis.[11]

If we also take advantage of (2.94) to express the eigenstate $|+\mathbf{z}\rangle$ in the S_x basis,

$$|+\mathbf{z}\rangle \xrightarrow[S_x \text{ basis}]{} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{2.101}$$

we can express the eigenvalue equation $\hat{J}_z|+\mathbf{z}\rangle = (\hbar/2)|+\mathbf{z}\rangle$ in the S_x basis:

$$\frac{\hbar}{2}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{\hbar}{2}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{2.102}$$

Compare (2.102) with (2.71), where the same equation is written in the S_z basis. Note that the eigenvalue equation is satisfied independently of the basis in which we choose to express it. This eigenvalue equation in its most basic form deals with operators and states, not with their representations, which we are free to choose in any way we want.

Before leaving this section, it is worth emphasizing again what we have learned. The S-matrices give us an easy way to transform both our states and our operators from one matrix representation to another. As the first line in both equations (2.94) and (2.95) shows, these S-matrices are composed of the amplitudes formed by taking the inner product of the basis kets of the representation we are transforming *from* with the basis bras of the representation we are transforming *to*. It is often convenient, however, to return to the active viewpoint with which we started our discussion. Instead of the S-matrices transforming a given state from one basis to another, we can view the S-matrix as the matrix representation of the rotation operator that rotates the given state into a different state within a fixed representation. This will be our starting point in Chapter 3. As we have seen, an active rotation that transforms the

---

[11] Alternatively, we could evaluate the matrix representation of $\hat{J}_z$ in the S_x basis by expressing the basis states $|\pm\mathbf{x}\rangle$ in terms of $|\pm\mathbf{z}\rangle$ so that we can let $\hat{J}_z$ act on them directly. For example, the element in the first row, second column of (2.100) is given by

$$\begin{aligned}\langle+\mathbf{x}|\hat{J}_z|-\mathbf{x}\rangle &= \frac{1}{2}\left(\langle+\mathbf{z}|+\langle-\mathbf{z}|\right)\hat{J}_z\left(-|+\mathbf{z}\rangle+|-\mathbf{z}\rangle\right) \\ &= \frac{1}{2}\left(\langle+\mathbf{z}|+\langle-\mathbf{z}|\right)\left(-\frac{\hbar}{2}|+\mathbf{z}\rangle-\frac{\hbar}{2}|-\mathbf{z}\rangle\right) = -\frac{\hbar}{2}\end{aligned}$$

state is just the inverse of the passive rotation that transforms the basis vectors used to form a particular representation.

EXAMPLE 2.5 The first lines of (2.94) and (2.95) as well as equation (2.98) and its inverse can be used to switch back and forth between the S_z and S_x bases for basis states such as

$$|+\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle \quad |-\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle - \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle$$

even though in this case the S-matrix

$$\mathbb{S} = \begin{pmatrix} \langle+\mathbf{z}|+\mathbf{x}\rangle & \langle+\mathbf{z}|-\mathbf{x}\rangle \\ \langle-\mathbf{z}|+\mathbf{x}\rangle & \langle-\mathbf{z}|-\mathbf{x}\rangle \end{pmatrix}$$

is *not* the matrix representation of the rotation operator. Determine $\mathbb{S}$ for these basis states and use it to repeat the calculations given in (2.100), (2.101), and (2.102).

SOLUTION

$$\mathbb{S} = \begin{pmatrix} \langle+\mathbf{z}|+\mathbf{x}\rangle & \langle+\mathbf{z}|-\mathbf{x}\rangle \\ \langle-\mathbf{z}|+\mathbf{x}\rangle & \langle-\mathbf{z}|-\mathbf{x}\rangle \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Thus in the S_x basis

$$\hat{J}_z \rightarrow \begin{pmatrix} \langle+\mathbf{x}|+\mathbf{z}\rangle & \langle+\mathbf{x}|-\mathbf{z}\rangle \\ \langle-\mathbf{x}|+\mathbf{z}\rangle & \langle-\mathbf{x}|-\mathbf{z}\rangle \end{pmatrix} \begin{pmatrix} \langle+\mathbf{z}|\hat{J}_z|+\mathbf{z}\rangle & \langle+\mathbf{z}|\hat{J}_z|-\mathbf{z}\rangle \\ \langle-\mathbf{z}|\hat{J}_z|+\mathbf{z}\rangle & \langle-\mathbf{z}|\hat{J}_z|-\mathbf{z}\rangle \end{pmatrix}$$
$$\times \begin{pmatrix} \langle+\mathbf{z}|+\mathbf{x}\rangle & \langle+\mathbf{z}|-\mathbf{x}\rangle \\ \langle-\mathbf{z}|+\mathbf{x}\rangle & \langle-\mathbf{z}|-\mathbf{x}\rangle \end{pmatrix}$$
$$= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The state $|+\mathbf{z}\rangle$ can be transformed into the S_x basis by the matrix $\mathbb{S}^\dagger$, which in this case is equal to the matrix $\mathbb{S}$:

$$|+\mathbf{z}\rangle \xrightarrow[S_x \text{ basis}]{} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Thus the eigenvalue equation $\hat{J}_z|+\mathbf{z}\rangle = (\hbar/2)|+\mathbf{z}\rangle$ in the S_x basis becomes

$$\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{\hbar}{2}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

As before, we see that the eigenvalue equation is satisfied with the same eigenvalue in either basis.

2.6 Expectation Values

It is interesting to see how we can use matrix mechanics to calculate expectation values of observables like the z component of the angular momentum with which we have associated the operator $\hat{J}_z$. If a spin-$\frac{1}{2}$ particle is in the state

$$|\psi\rangle = |+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + |-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle \tag{2.103}$$

then, as we saw in Section 1.4, the expectation value of S_z is given by

$$\langle S_z\rangle = \left(\frac{\hbar}{2}\right)|\langle+\mathbf{z}|\psi\rangle|^2 + \left(-\frac{\hbar}{2}\right)|\langle-\mathbf{z}|\psi\rangle|^2 \tag{2.104}$$

That is, the expectation value of S_z is the sum of the results $\hbar/2$ and $-\hbar/2$ of a measurement multiplied by the probability $|\langle+\mathbf{z}|\psi\rangle|^2$ and $|\langle-\mathbf{z}|\psi\rangle|^2$, respectively, of obtaining each result. We can express this expectation value in matrix mechanics as

$$\langle S_z\rangle = (\langle\psi|+\mathbf{z}\rangle, \langle\psi|-\mathbf{z}\rangle)\,\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\langle+\mathbf{z}|\psi\rangle\\ \langle-\mathbf{z}|\psi\rangle\end{pmatrix} \tag{2.105}$$

as can be verified by explicitly carrying out the matrix multiplication. The right-hand side of (2.105) is the representation in the S_z basis of $\langle\psi|\hat{J}_z|\psi\rangle$. Thus, we can also express the expectation value in the form

$$\langle S_z\rangle = \langle\psi|\hat{J}_z|\psi\rangle \tag{2.106}$$

In the language of eigenstates and eigenvalues, the expectation value (2.104) is the sum of the eigenvalues with each weighted by the probability of obtaining that eigenvalue. The advantage of expressing the expectation value in the form (2.106) is that we needn't evaluate it in a representation in which the basis states are the eigenstates of the operator in question. For example, we could evaluate (2.106) in the S_x basis by inserting the identity operator (2.53) between the bra vector and the operator and between the operator and the ket vector. Then we have

$$\langle S_z\rangle = (\langle\psi|+\mathbf{x}\rangle, \langle\psi|-\mathbf{x}\rangle)\begin{pmatrix}\langle+\mathbf{x}|\hat{J}_z|+\mathbf{x}\rangle & \langle+\mathbf{x}|\hat{J}_z|-\mathbf{x}\rangle\\ \langle-\mathbf{x}|\hat{J}_z|+\mathbf{x}\rangle & \langle-\mathbf{x}|\hat{J}_z|-\mathbf{x}\rangle\end{pmatrix}\begin{pmatrix}\langle+\mathbf{x}|\psi\rangle\\ \langle-\mathbf{x}|\psi\rangle\end{pmatrix} \tag{2.107}$$

You can verify that we can also go from (2.105) in the S_z basis to (2.107) in the S_x basis by inserting the identity operator $\mathbb{S}\mathbb{S}^\dagger$ before and after the 2×2 matrix in (2.105), provided we use the S-matrix (2.99) that transforms between these two basis sets.

As an example, let's return to (1.20), where we evaluated the expectation value of S_z for the state $|+\mathbf{x}\rangle$. Substituting the column vector representation (2.6) for this

ket in the S_z basis into (2.105), we see that the expectation value may be written in matrix form as

$$\langle S_z \rangle = \frac{1}{\sqrt{2}}(1,\ 1)\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0 \tag{2.108}$$

EXAMPLE 2.6 Use matrix mechanics to evaluate the expectation value $\langle S_z \rangle$ for the state $|+\mathbf{x}\rangle$ in the S_x basis states

$$|+\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle \quad |-\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle - \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle$$

SOLUTION In Example 2.5 we saw that in this basis

$$\hat{J}_z \rightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

then for the state $|+\mathbf{x}\rangle$

$$\langle S_z \rangle = (1,\ 0)\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

This result agrees of course with (2.108). In (2.108) the matrix form for the operator is especially straightforward, while here it is the representation for the state that is especially simple.

EXAMPLE 2.7 Use matrix mechanics to determine $\langle S_z \rangle$ for the state

$$|\psi\rangle = \frac{1}{2}|+\mathbf{z}\rangle + \frac{i\sqrt{3}}{2}|-\mathbf{z}\rangle$$

Compare your result with that of Example 1.2.

SOLUTION

$$\langle S_z \rangle = \langle \psi | \hat{S}_z | \psi \rangle = \frac{1}{2}(1,\ -i\sqrt{3})\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 \\ i\sqrt{3} \end{pmatrix} = -\frac{\hbar}{4}$$

in agreement with Example 1.2.

2.7 Photon Polarization and the Spin of the Photon

The previous discussion about representations of states and operators may seem somewhat mathematical in nature. The usefulness of this type of mathematics is just

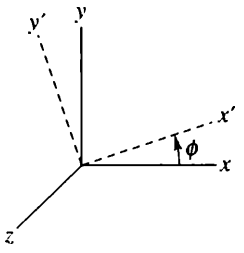


Figure 2.10 Two sets of transmission axes of a polarizer that may be used to create polarization states of photons traveling in the z direction.

a reflection of the fundamental underlying linear-vector-space structure of quantum mechanics. We conclude this chapter by looking at how we can apply this formalism to another physical two-state system, the polarization of the electromagnetic field. Many polarization effects can be described by classical physics, unlike the physics of spin-$\frac{1}{2}$ particles, which is a purely quantum phenomenon. Nonetheless, analyzing polarization effects using quantum mechanics can help to illuminate the differences between classical and quantum physics and at the same time tell us something fundamental about the quantum nature of the electromagnetic field.

Instead of a beam of spin-$\frac{1}{2}$ atoms passing through a Stern–Gerlach device, we consider a beam of photons, traveling in the z direction, passing through a linear polarizer. Those photons that pass through a polarizer with its transmission axis horizontal, that is, along the x axis, are said to be in the state $|x\rangle$, and those photons that pass through a polarizer with its transmission axis vertical are said to be in the state $|y\rangle$.[12] These two polarization states form a basis and the basis states satisfy $\langle x|y\rangle = 0$, since a beam of photons that passes through a polarizer whose transmission axis is vertical will be completely absorbed by a polarizer whose transmission axis is horizontal. Thus none of the photons will be found to be in the state $|x\rangle$ if they are put into the state $|y\rangle$ by virtue of having passed through the initial polarizer (assuming that our polarizers function with 100 percent efficiency).

We can also create polarized photons by sending the beam through a polarizer whose transmission axis is aligned at some angle to our original x-y axes. If the transmission axis is along the x' axis or y' axis shown in Fig. 2.10, the corresponding polarization states may be written as a superposition of the $|x\rangle$ and $|y\rangle$ polarization states as

$$\begin{aligned} |x'\rangle &= |x\rangle\langle x|x'\rangle + |y\rangle\langle y|x'\rangle \\ |y'\rangle &= |x\rangle\langle x|y'\rangle + |y\rangle\langle y|y'\rangle \end{aligned} \tag{2.109}$$

What are the amplitudes such as $\langle x|x'\rangle$, the amplitude for a photon linearly polarized along the x' axis to be found with its polarization along the x axis?

[12] These states are often referred to as $|x\rangle$ and $|y\rangle$. A different typeface is used to help distinguish these polarization states from position states, which will be introduced in Chapter 6.

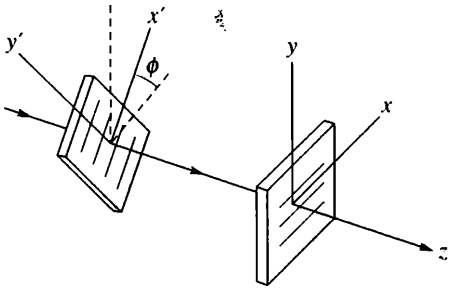


Figure 2.11 An x' polarizer followed by an x polarizer.

A classical physicist asked to determine the intensity of light passing through a polarizer with its transmission axis along either the x or the y axis *after* it has passed through a polarizer with its transmission axis along x', as pictured in Fig. 2.11, would calculate the component of the electric field along the x or the y axis and would square the *amplitude* of the field to determine the intensity passing through the second polarizer. If we denote the electric field after passage through the initial polarizer by $E_{x'}$, then the components of the field along the x and y axes are given by

$$E_x = E_{x'} \cos\phi \qquad E_y = E_{x'} \sin\phi$$

Thus the intensity of the light after passing through the second polarizer with its transmission axis along the x or y axis is proportional to $\cos^2\phi$ or $\sin^2\phi$, respectively. We can duplicate the classical results if we choose $\langle x|x'\rangle = \cos\phi$ and $\langle y|x'\rangle = \sin\phi$. Similarly, if the first polarizer has its transmission axis along the y' axis and we denote the electric field after passage through this polarizer by $E_{y'}$, then the components of the field along the x and y axes are given by

$$E_x = -E_{y'} \sin\phi \qquad E_y = E_{y'} \cos\phi$$

Again, we can duplicate the classical results if we choose $\langle x|y'\rangle = -\sin\phi$ and $\langle y|y'\rangle = \cos\phi$. Of course, the experiments outlined here alone do not give us any information about the phases of the amplitudes. However, since classical electromagnetic theory can account for interference phenomena such as the Young double-slit experiment, it is perhaps not too surprising that our conjectures about the amplitudes based on classical physics yield a valid quantum mechanical set, including phases:

$$\begin{aligned} |x'\rangle &= \cos\phi|x\rangle + \sin\phi|y\rangle \\ |y'\rangle &= -\sin\phi|x\rangle + \cos\phi|y\rangle \end{aligned} \tag{2.110}$$

Where do the quantum effects show up? Classical physics cannot account for the granular nature of the measurements, that a photomultiplier can detect photons coming in single lumps. Nor can it account for the inherently probabilistic nature of the measurements; we cannot do more than give a probability that a *single* photon in the state $|x'\rangle$ will pass through a polarizer with its transmission axis along x. For

example, if the angle $\phi = 60°$, then a single photon after having passed through an x' polarizer has a probability of $|\langle x|x'\rangle|^2 = \cos^2 60° = 0.25$ of passing through a second x polarizer. Knowing the polarization state of the photon does not, in general, determine whether it will pass through a subsequent polarizer. All we can determine is the probability, much to the discomfiture of the classical physicist who would like to believe that such results should be completely determined if enough information is known about the state of the system. The classical and quantum predictions are, however, in complete accord when the intensity of the beams is high so that the number of photons is large.

We can use (2.110) to calculate the matrix $\mathbb{S}^\dagger$ that transforms from the $|x\rangle$-$|y\rangle$ basis to the $|x'\rangle$-$|y'\rangle$ basis:

$$\mathbb{S}^\dagger = \begin{pmatrix} \langle x'|x\rangle & \langle x'|y\rangle \\ \langle y'|x\rangle & \langle y'|y\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \tag{2.111}$$

The matrix $\mathbb{S}$ that transforms from the $|x'\rangle$-$|y'\rangle$ basis to the $|x\rangle$-$|y\rangle$ basis is given by

$$\mathbb{S} = \begin{pmatrix} \langle x|x'\rangle & \langle x|y'\rangle \\ \langle y|x'\rangle & \langle y|y'\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \tag{2.112}$$

You can check that these matrices satisfy $\mathbb{S}^\dagger\mathbb{S} = \mathbb{I}$. All the elements of the matrix $\mathbb{S}$ are real. In fact, it is an example of an orthogonal matrix familiar from classical physics for rotating a vector in the x-y plane counterclockwise about the z axis by an angle ϕ. We can express $\mathbb{S}$ in terms of the rotation operator $\hat{R}(\phi\mathbf{k})$ that rotates the ket vectors themselves in this direction ($|x'\rangle = \hat{R}(\phi\mathbf{k})|x\rangle$ and $|y'\rangle = \hat{R}(\phi\mathbf{k})|y\rangle$):

$$\mathbb{S} = \begin{pmatrix} \langle x|\hat{R}(\phi\mathbf{k})|x\rangle & \langle x|\hat{R}(\phi\mathbf{k})|y\rangle \\ \langle y|\hat{R}(\phi\mathbf{k})|x\rangle & \langle y|\hat{R}(\phi\mathbf{k})|y\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \tag{2.113}$$

There is another set of basis vectors that have a great deal of physical significance but cannot be obtained from the $|x\rangle$-$|y\rangle$ basis by a simple rotation. We introduce

$$|R\rangle = \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle) \tag{2.114a}$$

$$|L\rangle = \frac{1}{\sqrt{2}}(|x\rangle - i|y\rangle) \tag{2.114b}$$

These states are referred to as right-circularly polarized and left-circularly polarized, respectively.

First, let's ask what the classical physicist would make of a right-circularly polarized electromagnetic plane wave of amplitude E_0 traveling in the z direction,

$$\mathbf{E} = E_0\mathbf{i}e^{i(kz-\omega t)} + iE_0\mathbf{j}e^{i(kz-\omega t)} \tag{2.115a}$$

Of course, the classical physicist uses complex numbers only as a convenient way to express a wave. The physics is determined by the real part of (2.115a), or

$$\mathbf{E} = E_0\mathbf{i}\cos(kz - \omega t) - E_0\mathbf{j}\sin(kz - \omega t) \tag{2.115b}$$

The "extra" factor of i in the y component of $\mathbf{E}$ in (2.115a) here means that the x and y components of the electric field are 90° out of phase, as (2.115b) shows. If we take $z = 0$ and examine the time dependence of the electromagnetic field, we see an $\mathbf{E}$ field that rotates in a circle as time progresses. If you curl your right hand in the direction of the changing $\mathbf{E}$, your thumb points in the direction of propagation along the positive z axis. The $\mathbf{E}$ field of the left-circularly polarized electromagnetic plane wave rotates in the opposite direction and thus would require you to curl your left hand in the direction of changing $\mathbf{E}$ to have your thumb point in the direction of propagation.

We can produce circularly polarized light by allowing linearly polarized light to fall on a birefringent crystal such as calcite that is cut so that the optic axis of the crystal lies in the x-y plane. Light polarized parallel to the optic axis in a birefringent crystal has a different index of refraction than does light perpendicular to the optic axis. We can orient our coordinate axes so that the optic axis is along x and the perpendicular axis is, of course, along y. Denoting the different indices of refraction by n_x and n_y, we see from (2.115a) that light polarized parallel to the x axis will pick up a phase $(n_x\omega/c)z$ in traversing a distance z through the crystal. Similarly, light polarized parallel to the y axis will gain a phase $(n_y\omega/c)z$. Thus a beam of linearly polarized light incident on such a crystal with its polarization axis inclined at 45° to the x axis will have equal magnitudes for the x and y components of the electric field, as indicated in Fig. 2.12, and there will be a phase difference $[(n_x - n_y)\omega/c]z$ between these two components that grows as the light passes through a distance z in the crystal. The crystal can be cut to a particular thickness, called a quarter-wave plate, so that the phase difference is 90° when the light of a particular wavelength exits the crystal, thus producing circularly polarized light.

What does the quantum physicist make of these circular polarization states (2.114)? Following the formalism of Section 2.2, it is instructive to ask how these states change under a rotation about the z axis. If we consider a right-circularly

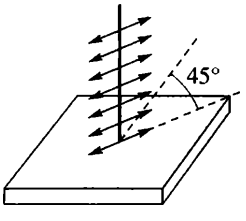


Figure 2.12 Plane-polarized light incident on a quarter-wave plate with its direction of polarization oriented at 45° to the optic axis will produce circularly polarized light.

polarized state that has been rotated by an angle ϕ counterclockwise about the z axis, we see that it can be expressed as

$$\begin{aligned}|R'\rangle &= \frac{1}{\sqrt{2}}\left(|x'\rangle + i|y'\rangle\right) \\ &= \frac{1}{\sqrt{2}}[\cos\phi|x\rangle + \sin\phi|y\rangle + i(-\sin\phi|x\rangle + \cos\phi|y\rangle)] \\ &= \frac{(\cos\phi - i\sin\phi)}{\sqrt{2}}(|x\rangle + i|y\rangle) \\ &= e^{-i\phi}|R\rangle \end{aligned} \tag{2.116}$$

Thus this state picks up only an overall phase factor when the state is rotated about the z axis. Based on our experience with the behavior of spin-$\frac{1}{2}$ states under rotations, (2.116) indicates that the state is one with definite angular momentum in the z direction. Since (2.32) shows that

$$|R'\rangle = \hat{R}(\phi\mathbf{k})|R\rangle = e^{-i\hat{J}_z\phi/\hbar}|R\rangle \tag{2.117}$$

consistency with the preceding equation requires that

$$\hat{J}_z|R\rangle = \hbar|R\rangle \tag{2.118}$$

Similarly, if we rotate the left-circularly polarized state by angle ϕ counterclockwise about the z axis, we obtain

$$|L'\rangle = e^{i\phi}|L\rangle \tag{2.119}$$

telling us that[13]

$$\hat{J}_z|L\rangle = -\hbar|L\rangle \tag{2.120}$$

Thus the right-circularly and left-circularly polarized states are eigenstates of $\hat{J}_z$, the operator that generates rotations about the z axis, but with eigenvalues $\pm\hbar$, not the $\pm\hbar/2$ characteristic of a spin-$\frac{1}{2}$ particle. In Chapter 3 we will see that the eigenvalues of $\hat{J}_z$ for a spin-1 particle are $+\hbar$, 0, and $-\hbar$. Photons have intrinsic spin of 1 instead of $\frac{1}{2}$. The absence of the 0 eigenvalue for $\hat{J}_z$ for a photon turns out to be a special characteristic of a massless particle, which moves at speed c.

[13] A particle with a positive (negative) projection of the intrinsic angular momentum along the direction of motion is said to have positive (negative) **helicity**. Photons thus come in two types, with both positive and negative helicity, corresponding to right- and left-circularly polarized light, respectively.

EXAMPLE 2.8 Determine the matrix representation of the angular momentum operator $\hat{J}_z$ using both the circular polarization vectors $|R\rangle$ and $|L\rangle$ and the linear polarization vectors $|x\rangle$ and $|y\rangle$ as a basis.

SOLUTION Let's start with the easy one first. Since the states $|R\rangle$ and $|L\rangle$ are eigenstates of $\hat{J}_z$ with eigenvalues $\hbar$ and $-\hbar$, respectively

$$\hat{J}_z \xrightarrow[|R\rangle\text{-}|L\rangle \text{ basis}] \begin{pmatrix} \langle R|\hat{J}_z|R\rangle & \langle R|\hat{J}_z|L\rangle \\ \langle L|\hat{J}_z|R\rangle & \langle L|\hat{J}_z|L\rangle \end{pmatrix} = \begin{pmatrix} \hbar & 0 \\ 0 & -\hbar \end{pmatrix}$$

The matrix is diagonal in this basis with the eigenvalues of the basis states on the diagonal. Switching to the linear polarization states $|x\rangle$ and $|y\rangle$:

$$\hat{J}_z \xrightarrow[|x\rangle\text{-}|y\rangle \text{ basis}] \begin{pmatrix} \langle x|R\rangle & \langle x|L\rangle \\ \langle y|R\rangle & \langle y|L\rangle \end{pmatrix} \begin{pmatrix} \langle R|\hat{J}_z|R\rangle & \langle R|\hat{J}_z|L\rangle \\ \langle L|\hat{J}_z|R\rangle & \langle L|\hat{J}_z|L\rangle \end{pmatrix} \begin{pmatrix} \langle R|x\rangle & \langle R|y\rangle \\ \langle L|x\rangle & \langle L|y\rangle \end{pmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \hbar & 0 \\ 0 & -\hbar \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

In this basis, the matrix has only off-diagonal elements. Since a Hermitian matrix is equal to its transpose, complex conjugate, both of these representations for $\hat{J}_z$ satisfy this condition, as they must.

2.8 Summary

In this chapter we have introduced operators in order to change a state into a different state. Since we are dealing here primarily with states of angular momentum, the natural operation is to rotate these states so that a state in which a component of the angular momentum has a definite value in a particular direction is rotated into a state in which the angular momentum has the same value in a different direction.[14] The operator that rotates states counterclockwise by angle ϕ about the z axis is

$$\hat{R}(\phi\mathbf{k}) = e^{-i\hat{J}_z\phi/\hbar} \tag{2.121}$$

where the operator $\hat{J}_z$ is called the generator of rotations about the z axis. In general, for an arbitrary operator $\hat{A}$, the bra corresponding to the ket

$$\hat{A}|\psi\rangle = |\varphi\rangle \tag{2.122a}$$

is

$$\langle\psi|\hat{A}^\dagger = \langle\varphi| \tag{2.122b}$$

[14] This way of describing a rotation of an angular momentum state may seem somewhat awkward, but in Chapter 3 we will see why we cannot say that the angular momentum simply points in a particular direction.

where the dagger denotes the adjoint operator. Thus the rotated bra corresponding to the rotated ket

$$\hat{R}(\phi\mathbf{k})|\psi\rangle = e^{-i\hat{J}_z\phi/\hbar}|\psi\rangle \tag{2.123a}$$

is given by

$$\langle\psi|\hat{R}^\dagger(\phi\mathbf{k}) = \langle\psi|e^{i\hat{J}_z^\dagger\phi/\hbar} \tag{2.123b}$$

In order for probability to be conserved under rotation,

$$\langle\psi|\hat{R}^\dagger(\phi\mathbf{k})\hat{R}(\phi\mathbf{k})|\psi\rangle = \langle\psi|e^{i\hat{J}_z^\dagger\phi/\hbar}e^{-i\hat{J}_z\phi/\hbar}|\psi\rangle = \langle\psi|\psi\rangle \tag{2.124}$$

which requires that the generators of rotation be Hermitian:

$$\hat{J}_z^\dagger = \hat{J}_z \tag{2.125}$$

An operator like the rotation operator that satisfies $\hat{R}^\dagger\hat{R} = 1$ is called a unitary operator.

For a spin-$\frac{1}{2}$ particle, the spin-up-along-z state $|+\mathbf{z}\rangle$ and spin-down-along-z state $|-\mathbf{z}\rangle$ satisfy

$$\hat{J}_z|\pm\mathbf{z}\rangle = \pm\frac{\hbar}{2}|\pm\mathbf{z}\rangle \tag{2.126}$$

showing that when the generator of rotations about the z axis acts on these states, the result is just the state itself multiplied by the value of S_z that these states are observed to have when a measurement of the intrinsic spin angular momentum in the z direction is carried out. Thus we can use a terminology in which we label the states $|\pm\mathbf{z}\rangle$ by $|S_z = \pm\hbar/2\rangle$, that is, we label the states by their values of S_z. Similarly, for example,

$$\hat{J}_x|\pm\mathbf{x}\rangle = \pm\frac{\hbar}{2}|\pm\mathbf{x}\rangle \tag{2.127}$$

where $\hat{J}_x$ is the generator of rotations about the x axis. In Chapter 3 we will argue on more general grounds that we should identify the generator of rotations with the component of the angular momentum along the axis about which the rotation is taking place. In subsequent chapters we will see that the operator that generates displacements in space is the linear momentum operator and the operator that generates time translations (moves the state forward in time) is the energy operator. Thus we will see repeated a pattern in which a Hermitian operator $\hat{A}$ is associated with a physical observable and the result a_n of a measurement for a particular state $|a_n\rangle$ satisfies

$$\hat{A}|a_n\rangle = a_n|a_n\rangle \tag{2.128}$$

Note that for a Hermitian, or self-adjoint, operator ($\hat{A} = \hat{A}^\dagger$), the bra equation corresponding to (2.128) is

$$\langle a_n|\hat{A} = \langle a_n|a_n^* \tag{2.129}$$

An equation in which an operator acting on a state yields a constant times the state is called an eigenvalue equation. In this case, the constant a_n in (2.128) is called the eigenvalue and the state $|a_n\rangle$ [or $\langle a_n|$ in (2.129)] is called the eigenstate.

We will now show that the eigenvalues of a Hermitian operator are real. Taking the inner product of the eigenvalue equation (2.128) with the bra $\langle a_k|$, we obtain

$$\langle a_k|\hat{A}|a_n\rangle = a_n\langle a_k|a_n\rangle \tag{2.130}$$

Taking advantage of (2.129), this equation becomes

$$a_k^*\langle a_k|a_n\rangle = a_n\langle a_k|a_n\rangle \tag{2.131a}$$

or

$$(a_k^* - a_n)\langle a_k|a_n\rangle = 0 \tag{2.131b}$$

Note that if we take $k = n$, we find

$$(a_n^* - a_n)\langle a_n|a_n\rangle = 0 \tag{2.132}$$

and therefore the eigenvalues of a Hermitian operator are real ($a_n^* = a_n$), a necessary condition if these are to be the values that we obtain for a measurement. Moreover, (2.131b) shows that

$$\langle a_k|a_n\rangle = 0 \quad a_k \neq a_n \tag{2.133}$$

as we argued in Chapter 1 must be true based on the fact that $\langle a_k|a_n\rangle$ is the amplitude to obtain a_k for a particle in the state $|a_n\rangle$. This shows that the eigenstates of a Hermitian operator corresponding to distinct eigenvalues are orthogonal. Thus our association of Hermitian operators with observables such as angular momentum forms a nice, self-consistent physical picture.

We also see that we can express the expectation value $\langle A\rangle$ of the observable A in terms of the operator $\hat{A}$ as

$$\langle A\rangle = \langle\psi|\hat{A}|\psi\rangle \tag{2.134}$$

For simplicity, let's consider the case where there are two eigenstates $|a_1\rangle$ and $|a_2\rangle$ with $a_1 \neq a_2$, as is the case for spin $\frac{1}{2}$. Since a general state can be written as

$$|\psi\rangle = c_1|a_1\rangle + c_2|a_2\rangle \tag{2.135}$$

then

$$\begin{aligned}\langle\psi|\hat{A}|\psi\rangle &= (c_1^*\langle a_1| + c_2^*\langle a_2|)\hat{A}(c_1|a_1\rangle + c_2|a_2\rangle) \\ &= (c_1^*\langle a_1| + c_2^*\langle a_2|)(c_1a_1|a_1\rangle + c_2a_2|a_2\rangle) \\ &= |c_1|^2a_1 + |c_2|^2a_2 \\ &= \langle A\rangle \end{aligned} \tag{2.136}$$

where the last step follows since the penultimate line of (2.136) is just the sum of the eigenvalues weighted by the probability of obtaining each of those values, which is just what we mean by the expectation value.

Also note that, as in (1.40), (2.135) can be expressed in the form

$$|\psi\rangle = |a_1\rangle\langle a_1|\psi\rangle + |a_2\rangle\langle a_2|\psi\rangle \tag{2.137}$$

This suggests that we can write the identity operator in the form

$$|a_1\rangle\langle a_1| + |a_2\rangle\langle a_2| = 1 \tag{2.138}$$

which is also known as a completeness relation, because it is equivalent to saying that we can express an arbitrary state $|\psi\rangle$ as a superposition of the states $|a_1\rangle$ and $|a_2\rangle$, as shown in (2.137). The identity operator can be decomposed into projection operators

$$\hat{P}_1 = |a_1\rangle\langle a_1| \quad \text{and} \quad \hat{P}_2 = |a_2\rangle\langle a_2| \tag{2.139}$$

that project out of the state $|\psi\rangle$ the component of the vector in the direction of the eigenvector. For example,

$$\hat{P}_1|\psi\rangle = |a_1\rangle\langle a_1|\psi\rangle \tag{2.140}$$

If we insert the identity operator (2.138) between the ket and the bra in the amplitude $\langle\varphi|\psi\rangle$, we obtain

$$\langle\varphi|\psi\rangle = \langle\varphi|a_1\rangle\langle a_1|\psi\rangle + \langle\varphi|a_2\rangle\langle a_2|\psi\rangle \tag{2.141}$$

Thus, if a particle is in the state $|\psi\rangle$ and a measurement is carried out, the probability of finding the particle in the state $|\varphi\rangle$ can be written as

$$|\langle\varphi|\psi\rangle|^2 = |\langle\varphi|a_1\rangle\langle a_1|\psi\rangle + \langle\varphi|a_2\rangle\langle a_2|\psi\rangle|^2 \tag{2.142}$$

Note that the amplitudes $\langle\varphi|a_1\rangle\langle a_1|\psi\rangle$ and $\langle\varphi|a_2\rangle\langle a_2|\psi\rangle$ can interfere with each other. Equation (2.142) presumes that no measurement of the observable A has actually taken place. If we were to actually insert a device that measured the observable A for the state $|\psi\rangle$, we would then find the probability to obtain the state $|\varphi\rangle$ given by

$$|\langle\varphi|a_1\rangle|^2|\langle a_1|\psi\rangle|^2 + |\langle\varphi|a_2\rangle|^2|\langle a_2|\psi\rangle|^2 \tag{2.143}$$

which is just the sum of the probabilities of finding $|\psi\rangle$ in the states $|a_1\rangle$ and $|a_2\rangle$ times the probability that each of these states is found in the state $|\varphi\rangle$. Equations (2.142) and (2.143) illustrate one of the fundamental principles of quantum mechanics: When we do not make a measurement that permits us to distinguish the intermediate states $|a_1\rangle$ and $|a_2\rangle$, we add the amplitudes and then square to get the probability, while if we do make a measurement that can distinguish which of the states $|a_1\rangle$ and $|a_2\rangle$ the particle is in, we add the individual probabilities, not the amplitudes. For a specific example, see the discussion at the end of Section 2.3.

A convenient shorthand notation is to use the eigenstates $|a_1\rangle$ and $|a_2\rangle$ as a basis and represent a ket such as (2.135) by a column vector

$$|\psi\rangle \xrightarrow[|a_1\rangle\text{-}|a_2\rangle \text{ basis}]{} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \langle a_1|\psi\rangle \\ \langle a_2|\psi\rangle \end{pmatrix} \tag{2.144}$$

a bra by a row vector

$$\langle\psi| \xrightarrow[|a_1\rangle\text{-}|a_2\rangle \text{ basis}]{} (c_1^*, c_2^*) = \big(\langle\psi|a_1\rangle, \langle\psi|a_2\rangle\big) \tag{2.145}$$

and an operator by a matrix

$$\hat{B} \xrightarrow[|a_1\rangle\text{-}|a_2\rangle \text{ basis}]{} \begin{pmatrix} \langle a_1|\hat{B}|a_1\rangle & \langle a_1|\hat{B}|a_2\rangle \\ \langle a_2|\hat{B}|a_1\rangle & \langle a_2|\hat{B}|a_2\rangle \end{pmatrix} \tag{2.146}$$

In this notation, an equation such as

$$\hat{B}|\psi\rangle = |\varphi\rangle \tag{2.147}$$

becomes

$$\begin{pmatrix} \langle a_1|\hat{B}|a_1\rangle & \langle a_1|\hat{B}|a_2\rangle \\ \langle a_2|\hat{B}|a_1\rangle & \langle a_2|\hat{B}|a_2\rangle \end{pmatrix} \begin{pmatrix} \langle a_1|\psi\rangle \\ \langle a_2|\psi\rangle \end{pmatrix} = \begin{pmatrix} \langle a_1|\varphi\rangle \\ \langle a_2|\varphi\rangle \end{pmatrix} \tag{2.148}$$

Knowing the matrix elements $\langle a_i|\hat{B}|a_j\rangle$ permits us to evaluate the action of the operator $\hat{B}$ on any state $|\psi\rangle$. As an example, we can use matrix mechanics to evaluate the expectation value of B in the state $|\psi\rangle$:

$$\langle B\rangle = \langle\psi|\hat{B}|\psi\rangle = \big(\langle\psi|a_1\rangle, \langle\psi|a_2\rangle\big) \begin{pmatrix} \langle a_1|\hat{B}|a_1\rangle & \langle a_1|\hat{B}|a_2\rangle \\ \langle a_2|\hat{B}|a_1\rangle & \langle a_2|\hat{B}|a_2\rangle \end{pmatrix} \begin{pmatrix} \langle a_1|\psi\rangle \\ \langle a_2|\psi\rangle \end{pmatrix} \tag{2.149}$$

where the last step follows from inserting the identity operator (2.138) between the bra $\langle\psi|$ and the operator $\hat{B}$ and between the operator $\hat{B}$ and the ket $|\psi\rangle$. Finally, note

that if basis states are the eigenstates of the operator, the matrix representation is diagonal with the eigenvalues forming the diagonal matrix elements:[15]

$$\hat{A} \xrightarrow[|a_1\rangle\text{-}|a_2\rangle \text{ basis}]{} \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \tag{2.150}$$

All of the results (2.135) through (2.150) can be extended in a straightforward fashion to larger dimensional bases, as introduced in Section 1.6. For example, the identity operator is given by $\sum_n |a_n\rangle\langle a_n|$ in the more general case.

Problems

2.1. Show that

$$\lim_{N\to\infty} \left(1 + \frac{x}{N}\right)^N = e^x$$

by comparing the Taylor series expansions for the two functions.

2.2. Use Dirac notation (the properties of kets, bras, and inner products) directly without explicitly using matrix representations to establish that the projection operator $\hat{P}_+$ is Hermitian. Use the fact that $\hat{P}_+^2 = \hat{P}_+$ to establish that the eigenvalues of the projection operator are 1 and 0.

2.3. Determine the matrix representation of the rotation operator $\hat{R}(\phi\mathbf{k})$ using the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ as a basis. Using your matrix representation, verify that the rotation operator is unitary, that is, it satisfies $\hat{R}^\dagger(\phi\mathbf{k})\hat{R}(\phi\mathbf{k}) = 1$.

2.4. Determine the column vectors representing the states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ using the states $|+\mathbf{y}\rangle$ and $|-\mathbf{y}\rangle$ as a basis.

2.5. What is the matrix representation of $\hat{J}_z$ using the states $|+\mathbf{y}\rangle$ and $|-\mathbf{y}\rangle$ as a basis? Use this representation to evaluate the expectation value of S_z for a collection of particles each in the state $|-\mathbf{y}\rangle$.

2.6. Evaluate $\hat{R}(\theta\mathbf{j})|+\mathbf{z}\rangle$, where $\hat{R}(\theta\mathbf{j}) = e^{-i\hat{J}_y\theta/\hbar}$ is the operator that rotates kets counterclockwise by angle θ about the y axis. Show that $\hat{R}(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle = |+\mathbf{x}\rangle$. *Sug-*

[15] In general, there are an infinite number of sets of basis states that may be used to form representations in matrix mechanics. For example, in addition to the states $|\pm\mathbf{z}\rangle$, the states $|\pm\mathbf{x}\rangle$ can be used as a basis to represent states and operators for spin-$\frac{1}{2}$ particles. However, since $|\pm\mathbf{x}\rangle$ are not eigenstates of $\hat{J}_z$, the matrix representation of this operator using these states as a basis is not diagonal, as (2.100) shows.

gestion: Express the ket $|+\mathbf{z}\rangle$ as a superposition of the kets $|+\mathbf{y}\rangle$ and $|-\mathbf{y}\rangle$ and take advantage of the fact that $\hat{J}_y|\pm\mathbf{y}\rangle = (\pm\hbar/2)|\pm\mathbf{y}\rangle$; then switch back to the $|+\mathbf{z}\rangle$-$|-\mathbf{z}\rangle$ basis.

2.7. Work out the matrix representations of the projection operators $\hat{P}_+ = |+\mathbf{z}\rangle\langle+\mathbf{z}|$ and $\hat{P}_- = |-\mathbf{z}\rangle\langle-\mathbf{z}|$ using the states $|+\mathbf{y}\rangle$ and $|-\mathbf{y}\rangle$ of a spin-$\frac{1}{2}$ particle as a basis. Check that the results (2.51) and (2.52) are satisfied using these matrix representations.

2.8. The column vector representing the state $|\psi\rangle$ is given by

$$|\psi\rangle \xrightarrow[S_z \text{ basis}]{} \frac{1}{\sqrt{5}}\begin{pmatrix} i \\ 2 \end{pmatrix}$$

Using matrix mechanics, show that $|\psi\rangle$ is properly normalized and calculate the probability that a measurement of S_x yields $\hbar/2$. Also determine the probability that a measurement of S_y yields $\hbar/2$.

2.9. Suppose in a two-dimensional basis that the operators $\hat{A}$ and $\hat{B}$ are represented by the 2×2 matrices

$$\hat{A} \rightarrow \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \hat{B} \rightarrow \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

Show that $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$.

2.10. Determine the matrix representation of $\hat{J}_x$ in the S_z basis. *Suggestion:* Start with the matrix representation of the operator $\hat{S}_x$ using the states

$$|+\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle \quad |-\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle - \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle$$

as a basis and then transform to the S_z basis.

2.11. The column vector representing the state $|\psi\rangle$ is given by

$$|\psi\rangle \xrightarrow[S_z \text{ basis}]{} \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}$$

Use matrix mechanics and the result of Problem 2.10 to determine $\langle S_x\rangle$ for this state.

2.12. A photon polarization state for a photon propagating in the z direction is given by

$$|\psi\rangle = \sqrt{\frac{2}{3}}|x\rangle + \frac{i}{\sqrt{3}}|y\rangle$$

(a) What is the probability that a photon in this state will pass through an ideal polarizer with its transmission axis oriented in the y direction?

(b) What is the probability that a photon in this state will pass through an ideal polarizer with its transmission axis y' making an angle ϕ with the y axis?

(c) A beam carrying N photons per second, each in the state $|\psi\rangle$, is totally absorbed by a black disk with its normal to the surface in the z direction. How large is the torque exerted on the disk? In which direction does the disk rotate? *Reminder:* The photon states $|R\rangle$ and $|L\rangle$ each carry a unit $\hbar$ of angular momentum parallel and antiparallel, respectively, to the direction of propagation of the photons.

(d) How would the result for each of these questions differ if the polarization state were

$$|\psi'\rangle = \sqrt{\frac{2}{3}}|x\rangle + \frac{1}{\sqrt{3}}|y\rangle$$

that is, the "i" in the state $|\psi\rangle$ is absent?

2.13. A system of N ideal linear polarizers is arranged in sequence, as shown in Fig. 2.13. The transmission axis of the first polarizer makes an angle of ϕ/N with the y axis. The transmission axis of every other polarizer makes an angle of ϕ/N with respect to the axis of the preceding one. Thus, the transmission axis of the final polarizer makes an angle ϕ with the y axis. A beam of y-polarized photons is incident on the first polarizer.

(a) What is the probability that an incident photon is transmitted by the array?

(b) Evaluate the probability of transmission in the limit of large N.

(c) Consider the special case with the angle $\phi = 90°$. Explain why your result is not in conflict with the fact that $\langle x|y\rangle = 0$.[16]

2.14.

(a) Determine a 2×2 matrix $\mathbb{S}$ that can be used to transform a column vector representing a photon polarization state using the linear polarization vectors $|x\rangle$ and $|y\rangle$ as a basis to one using the circular polarization vectors $|R\rangle$ and $|L\rangle$ as a basis.

(b) Using matrix multiplication, verify explicitly that the matrix $\mathbb{S}$ that you found in (a) is unitary.

---

[16] A nice discussion of the quantum state using photon polarization states as a basis is given by A. P. French and E. F. Taylor, *An Introduction to Quantum Physics*, Norton, New York, 1978, Chapters 6 and 7. Problem 2.9 is adapted from this source.

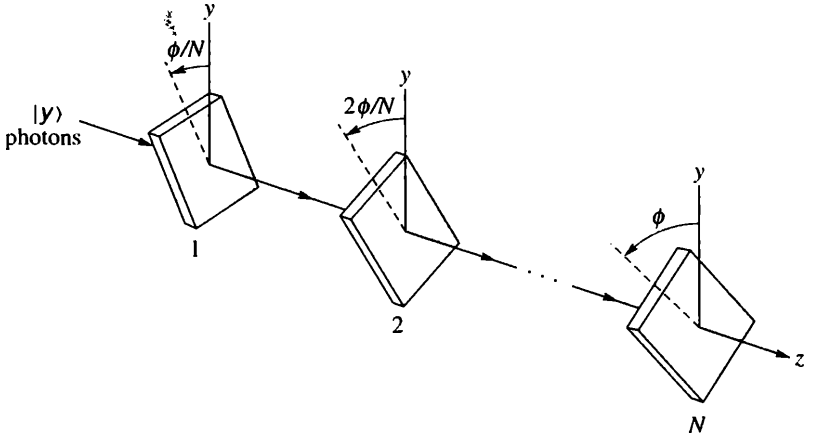


Figure 2.13 An array of N linear polarizers.

2.15. Evaluate the matrix elements

$$\begin{pmatrix} \langle x|\hat{J}_z|x\rangle & \langle x|\hat{J}_z|y\rangle \\ \langle y|\hat{J}_z|x\rangle & \langle y|\hat{J}_z|y\rangle \end{pmatrix}$$

by expressing the linear polarization states $|x\rangle$ and $|y\rangle$ in terms of the circular polarization states $|R\rangle$ and $|L\rangle$. Compare your result with that given in Example 2.8.

2.16. Use both the matrix representations of the angular momentum operator $\hat{J}_z$ from Example 2.8 to determine the expectation value of the angular momentum for the photon state $a|R\rangle + b|L\rangle$.

2.17. Use the matrix representation of the rotation operator $\hat{R}(\phi\mathbf{k})$ in the $|x\rangle$-$|y\rangle$ basis as given in (2.113) to establish that the photon circular polarization states (2.114), expressed as column vectors in the $|x\rangle$-$|y\rangle$ basis, are eigenstates of the rotation operator with the eigenvalues that appear in (2.116) and (2.119).

2.18. Construct projection operators out of bras and kets for x-polarized and y-polarized photons. Give physical examples of devices that can serve as these projection operators. Use (a) the properties of bras and kets and (b) the properties of the physical devices to show that the projection operators satisfy $\hat{P}_x^2 = \hat{P}_x$, $\hat{P}_y^2 = \hat{P}_y$, and $\hat{P}_x\hat{P}_y = \hat{P}_y\hat{P}_x = 0$.

2.19. Show that $\hat{J}_z = \hbar|R\rangle\langle R| - \hbar|L\rangle\langle L|$ for photons.

2.20. What is the probability that a right-circularly polarized photon will pass through a linear polarizer with its transmission axis along the x' axis, which makes an angle ϕ with the x axis?

2.21. Linearly polarized light of wavelength 5890 Å is incident normally on a birefringent crystal that has its optic axis parallel to the face of the crystal, along the x axis. If the incident light is polarized at an angle of 45° to the x and y axes, what is the probability that the photons exiting a crystal of thickness 100.0 microns will be right-circularly polarized? The index of refraction for light of this wavelength polarized along y (perpendicular to the optic axis) is 1.66 and the index of refraction for light polarized along x (parallel to the optic axis) is 1.49.

2.22. A beam of linearly polarized light is incident on a quarter-wave plate with its direction of polarization oriented at 30° to the optic axis. Subsequently, the beam is absorbed by a black disk. Determine the rate at which angular momentum is transferred to the disk, assuming the beam carries N photons per second.

2.23.

(a) Show that if the states $|a_n\rangle$ form an orthonormal basis, so do the states $\hat{U}|a_n\rangle$, provided $\hat{U}$ is unitary.

(b) Show that the eigenvalues of a unitary operator can be written as $e^{i\theta}$.

2.24. The Hermitian operator $\hat{A}$ corresponding to the observable A has two eigenstates $|a_1\rangle$ and $|a_2\rangle$ with eigenvalues a_1 and a_2, respectively. Assume $a_1 \neq a_2$. Show that $\hat{A}$ can be written in the form

$$\hat{A} = a_1|a_1\rangle\langle a_1| + a_2|a_2\rangle\langle a_2|$$

and that

$$\langle\psi|\hat{A}|\psi\rangle = \langle A\rangle$$

CHAPTER 3

Angular Momentum

In this chapter we will see that the order in which we carry out rotations about different axes matters. Therefore, the operators that generate rotations about these different axes do not commute, leading to commutation relations that may be viewed as the defining relations for the angular momentum operators. We will use these commutation relations to determine the angular momentum eigenstates and eigenvalues. We will also see that the spin-$\frac{1}{2}$ states that have occupied much of our attention so far appear as a particular case of this general analysis of angular momentum in quantum mechanics.

3.1 Rotations Do Not Commute and Neither Do the Generators

Take your textbook and set up a convenient coordinate system centered on the book, as shown in Fig. 3.1. Rotate your text by 90° about the x axis and then rotate it by 90° about the y axis. Either note carefully the orientation of the text or, better still, borrow a copy of the text from a friend and perform the two rotations again, but this time first rotate about the y axis by 90° and then about the x axis by 90°. The orientations of the two texts are different. Clearly, the order in which you carry out the rotations matters. We say that finite rotations about different axes do not commute.

In Section 2.7 we determined the matrix $\mathbb{S}$ that transforms a basis set of polarization states to another set that are related to the initial set by a rotation by angle ϕ counterclockwise about the z axis. The matrix (2.112) is also the matrix that is used to rotate the components of an ordinary vector in the x-y plane. Our familiarity with this example makes it a good one to use to analyze in more detail what happens when we make rotations about different axes. Rather than working directly with the actual operators that perform these rotations in our quantum mechanical vector space, we will initially work in a specific representation and infer from the behavior that we see some fundamental properties about the operators themselves. The results we are

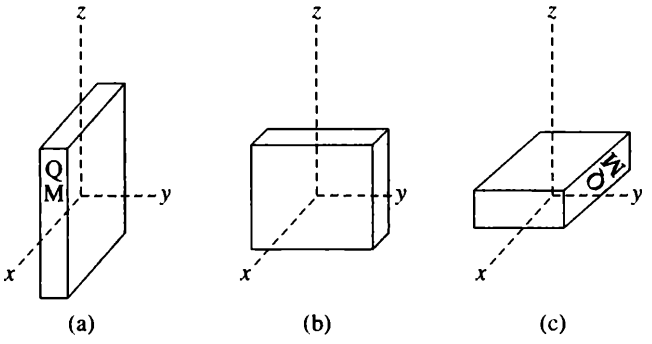


Figure 3.1 Noncommutativity of rotations. A book, shown in (a), is rotated in (b) by 90° around the x axis, then 90° about the y axis; in (c) the order of the rotations is reversed.

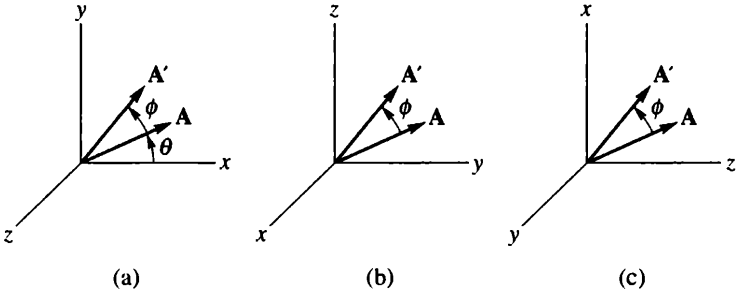


Figure 3.2 Rotating vector **A** into vector **A′** by angle ϕ counterclockwise about (a) the z axis, (b) the x axis, and (c) the y axis. For simplicity, only the components of the vector in the plane perpendicular to the axis of rotation are shown.

interested in depend on the three-dimensional structure of space and are properties that manifest themselves in all nontrivial representations.

Let's consider an ordinary three-dimensional vector **A** and a vector **A′** that is obtained by rotating **A** counterclockwise by an angle ϕ about the z axis. How are the components of **A** and **A′** related to each other? Denoting by θ the angle between the projection of **A** in the x-y plane and the x axis, as in Fig. 3.2a, we have

$$A'_x = \sqrt{A_x^2 + A_y^2}\,\cos(\phi+\theta) = \sqrt{A_x^2 + A_y^2}\,(\cos\phi\cos\theta - \sin\phi\sin\theta)$$
$$= A_x\cos\phi - A_y\sin\phi \tag{3.1a}$$

$$A'_y = \sqrt{A_x^2 + A_y^2}\,\sin(\phi+\theta) = \sqrt{A_x^2 + A_y^2}\,(\sin\phi\cos\theta + \sin\theta\cos\phi)$$
$$= A_x\sin\phi + A_y\cos\phi \tag{3.1b}$$

$$A'_z = A_z \tag{3.1c}$$

or, in matrix form,

$$\begin{pmatrix} A'_x \\ A'_y \\ A'_z \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \tag{3.2}$$

Thus the matrix that rotates the vector by angle ϕ counterclockwise about the z axis is given by

$$\mathbb{S}(\phi\mathbf{k}) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.3}$$

The 2×2 matrix in the upper left-hand corner is just the matrix (2.112). Because we are dealing here with a vector that has three components, the rotation matrix is a 3×3 matrix instead of the 2×2 matrix that we found for rotating polarization states. The additional elements in this matrix (3.3) simply show that the component of the vector in the z direction is unaffected by a rotation about the z axis.

We consider the special case where the angle is a small angle $\Delta\phi$ and retain terms in the Taylor series expansions for $\sin\Delta\phi$ and $\cos\Delta\phi$ through second order. It is necessary to work to at least this order to see the noncommutativity of the rotations. Thus

$$\mathbb{S}(\Delta\phi\mathbf{k}) = \begin{pmatrix} 1-\Delta\phi^2/2 & -\Delta\phi & 0 \\ \Delta\phi & 1-\Delta\phi^2/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.4}$$

From Fig. 3.2b we see that for a rotation about the x axis by angle ϕ, the matrix for the rotation can be obtained from the matrix (3.3) by letting $x \to y$, $y \to z$, and $z \to x$, that is, by a cyclic substitution. Therefore, the rotation matrix is

$$\mathbb{S}(\phi\mathbf{i}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} \tag{3.5}$$

and consequently

$$\mathbb{S}(\Delta\phi\mathbf{i}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-\Delta\phi^2/2 & -\Delta\phi \\ 0 & \Delta\phi & 1-\Delta\phi^2/2 \end{pmatrix} \tag{3.6}$$

Finally, we can obtain the matrix for a rotation about the y axis from the matrix for a rotation about the x axis by another cyclic substitution (see Fig. 3.2c). Thus

$$\mathbb{S}(\Delta\phi\mathbf{j}) = \begin{pmatrix} 1-\Delta\phi^2/2 & 0 & \Delta\phi \\ 0 & 1 & 0 \\ -\Delta\phi & 0 & 1-\Delta\phi^2/2 \end{pmatrix} \tag{3.7}$$

We now consider a rotation by $\Delta\phi$ about the y axis followed by a rotation by the same angle about the x axis. We subtract from it a rotation about the x axis followed by a rotation about the y axis. Multiplying the matrices (3.6) and (3.7), we obtain

$$\mathbb{S}(\Delta\phi\mathbf{i})\mathbb{S}(\Delta\phi\mathbf{j}) - \mathbb{S}(\Delta\phi\mathbf{j})\mathbb{S}(\Delta\phi\mathbf{i}) = \begin{pmatrix} 0 & -\Delta\phi^2 & 0 \\ \Delta\phi^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$= \mathbb{S}(\Delta\phi^2\mathbf{k}) - \mathbb{I} \tag{3.8}$$

where in the last step we have taken advantage of the explicit form of the matrix (3.4) when the rotation angle is $\Delta\phi^2$ and terms through order $\Delta\phi^2$ are retained.

From Section 2.5 we know that these S-matrices are the matrix representations of the rotation operators. For example, the matrix (3.3) is the representation of the rotation operator $\hat{R}(\phi\mathbf{k})$ in a particular basis.[1] Equation (3.8) shows that when we retain terms through second order in $\Delta\phi$, the operators themselves do not commute. Recall from (2.32) that the operator that rotates states by angle ϕ about the z axis is

$$\hat{R}(\phi\mathbf{k}) = e^{-i\hat{J}_z\phi/\hbar} \tag{3.9}$$

where $\hat{J}_z$ is the generator of rotations. We can think of this as a special case of the more general rotation operator

$$\hat{R}(\phi\mathbf{n}) = e^{-i\hat{\mathbf{J}}\cdot\mathbf{n}\phi/\hbar} \tag{3.10}$$

that rotates states by angle ϕ about the axis defined by the unit vector $\mathbf{n}$. Thus the operators that rotate states by angle ϕ about the x axis and the y axis are given by

$$\hat{R}(\phi\mathbf{i}) = e^{-i\hat{J}_x\phi/\hbar} \quad \text{and} \quad \hat{R}(\phi\mathbf{j}) = e^{-i\hat{J}_y\phi/\hbar} \tag{3.11}$$

with generators $\hat{J}_x$ and $\hat{J}_y$, respectively. Thus, if we take the angle of rotation to be the small angle $\Delta\phi$ and expand the rotation operators through second order in $\Delta\phi$, (3.8) tells us that

[1] Although we have phrased our discussion so far in terms of how ordinary vectors change under rotations, we are effectively using spin-1 states like the ones we saw in Section 2.7 as a basis, but with three states instead of just the two states that are necessary to describe photon polarization. We argued in that section that the way the photon polarization states changed under rotation told us that photons are spin-1 particles. If photons traveling in the z direction were to have a $|z\rangle$ polarization state as well as $|x\rangle$ and $|y\rangle$, this $|z\rangle$ polarization state would not be changed by performing a rotation about the z axis, and the matrix representation of the rotation operator $\hat{R}(\phi\mathbf{k})$ using the $|x\rangle$, $|y\rangle$, and $|z\rangle$ states as a basis would look like (3.3) instead of (2.113). Later in this chapter we will see how spin-1 states do form a three-dimensional basis. Again, particles like photons that move at c require special treatment.

$$\left\{1-\frac{i\hat{J}_x\Delta\phi}{\hbar}-\frac{1}{2}\left(\frac{\hat{J}_x\Delta\phi}{\hbar}\right)^2\right\}\left\{1-\frac{i\hat{J}_y\Delta\phi}{\hbar}-\frac{1}{2}\left(\frac{\hat{J}_y\Delta\phi}{\hbar}\right)^2\right\}$$
$$-\left\{1-\frac{i\hat{J}_y\Delta\phi}{\hbar}-\frac{1}{2}\left(\frac{\hat{J}_y\Delta\phi}{\hbar}\right)^2\right\}\left\{1-\frac{i\hat{J}_x\Delta\phi}{\hbar}-\frac{1}{2}\left(\frac{\hat{J}_x\Delta\phi}{\hbar}\right)^2\right\}$$
$$=\left(1-\frac{i\hat{J}_z\Delta\phi^2}{\hbar}\right)-1 \tag{3.12}$$

The lowest order nonvanishing terms involve $\Delta\phi^2$. Equating these terms, we obtain

$$\hat{J}_x\hat{J}_y-\hat{J}_y\hat{J}_x=i\hbar\hat{J}_z \tag{3.13}$$

or

$$[\hat{J}_x,\hat{J}_y]=i\hbar\hat{J}_z \tag{3.14a}$$

where the left-hand side of the equation is called the **commutator** of the two operators $\hat{J}_x$ and $\hat{J}_y$. The commutator of two operators is just the product of the two operators subtracted from the product of the two operators with the order of the operators reversed. Notice how Planck's constant enters on the right-hand side of (3.14a).

If we were to repeat this whole procedure for rotations about the y and z axes and for rotations about the z and x axes, we would obtain two other commutation relations related to (3.14a) by the cyclic permutation $x\rightarrow y$, $y\rightarrow z$, and $z\rightarrow x$:

$$[\hat{J}_y,\hat{J}_z]=i\hbar\hat{J}_x \tag{3.14b}$$

and

$$[\hat{J}_z,\hat{J}_x]=i\hbar\hat{J}_y \tag{3.14c}$$

It would be difficult to overemphasize the importance of these commutation relations. In Section 3.3 we will see that they alone are sufficient to determine the eigenstates and the eigenvalues of the angular momentum operators. So far, our arguments to establish that these generators of rotations should be identified with the angular momentum operators are probably at best suggestive. The proof is in the results and the comparison with experiment.

Later we will see that the *orbital* angular momentum operators

$$\hat{\mathbf{L}}=\hat{\mathbf{r}}\times\hat{\mathbf{p}} \tag{3.15}$$

also obey these same commutation relations, that is, for example,

$$[\hat{L}_x,\hat{L}_y]=i\hbar\hat{L}_z \tag{3.16}$$

However, we have not introduced angular momentum operators through (3.15), but rather simply as the generators of rotations. Although this approach may seem more abstract and initially less physical, it is also more general and, in fact, essential. In Chapter 9 we will see that the eigenvalues of orbital angular momentum, as defined by (3.15), do *not* include the half-integral values that characterize spin-$\frac{1}{2}$ particles such as electrons, protons, neutrons, and neutrinos.

3.2 Commuting Operators

The commutation relations of the generators of rotations show that the generators of rotations about different axes do not commute with each other. As we saw in Chapter 2, these generators are Hermitian operators. Before turning our attention toward solving the angular momentum eigenvalue problem, we need to ask what happens when two operators *do* commute. Consider two such linear Hermitian operators $\hat{A}$ and $\hat{B}$ that satisfy

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} = 0 \tag{3.17}$$

Suppose there exists only a single state $|a\rangle$ that is an eigenstate of $\hat{A}$ with eigenvalue a:

$$\hat{A}|a\rangle = a|a\rangle \tag{3.18}$$

If we apply the operator $\hat{B}$ to (3.18), we obtain

$$\hat{B}\hat{A}|a\rangle = \hat{B}a|a\rangle \tag{3.19}$$

On the left-hand side we take advantage of (3.17) and on the right-hand side we take advantage of the fact that $\hat{B}$ is a linear operator to write

$$\hat{A}\hat{B}|a\rangle = a\hat{B}|a\rangle \tag{3.20a}$$

or

$$\hat{A}(\hat{B}|a\rangle) = a(\hat{B}|a\rangle) \tag{3.20b}$$

where we have inserted the parentheses to isolate the state $\hat{B}|a\rangle$ on both sides. Equation (3.20) says that the state $\hat{B}|a\rangle$ is an eigenstate of the operator $\hat{A}$ with eigenvalue a. Since we have presumed there is only one such state, we conclude that

$$\hat{B}|a\rangle = b|a\rangle \tag{3.21}$$

where b is a constant, since if $|a\rangle$ satisfies (3.18), so does $b|a\rangle$ for any constant b. But (3.21) says that $|a\rangle$ is an eigenstate of $\hat{B}$ as well with eigenvalue b. Therefore,

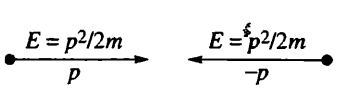


Figure 3.3 A free particle with momentum p has the same energy as one with momentum $-p$.

we can relabel the state $|a\rangle$ as $|a, b\rangle$ to show both of the eigenvalues and say that $\hat{A}$ and $\hat{B}$ have the eigenstate $|a, b\rangle$ in common. An example of a state that can be labeled by two eigenvalues is the state $|E, p\rangle$ of a *free* particle in one dimension, where E is the energy and p is the momentum of the particle.

If there is more than one eigenstate of the operator $\hat{A}$ with eigenvalue a, we say that there is **degeneracy**. Our proof has established that each eigenstate of $\hat{A}$ is also an eigenstate of $\hat{B}$ for those states that are not degenerate. If there is degeneracy, one can always find linear combinations of the degenerate eigenstates of $\hat{A}$ that are eigenstates of the Hermitian operator $\hat{B}$. Thus two Hermitian operators that commute have a complete set of eigenstates in common. This result follows from the fundamental spectral theorem of linear algebra. We will not prove it here, but we will have a number of opportunities in later chapters to verify that it holds in special cases. In fact, the example of the one-dimensional free particle can serve as an illustration, since for a particular energy $E = p^2/2m$ there is two-fold degeneracy: the states $|E, p\rangle$ and $|E, -p\rangle$ have the same energy but momenta p and $-p$, respectively, corresponding to a particle moving to the right or the left (see Fig. 3.3). Note that you can certainly form states that are superpositions of the states $|E, p\rangle$ and $|E, -p\rangle$ (such as standing waves), so states with a definite energy need not have a definite momentum.

EXAMPLE 3.1 Equation (2.113) gives the matrix representation

$$\mathbb{S} = \begin{pmatrix} \langle x|\hat{R}(\phi\mathbf{k})|x\rangle & \langle x|\hat{R}(\phi\mathbf{k})|y\rangle \\ \langle y|\hat{R}(\phi\mathbf{k})|x\rangle & \langle y|\hat{R}(\phi\mathbf{k})|y\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

of the rotation operator $\hat{R}(\phi\mathbf{k})$ using the linear polarization vectors $|x\rangle$ and $|y\rangle$ for photons as a basis. Example 2.8 shows that

$$\hat{J}_z \to \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

in the same basis. Show that these operators commute and therefore have eigenstates in common. What are these eigenstates and what are the matrix representations for $\hat{R}(\phi\mathbf{k})$ and $\hat{J}_z$ using these eigenstates as a basis?

SOLUTION It is straightforward to verify that these operators commute:

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} = 0$$

We know from Section 2.7 that the eigenstates of $\hat{J}_z$ are the circular polarization states $|R\rangle$ and $|L\rangle$ with eigenvalues $\hbar$ and $-\hbar$, respectively. Consequently, as given in Example 2.8,

$$\begin{pmatrix} \langle R|\hat{J}_z|R\rangle & \langle R|\hat{J}_z|L\rangle \\ \langle L|\hat{J}_z|R\rangle & \langle L|\hat{J}_z|L\rangle \end{pmatrix} = \begin{pmatrix} \hbar & 0 \\ 0 & -\hbar \end{pmatrix}$$

Since $\hat{R}(\phi\mathbf{k}) = e^{-i\hat{J}_z\phi/\hbar}$, we also see that

$$\begin{pmatrix} \langle R|\hat{R}(\phi\mathbf{k})|R\rangle & \langle R|\hat{R}(\phi\mathbf{k})|L\rangle \\ \langle L|\hat{R}(\phi\mathbf{k})|R\rangle & \langle L|\hat{R}(\phi\mathbf{k})|L\rangle \end{pmatrix} = \begin{pmatrix} e^{-i\phi} & 0 \\ 0 & e^{i\phi} \end{pmatrix}$$

consistent with the fact that these two operators have the eigenstates $|R\rangle$ and $|L\rangle$ in common. Using these eigenstates as a basis, the matrix representations of both operators are diagonal with the corresponding eigenvalues as the diagonal matrix elements.

3.3 The Eigenvalues and Eigenstates of Angular Momentum

Although the commutation relations (3.14) show us that the generators of rotations about different axes do not commute with each other, the operator

$$\hat{\mathbf{J}}^2 = \hat{\mathbf{J}} \cdot \hat{\mathbf{J}} = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \tag{3.22}$$

does commute with each of the generators.[2] In order to verify this, we choose $\hat{J}_z$, the generator of rotations about the z axis, and use the identity (see Problem 3.1)

$$[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} \tag{3.23}$$

to obtain[3]

$$\begin{aligned} [\hat{J}_z, \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2] &= [\hat{J}_z, \hat{J}_x^2] + [\hat{J}_z, \hat{J}_y^2] + [\hat{J}_z, \hat{J}_z^2] \\ &= \hat{J}_x[\hat{J}_z, \hat{J}_x] + [\hat{J}_z, \hat{J}_x]\hat{J}_x + \hat{J}_y[\hat{J}_z, \hat{J}_y] + [\hat{J}_z, \hat{J}_y]\hat{J}_y \\ &= i\hbar(\hat{J}_x\hat{J}_y + \hat{J}_y\hat{J}_x - \hat{J}_y\hat{J}_x - \hat{J}_x\hat{J}_y) = 0 \end{aligned} \tag{3.24}$$

---

[2] The operator $\hat{\mathbf{J}} = \hat{J}_x\mathbf{i} + \hat{J}_y\mathbf{j} + \hat{J}_z\mathbf{k}$ is a **vector operator**. For vector operators such as $\hat{\mathbf{J}}$ we use the notation $\hat{\mathbf{J}}^2 = (\hat{J}_x\mathbf{i} + \hat{J}_y\mathbf{j} + \hat{J}_z\mathbf{k}) \cdot (\hat{J}_x\mathbf{i} + \hat{J}_y\mathbf{j} + \hat{J}_z\mathbf{k}) = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$.

[3] We will use commutator identity (3.23) as well as its analogue $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$ often when evaluating a commutator that involves a product of operators. In general, this is much easier than starting by expanding the commutator using the defining relationship $[\hat{A}, \hat{B}\hat{C}] = \hat{A}\hat{B}\hat{C} - \hat{B}\hat{C}\hat{A}$. You are encouraged to work out Problem 3.1 so you feel comfortable with these commutator identities.

Because the operator $\hat{\mathbf{J}}^2$ commutes with $\hat{J}_z$, these operators have simultaneous eigenstates in common. We label the kets $|\lambda, m\rangle$, where

$$\hat{\mathbf{J}}^2|\lambda, m\rangle = \lambda\hbar^2|\lambda, m\rangle \tag{3.25a}$$

$$\hat{J}_z|\lambda, m\rangle = m\hbar|\lambda, m\rangle \tag{3.25b}$$

We have explicitly included the dimensions of the operators in the factors of $\hbar$ so that λ and m are dimensionless. Thus $|\lambda, m\rangle$ is a state for which a measurement of the z component of the angular momentum yields the value $m\hbar$ and the magnitude squared of the angular momentum is $\lambda\hbar^2$.

We can see that $\lambda \geq 0$, as we would expect physically since λ specifies the magnitude squared of the angular momentum in the state $|\lambda, m\rangle$. Consider

$$\langle\lambda, m|\hat{\mathbf{J}}^2|\lambda, m\rangle = \lambda\hbar^2\langle\lambda, m|\lambda, m\rangle \tag{3.26}$$

Like all physical states, the eigenstates satisfy $\langle\lambda, m|\lambda, m\rangle = 1$. A typical term in the left-hand side of (3.26) is of the form

$$\langle\lambda, m|\hat{J}_x^2|\lambda, m\rangle = \langle\psi|\psi\rangle \tag{3.27}$$

where we have defined $\hat{J}_x|\lambda, m\rangle = |\psi\rangle$, and $\langle\psi| = \langle\lambda, m|\hat{J}_x$ since $\hat{J}_x$ is Hermitian. Although the ket $|\psi\rangle$ is not normalized, we can always write it as $|\psi\rangle = c|\varphi\rangle$, where c is a complex constant (that must have the dimensions of $\hbar$) and $|\varphi\rangle$ is a physical state satisfying $\langle\varphi|\varphi\rangle = 1$. In other words, the action of the operator $\hat{J}_x$ on a ket vector must yield another ket vector that belongs to the vector space.[4] Since $\langle\psi| = c^*\langle\varphi|$, we see that $\langle\psi|\psi\rangle = c^*c\langle\varphi|\varphi\rangle \geq 0$, where the equality would hold if $c = 0$. Our argument that (3.27) is positive semidefinite holds for each of the three pieces [see the form (3.22) of $\hat{\mathbf{J}}^2$] on the left-hand side of (3.26), and therefore $\lambda \geq 0$.

AN EXAMPLE: SPIN 1

To illustrate what we have discovered so far and suggest the next step, let's take the specific example involving the following three 3×3 matrices:

$$\hat{J}_x \to \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \hat{J}_y \to \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \hat{J}_z \to \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{3.28}$$

[4] Because $\hat{J}_x$ is the generator of rotations about the x axis, the ket $(1 - i\hat{J}_x d\phi/\hbar)|\lambda, m\rangle$ is just the ket that is produced by rotating the ket $|\lambda, m\rangle$ by angle $d\phi$ about the x axis. Thus the ket $|\psi\rangle$ can be viewed as a linear combination of the rotated ket and the ket $|\lambda, m\rangle$, that is, a superposition of two physical states.

For now, don't worry about how we have obtained these matrices. Later in this chapter we will see how we can deduce the form of these matrices (see Example 3.3 and Problem 3.14). In the meantime, let's see what we can learn from the matrices themselves.

To begin, how can we be sure that these three matrices really *represent* angular momentum operators? Following our earlier discussion, it is sufficient to check (see Problem 3.13) that these matrices do indeed satisfy the commutation relations (3.14). We next calculate

$$\hat{\mathbf{J}}^2 = \hat{\mathbf{J}} \cdot \hat{\mathbf{J}} = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \to 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.29}$$

We see explicitly that $\hat{\mathbf{J}}^2$ is just a constant times the identity matrix and thus commutes with each of the components of $\hat{\mathbf{J}}$. The operator $\hat{J}_z$ is diagonal as well, suggesting that the matrix representations (3.28) are formed using the eigenstates of $\hat{J}_z$ as well as $\hat{\mathbf{J}}^2$ as a basis. The column vectors representing these eigenstates are given by[5]

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{3.30}$$

which have eigenvalues $\hbar$, 0, and $-\hbar$, respectively, as can be verified by operating on them with the matrix representing $\hat{J}_z$. For example,

$$\hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \hbar \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{3.31}$$

Similarly, we see that each of these states is an eigenstate of $\hat{\mathbf{J}}^2$ with eigenvalue $2\hbar^2$.

Since the matrix representations of $\hat{J}_x$ and $\hat{J}_y$ are not diagonal, the states (3.30) are not eigenstates of these operators. It is straightforward to evaluate the action of the operators $\hat{J}_x$ and $\hat{J}_y$ on the basis states. There is, however, a linear combination of these two operators, namely,

$$\hat{J}_x + i\hat{J}_y \to \sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.32}$$

whose action on the basis states exhibits an interesting pattern. Applying this operator to the basis states (3.30), we obtain

[5] Compare these results with (2.70), (2.71), and (2.72) for a spin-$\frac{1}{2}$ particle.

$$\sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \sqrt{2}\hbar \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{3.33}$$

$$\sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \sqrt{2}\hbar \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{3.34}$$

$$\sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{3.35}$$

Thus, according to (3.33), the operator $\hat{J}_x + i\hat{J}_y$ acting on the state with eigenvalue $-\hbar$ for $\hat{J}_z$ turns it into a state with eigenvalue 0, multiplied by $\sqrt{2}\hbar$. Similarly, as (3.34) shows, when the operator acts on the state with eigenvalue 0 for $\hat{J}_z$, it turns it into a state with eigenvalue $\hbar$, multiplied by $\sqrt{2}\hbar$. This raising action terminates when the operator $\hat{J}_x + i\hat{J}_y$ acts on the state with eigenvalue $\hbar$, the maximum eigenvalue for $\hat{J}_z$. See (3.35). It can be similarly verified that the operator

$$\hat{J}_x - i\hat{J}_y \rightarrow \sqrt{2}\hbar \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{3.36}$$

has a lowering action when it acts on the states with eigenvalues $\hbar$ and 0, turning them into states with eigenvalues 0 and $-\hbar$, respectively. In this case, the lowering action terminates when the operator (3.36) acts on the state with eigenvalue $-\hbar$, the lowest eigenvalue for $\hat{J}_z$.

RAISING AND LOWERING OPERATORS

Let's return to our general analysis of angular momentum. The example suggests that it is convenient to introduce the two operators

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y \tag{3.37}$$

in the general case. Notice that these are not Hermitian operators since

$$\hat{J}_+^\dagger = \hat{J}_x^\dagger + (-i)\hat{J}_y^\dagger = \hat{J}_x - i\hat{J}_y = \hat{J}_- \tag{3.38}$$

The utility of these operators derives from their commutation relations with $\hat{J}_z$:

$$[\hat{J}_z, \hat{J}_\pm] = [\hat{J}_z, \hat{J}_x \pm i\hat{J}_y] = i\hbar\hat{J}_y \pm i(-i\hbar\hat{J}_x) = \pm\hbar\hat{J}_\pm \tag{3.39}$$

To see the effect of $\hat{J}_+$ on the eigenstates, we evaluate $\hat{J}_z\hat{J}_+|\lambda, m\rangle$. We can use the commutation relation (3.39) to invert the order of the operators so that $\hat{J}_z$ can act

directly on its eigenstate $|\lambda, m\rangle$. However, since the commutator of $\hat{J}_z$ and $\hat{J}_+$ is not zero but rather is proportional to the operator $\hat{J}_+$ itself, we pick up an additional contribution:

$$\begin{aligned}\hat{J}_z\hat{J}_+|\lambda, m\rangle &= (\hat{J}_+\hat{J}_z + \hbar\hat{J}_+)|\lambda, m\rangle \\ &= (\hat{J}_+ m\hbar + \hbar\hat{J}_+)|\lambda, m\rangle \\ &= (m+1)\hbar\hat{J}_+|\lambda, m\rangle \end{aligned} \tag{3.40a}$$

Inserting some parentheses to help guide the eye:

$$\hat{J}_z(\hat{J}_+|\lambda, m\rangle) = (m+1)\hbar(\hat{J}_+|\lambda, m\rangle) \tag{3.40b}$$

we see that $\hat{J}_+|\lambda, m\rangle$ is an eigenstate of $\hat{J}_z$ with eigenvalue $(m+1)\hbar$. Hence $\hat{J}_+$ is referred to as a **raising operator**. The action of $\hat{J}_+$ on the state $|\lambda, m\rangle$ is to produce a new state with eigenvalue $(m+1)\hbar$.

Also

$$\begin{aligned}\hat{J}_z\hat{J}_-|\lambda, m\rangle &= (\hat{J}_-\hat{J}_z - \hbar\hat{J}_-)|\lambda, m\rangle \\ &= (\hat{J}_- m\hbar - \hbar\hat{J}_-)|\lambda, m\rangle \\ &= (m-1)\hbar\hat{J}_-|\lambda, m\rangle \end{aligned} \tag{3.41a}$$

Again, inserting some parentheses,

$$\hat{J}_z(\hat{J}_-|\lambda, m\rangle) = (m-1)\hbar(\hat{J}_-|\lambda, m\rangle) \tag{3.41b}$$

showing that $\hat{J}_-|\lambda, m\rangle$ is an eigenstate of $\hat{J}_z$ with eigenvalue $(m-1)\hbar$; hence $\hat{J}_-$ is a **lowering operator**. Notice that since $\hat{J}_+$ and $\hat{J}_-$ commute with $\hat{\mathbf{J}}^2$, the states $\hat{J}_\pm|\lambda, m\rangle$ are still eigenstates of the operator $\hat{\mathbf{J}}^2$ with eigenvalue $\lambda\hbar^2$:

$$\hat{\mathbf{J}}^2(\hat{J}_\pm|\lambda, m\rangle) = \hat{J}_\pm\hat{\mathbf{J}}^2|\lambda, m\rangle = \lambda\hbar^2(\hat{J}_\pm|\lambda, m\rangle) \tag{3.42}$$

THE EIGENVALUE SPECTRUM

We now have enough information to determine the eigenvalues λ and m, because there are bounds on how far we can raise or lower m. Physically (see Fig. 3.4), we expect that the square of the projection of the angular momentum on any axis should not exceed the magnitude of $\mathbf{J}^2$ and hence

$$m^2 \leq \lambda \tag{3.43}$$

Formally, since

$$\langle\lambda, m|(\hat{J}_x^2 + \hat{J}_y^2)|\lambda, m\rangle \geq 0 \tag{3.44}$$

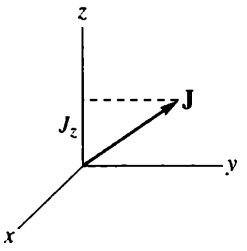


Figure 3.4 The projection of the angular momentum on the axis never exceeds the magnitude of the angular momentum. *Caution:* This is a classical picture; the angular momentum cannot point in any definite direction.

we have

$$\langle \lambda, m|(\hat{\mathbf{J}}^2 - \hat{J}_z^2)|\lambda, m\rangle = (\lambda - m^2)\hbar^2\langle \lambda, m|\lambda, m\rangle \geq 0 \tag{3.45}$$

establishing (3.43).

Let's call the *maximum m* value j. Then we must have

$$\hat{J}_+|\lambda, j\rangle = 0 \tag{3.46}$$

since otherwise $\hat{J}_+$ would create a state $|\lambda, j+1\rangle$, violating our assumption that j is the maximum eigenvalue for $\hat{J}_z$.[6] Using

$$\begin{aligned}\hat{J}_-\hat{J}_+ &= (\hat{J}_x - i\hat{J}_y)(\hat{J}_x + i\hat{J}_y)\\ &= \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_x, \hat{J}_y]\\ &= \hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z\end{aligned} \tag{3.47}$$

we see that

$$\begin{aligned}\hat{J}_-\hat{J}_+|\lambda, j\rangle &= (\hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|\lambda, j\rangle\\ &= (\lambda - j^2 - j)\hbar^2|\lambda, j\rangle = 0\end{aligned} \tag{3.48}$$

or $\lambda = j(j+1)$.

Similarly, if we call the *minimum m* value j', then

$$\hat{J}_-|\lambda, j'\rangle = 0 \tag{3.49}$$

and we find that

$$\begin{aligned}\hat{J}_+\hat{J}_-|\lambda, j'\rangle &= (\hat{\mathbf{J}}^2 - \hat{J}_z^2 + \hbar\hat{J}_z)|\lambda, j'\rangle\\ &= (\lambda - j'^2 + j')\hbar^2|\lambda, j'\rangle = 0\end{aligned} \tag{3.50}$$

[6] Equation (3.35) demonstrates how this works for the special case of spin 1.

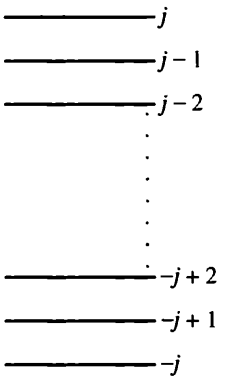


Figure 3.5 The possible m values for a fixed magnitude $\sqrt{j(j+1)}\hbar$ of the angular momentum.

In deriving this result, we have used

$$\begin{aligned}\hat{J}_+\hat{J}_- &= (\hat{J}_x + i\hat{J}_y)(\hat{J}_x - i\hat{J}_y)\\ &= \hat{J}_x^2 + \hat{J}_y^2 - i[\hat{J}_x, \hat{J}_y]\\ &= \hat{\mathbf{J}}^2 - \hat{J}_z^2 + \hbar\hat{J}_z\end{aligned} \tag{3.51}$$

Thus $\lambda = j'^2 - j'$. The solutions to the equation $j^2 + j = j'^2 - j'$, which results from setting these two values of λ equal to each other, are $j' = -j$ and $j' = j + 1$. The second solution violates our assumption that the maximum m value is j. Thus we find the minimum m value is $-j$.

If we start at the $m = j$ state, the state with the maximum m, and apply the lowering operator a sufficient number of times, we must reach the state with $m = -j$, the state with the minimum m. If this were not the case, we would either reach a state with an m value not equal to $-j$ for which (3.49) is satisfied or we would violate the bound on the m values. But (3.49) determines uniquely the value of j' to be $-j$. Since we lower an integral number of times, $j - j' = j - (-j) = 2j =$ an integer, and we deduce that the allowed values of j are given by

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \tag{3.52}$$

As indicated in Fig. 3.5, the m values for each j run from j to $-j$ in integral steps:

$$m = \underbrace{j, j-1, j-2, \ldots, -j+1, -j}_{2j+1 \text{ states}} \tag{3.53}$$

Given these results, we now change our notation slightly. It is conventional to denote a simultaneous eigenstate of the operators $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ by $|j, m\rangle$ instead of $|\lambda, m\rangle = |j(j+1), m\rangle$. It is important to remember in this shorthand notation that

$$\hat{\mathbf{J}}^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle \tag{3.54a}$$

as well as

$$\hat{J}_z|j, m\rangle = m\hbar|j, m\rangle \tag{3.54b}$$

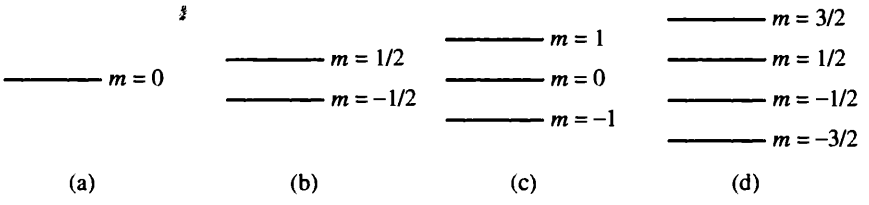


Figure 3.6 The m values for (a) spin 0, (b) spin $\frac{1}{2}$, (c) spin 1, and (d) spin $\frac{3}{2}$.

Let's examine a few of these states, for which the m values are shown in Fig. 3.6.

1. The $j = 0$ state is denoted by $|0, 0\rangle$. Since the magnitude of the angular momentum is zero for this state, it is not surprising that the projection of the angular momentum on the z axis vanishes as well.
2. The $j = \frac{1}{2}$ states are given by $|\frac{1}{2}, \frac{1}{2}\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle$. Note that the eigenvalues of $\hat{J}_z$ for these states are $\hbar/2$ and $-\hbar/2$, respectively. These states are just the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ that have concerned us for much of Chapters 1 and 2. We now see the rationale for calling these states spin-$\frac{1}{2}$ states: the constant j takes on the value $\frac{1}{2}$. However, the magnitude of the spin of the particle in these states is given by $\sqrt{\frac{1}{2}(\frac{1}{2}+1)}\,\hbar = \sqrt{3}\,\hbar/2$.
3. The angular momentum $j = 1$ states are denoted by $|1, 1\rangle$, $|1, 0\rangle$, and $|1, -1\rangle$. These spin-1 states are represented by the column vectors (3.30) in the example of this section. The eigenvalues of $\hat{J}_z$ are $\hbar$, 0, and $-\hbar$, which are the diagonal matrix elements of the matrix representing $\hat{J}_z$ in (3.28). The magnitude of the angular momentum for these states is given by $\sqrt{1(1+1)}\,\hbar = \sqrt{2}\,\hbar$.
4. There are four $j = \frac{3}{2}$ states: $|\frac{3}{2}, \frac{3}{2}\rangle$, $|\frac{3}{2}, \frac{1}{2}\rangle$, $|\frac{3}{2}, -\frac{1}{2}\rangle$, and $|\frac{3}{2}, -\frac{3}{2}\rangle$. The magnitude of the angular momentum is $\sqrt{\frac{3}{2}(\frac{3}{2}+1)}\,\hbar = \sqrt{15}\,\hbar/2$.

As these examples illustrate, the magnitude $\sqrt{j(j+1)}\,\hbar$ of the angular momentum is always bigger than the maximum projection $j\hbar$ on the z axis for any nonzero angular momentum. In Section 3.5 we will see how the uncertainty relations for angular momentum allow us to understand why the angular momentum does not line up along an axis.

EXAMPLE 3.2 An atom passes straight through an SG**z** device without deflecting. What can you deduce about the angular momentum of the atom?

SOLUTION Since the atom is not deflected, it must have $J_z = 0$. Thus the atom has an integral value j for its angular momentum, since only for integral values of j is $m = 0$ one of the eigenvalues for $\hat{J}_z$.

3.4 The Matrix Elements of the Raising and Lowering Operators

We have seen in (3.40) and (3.42) that the action of the raising operator $\hat{J}_+$ on a state of angular momentum j is to create a state with the same magnitude of the angular momentum but with the z component increased by one unit of $\hbar$:

$$\hat{J}_+|j, m\rangle = c_+\hbar|j, m+1\rangle \tag{3.55}$$

while the action of the lowering operator is

$$\hat{J}_-|j, m\rangle = c_-\hbar|j, m-1\rangle \tag{3.56}$$

It is useful to determine the values of c_+ and c_-. Taking the inner product of the ket (3.55) with the corresponding bra and making use of (3.38), we obtain

$$\langle j, m|\hat{J}_-\hat{J}_+|j, m\rangle = c_+^* c_+ \hbar^2 \langle j, m+1|j, m+1\rangle \tag{3.57}$$

Substituting (3.47) for the operators $\hat{J}_-\hat{J}_+$, we find

$$\begin{aligned}\langle j, m|(\hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|j, m\rangle &= [j(j+1) - m^2 - m]\hbar^2\langle j, m|j, m\rangle \\ &= c_+^* c_+ \hbar^2 \langle j, m+1|j, m+1\rangle \end{aligned} \tag{3.58}$$

Assuming the angular momentum states satisfy $\langle j, m|j, m\rangle = \langle j, m+1|j, m+1\rangle$, we can choose $c_+ = \sqrt{j(j+1) - m(m+1)}$, or

$$\hat{J}_+|j, m\rangle = \sqrt{j(j+1) - m(m+1)}\,\hbar|j, m+1\rangle \tag{3.59}$$

Note that when $m = j$, the square root factor vanishes and the raising action terminates, as it must. Similarly, we can establish that

$$\hat{J}_-|j, m\rangle = \sqrt{j(j+1) - m(m-1)}\,\hbar|j, m-1\rangle \tag{3.60}$$

for which the square root factor vanishes when $m = -j$, as it must.

These results determine the matrix elements of the raising and lowering operators using the states $|j, m\rangle$ as a basis:

$$\begin{aligned}\langle j, m'|\hat{J}_+|j, m\rangle &= \sqrt{j(j+1) - m(m+1)}\,\hbar\,\langle j, m'|j, m+1\rangle \\ &= \sqrt{j(j+1) - m(m+1)}\,\hbar\,\delta_{m',m+1}\end{aligned} \tag{3.61}$$

and

$$\begin{aligned}\langle j, m'|\hat{J}_-|j, m\rangle &= \sqrt{j(j+1) - m(m-1)}\,\hbar\,\langle j, m'|j, m-1\rangle \\ &= \sqrt{j(j+1) - m(m-1)}\,\hbar\,\delta_{m',m-1}\end{aligned} \tag{3.62}$$

In obtaining these matrix elements, we have made use of $\langle j, m'|j, m\rangle = \delta_{m',m}$, since the amplitude to find a state having $J_z = m\hbar$ with $J_z = m'\hbar$, $m' \neq m$, is zero. In

Section 3.6 we will see how useful the matrix elements (3.61) and (3.62) are for obtaining matrix representations of $\hat{J}_x$ and $\hat{J}_y$.

EXAMPLE 3.3 Obtain the matrix representation of the raising and lowering operators using the $j = 1$ states as a basis.

SOLUTION The three $j = 1$ basis states are $|1\rangle = |1, 1\rangle$, $|2\rangle = |1, 0\rangle$, and $|3\rangle = |1, -1\rangle$. Using (3.61), we see that $\hat{J}_+|1, 1\rangle = 0$, $\hat{J}_+|1, 0\rangle = \sqrt{2}\,\hbar|1, 1\rangle$, and $\hat{J}_+|1, -1\rangle = \sqrt{2}\,\hbar|1, 0\rangle$. Thus the only nonzero matrix elements are $\langle 1|\hat{J}_+|2\rangle = \langle 1, 1|\hat{J}_+|1, 0\rangle = \sqrt{2}\,\hbar$ in the first row, second column and $\langle 2|\hat{J}_+|3\rangle = \langle 1, 0|\hat{J}_+|1, -1\rangle = \sqrt{2}\,\hbar$ in the second row, third column:

$$\hat{J}_+ \xrightarrow[J_z \text{ basis}]{} \sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Since $\hat{J}_- = \hat{J}_+^\dagger$, the matrix representation of $\hat{J}_-$ is the transpose, complex conjugate of the matrix representation for $\hat{J}_+$:

$$\hat{J}_- \xrightarrow[J_z \text{ basis}]{} \sqrt{2}\hbar \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

These results are in agreement with (3.32) and (3.36), showing that the 3×3 matrix representations in Section 3.3 are indeed those for $j = 1$.

3.5 Uncertainty Relations and Angular Momentum

In solving the angular momentum problem in Section 3.3, we took advantage of the commutation relation (3.24) to form simultaneous eigenstates of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$. Since $[\hat{\mathbf{J}}^2, \hat{J}_x] = 0$ as well, we can also form simultaneous eigenstates of $\hat{\mathbf{J}}^2$ and $\hat{J}_x$. For the $j = \frac{1}{2}$ sector, the two eigenstates would be the states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ that we discussed in the earlier chapters. We did not, however, try to form simultaneous eigenstates of $\hat{\mathbf{J}}^2$, $\hat{J}_x$, and $\hat{J}_z$. We now want to show that such simultaneous eigenstates are prohibited by the commutation relations of the angular momentum operators themselves, such as

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z \tag{3.63}$$

This is why in Section 3.3 we chose *only one* of the components of $\hat{\mathbf{J}}$, together with the operator $\hat{\mathbf{J}}^2$, to label the eigenstates.

The commutation relation (3.63) is an example of two operators that do not commute and whose commutator can be expressed in the form

$$[\hat{A}, \hat{B}] = i\hat{C} \tag{3.64}$$

where $\hat{A}$, $\hat{B}$, and $\hat{C}$ are Hermitian operators. We will now demonstrate that a commutation relation of the form (3.64) implies a fundamental uncertainty relation. To derive the uncertainty relation, we use the Schwarz inequality

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2 \tag{3.65}$$

This is the analogue of the relation $(\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b}) \geq (\mathbf{a}\cdot\mathbf{b})^2$, familiar from the ordinary real three-dimensional vector space. See Problem 3.7 for a derivation of (3.65).

We substitute

$$|\alpha\rangle = (\hat{A} - \langle A\rangle)|\psi\rangle \tag{3.66a}$$

$$|\beta\rangle = (\hat{B} - \langle B\rangle)|\psi\rangle \tag{3.66b}$$

into (3.65), where the expectation values

$$\langle A\rangle = \langle\psi|\hat{A}|\psi\rangle \tag{3.67a}$$

and

$$\langle B\rangle = \langle\psi|\hat{B}|\psi\rangle \tag{3.67b}$$

are real numbers because the operators are Hermitian. Notice that

$$\langle\alpha|\alpha\rangle = \langle\psi|(\hat{A} - \langle A\rangle)^2|\psi\rangle = (\Delta A)^2 \tag{3.68a}$$

$$\langle\beta|\beta\rangle = \langle\psi|(\hat{B} - \langle B\rangle)^2|\psi\rangle = (\Delta B)^2 \tag{3.68b}$$

where we have used the familiar definition of the uncertainty (see Section 1.4 or Section 1.6) and the fact that $\hat{A}$ and $\hat{B}$ are Hermitian operators. The right-hand side of the Schwarz inequality (3.65) for the states (3.66) becomes

$$\langle\alpha|\beta\rangle = \langle\psi|(\hat{A} - \langle A\rangle)(\hat{B} - \langle B\rangle)|\psi\rangle \tag{3.69}$$

For any operator $\hat{O}$, we may write

$$\hat{O} = \frac{\hat{O} + \hat{O}^\dagger}{2} + \frac{\hat{O} - \hat{O}^\dagger}{2} = \frac{\hat{F}}{2} + \frac{i\hat{G}}{2} \tag{3.70}$$

where $\hat{F} = \hat{O} + \hat{O}^\dagger$ and $\hat{G} = -i(\hat{O} - \hat{O}^\dagger)$ are Hermitian operators. If we take the operator $\hat{O}$ to be $(\hat{A} - \langle A\rangle)(\hat{B} - \langle B\rangle)$, we find

$$\hat{O} - \hat{O}^\dagger = [\hat{A}, \hat{B}] = i\hat{C} \tag{3.71}$$

and therefore $\hat{G} = \hat{C}$ in (3.70). Thus

$$|\langle\alpha|\beta\rangle|^2 = \left|\frac{1}{2}\langle\psi|\hat{F}|\psi\rangle + \frac{i}{2}\langle\psi|\hat{C}|\psi\rangle\right|^2$$
$$= \frac{|\langle\psi|\hat{F}|\psi\rangle|^2}{4} + \frac{|\langle\psi|\hat{C}|\psi\rangle|^2}{4} \geq \frac{|\langle C\rangle|^2}{4} \tag{3.72}$$

where we have made use of the fact that the expectation values of the Hermitian operators $\hat{F}$ and $\hat{C}$ are real. Combining (3.65), (3.68), and (3.72), we obtain

$$(\Delta A)^2(\Delta B)^2 \geq \frac{|\langle C\rangle|^2}{4} \tag{3.73}$$

or simply

$$\Delta A \Delta B \geq \frac{|\langle C\rangle|}{2} \tag{3.74}$$

which is a very important result.

If we apply this uncertainty relation to the specific commutation relation (3.63), we find[7]

$$\Delta J_x \Delta J_y \geq \frac{\hbar}{2}|\langle J_z\rangle| \tag{3.75}$$

This uncertainty relation helps to explain a number of our earlier results. If a spin-$\frac{1}{2}$ particle is in a state with a definite value of J_z, $\langle J_z\rangle$ is either $\hbar/2$ or $-\hbar/2$, which is certainly nonzero. But (3.75) says that ΔJ_x must then also be nonzero, and thus the particle cannot have a definite value of J_x when it has a definite value of J_z. We now see why making a measurement of S_z in the Stern–Gerlach experiments is bound to modify subsequent measurements of S_x. We cannot know both the x and the z components of the angular momentum of the particle with definite certainty. We can also see why in general the angular momentum doesn't line up along any axis: If the angular momentum were aligned completely along the z axis, both the x and y components of the angular momentum would vanish. We would then know all three components of the angular momentum, in disagreement with the uncertainty relation (3.75), which requires that both ΔJ_x and ΔJ_y are nonzero in a state with a definite nonzero value of J_z. Thus the angular momentum never really "points" in any definite direction.

---

[7] In Chapter 6 we will see that the position and momentum operators satisfy

$$[\hat{x}, \hat{p}_x] = i\hbar$$

Thus (3.74) leads directly to the famous Heisenberg uncertainty relation $\Delta x \Delta p_x \geq \hbar/2$ as well.

3.6 The Spin-$\frac{1}{2}$ Eigenvalue Problem

In this section we will see how we can use the results of this chapter to derive the spin states of a spin-$\frac{1}{2}$ particle that we deduced from the results of Stern–Gerlach experiments in Chapter 1. First we will make a small change in notation. It is customary in discussing angular momentum to call the angular momentum operators $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$ in general. We have introduced these operators as the generators of rotations. The commutation relations that we used in Section 3.3 depended only on the fact that rotations about different axes do not commute in a well-defined way. Our formulation is general enough to include all kinds of angular momentum, both intrinsic spin angular momentum and orbital angular momentum. That is one of the major virtues of introducing angular momentum in this way. In Chapter 9 we will see that for *orbital* angular momentum—angular momentum of the $\mathbf{r} \times \mathbf{p}$ type—only integral j's are permitted. If our discussion of angular momentum is restricted to purely orbital angular momentum, it is conventional to denote the angular momentum operators by $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$. On the other hand, if our discussion is restricted to intrinsic spin angular momentum, it is customary to call the spin angular momentum operators $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$. Our discussion in Chapters 1 and 2 of the intrinsic spin angular momentum of particles like electrons and photons was restricted to angular momentum of the latter sort. Thus, we could return to Chapter 2, where we first introduced the generator of rotations about the z axis, and relabel $\hat{J}_z$ to $\hat{S}_z$, because we were strictly concerned with rotating intrinsic spin states. In addition to renaming the operators for intrinsic spin, it is also common to relabel the basis states as $|s, m\rangle$, where

$$\hat{\mathbf{S}}^2|s, m\rangle = s(s+1)\hbar^2|s, m\rangle \tag{3.76a}$$

$$\hat{S}_z|s, m\rangle = m\hbar|s, m\rangle \tag{3.76b}$$

For a spin-$\frac{1}{2}$ particle, $s = \frac{1}{2}$ and there are two spin states, $|\frac{1}{2}, \frac{1}{2}\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle$.

Before solving the eigenvalue problem for a spin-$\frac{1}{2}$ particle, it is useful to determine the matrix representations of the spin operators $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$. We will use as a basis the states $|\frac{1}{2}, \frac{1}{2}\rangle = |+\mathbf{z}\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle = |-\mathbf{z}\rangle$ that we found in Section 3.3. In fact, we already determined the matrix representation of $\hat{S}_z$ in this basis in Section 2.5. Of course, we were calling the operator $\hat{J}_z$ then. In agreement with (2.70) we have

$$\hat{S}_z \to \begin{pmatrix} \langle +\mathbf{z}|\hat{S}_z|+\mathbf{z}\rangle & \langle +\mathbf{z}|\hat{S}_z|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|\hat{S}_z|+\mathbf{z}\rangle & \langle -\mathbf{z}|\hat{S}_z|-\mathbf{z}\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.77}$$

in the S_z basis.

In order to determine the matrix representations for $\hat{S}_x$ and $\hat{S}_y$, we start with the matrix representations of the raising and lowering operators $\hat{S}_+$ and $\hat{S}_-$, whose

action on the basis states we already know. Forming the matrix representation in the S_z basis for the raising operator using (3.61), we have

$$\hat{S}_+ \rightarrow \begin{pmatrix} \langle +\mathbf{z}|\hat{S}_+|+\mathbf{z}\rangle & \langle +\mathbf{z}|\hat{S}_+|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|\hat{S}_+|+\mathbf{z}\rangle & \langle -\mathbf{z}|\hat{S}_+|-\mathbf{z}\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{3.78}$$

reflecting the fact that

$$\hat{S}_+|+\mathbf{z}\rangle = \hat{S}_+|\tfrac{1}{2}, \tfrac{1}{2}\rangle = 0 \tag{3.79}$$

and

$$\begin{aligned} \hat{S}_+|-\mathbf{z}\rangle &= \hat{S}_+|\tfrac{1}{2}, -\tfrac{1}{2}\rangle \\ &= \sqrt{\tfrac{1}{2}\left(\tfrac{1}{2}+1\right) - \left(-\tfrac{1}{2}\right)\left(-\tfrac{1}{2}+1\right)}\,\hbar|\tfrac{1}{2}, \tfrac{1}{2}\rangle \\ &= \hbar|\tfrac{1}{2}, \tfrac{1}{2}\rangle = \hbar|+\mathbf{z}\rangle \end{aligned} \tag{3.80}$$

Also, the matrix representation of the lowering operator in the S_z basis can be obtained from (3.62):

$$\hat{S}_- \rightarrow \begin{pmatrix} \langle +\mathbf{z}|\hat{S}_-|+\mathbf{z}\rangle & \langle +\mathbf{z}|\hat{S}_-|-\mathbf{z}\rangle \\ \langle -\mathbf{z}|\hat{S}_-|+\mathbf{z}\rangle & \langle -\mathbf{z}|\hat{S}_-|-\mathbf{z}\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.81}$$

reflecting the fact that

$$\hat{S}_-|-\mathbf{z}\rangle = \hat{S}_-|\tfrac{1}{2}, -\tfrac{1}{2}\rangle = 0 \tag{3.82}$$

and

$$\begin{aligned} \hat{S}_-|+\mathbf{z}\rangle &= \hat{S}_-|\tfrac{1}{2}, \tfrac{1}{2}\rangle \\ &= \sqrt{\tfrac{1}{2}\left(\tfrac{1}{2}+1\right) - \left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}-1\right)}\,\hbar|\tfrac{1}{2}, -\tfrac{1}{2}\rangle \\ &= \hbar|\tfrac{1}{2}, -\tfrac{1}{2}\rangle = \hbar|-\mathbf{z}\rangle \end{aligned} \tag{3.83}$$

As a check, note that since $\hat{S}_+^\dagger = \hat{S}_-$, we could also obtain (3.81) as the transpose, complex conjugate of the matrix (3.78). Recall (2.80).

With the matrix representations for $\hat{S}_+$ and $\hat{S}_-$, determining the matrix representations of $\hat{S}_x$ and $\hat{S}_y$ is straightforward. Since

$$\hat{S}_+ = \hat{S}_x + i\hat{S}_y \tag{3.84}$$

$$\hat{S}_- = \hat{S}_x - i\hat{S}_y \tag{3.85}$$

then

$$\hat{S}_x = \frac{\hat{S}_+ + \hat{S}_-}{2} \tag{3.86}$$

and

$$\hat{S}_y = \frac{\hat{S}_+ - \hat{S}_-}{2i} \tag{3.87}$$

Using the matrix representations (3.78) and (3.81) in the S_z basis, we obtain

$$\hat{S}_x \to \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.88}$$

and

$$\hat{S}_y \to \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{3.89}$$

The three 2×2 matrices in (3.88), (3.89), and (3.77) (without the factors of $\hbar/2$) are often referred to as **Pauli spin matrices** and are denoted by σ_x, σ_y, and σ_z, respectively. These three equations can then be expressed in the vector notation

$$\hat{\mathbf{S}} \to \frac{\hbar}{2}\boldsymbol{\sigma} \tag{3.90}$$

where $\hat{\mathbf{S}} = \hat{S}_x\mathbf{i} + \hat{S}_y\mathbf{j} + \hat{S}_z\mathbf{k}$ and $\boldsymbol{\sigma} = \sigma_x\mathbf{i} + \sigma_y\mathbf{j} + \sigma_z\mathbf{k}$.

We are now ready to find the eigenstates of $\hat{S}_x$ or $\hat{S}_y$. In fact, we can use the matrix representations (3.90) to determine the eigenstates of $\hat{S}_n = \hat{\mathbf{S}} \cdot \mathbf{n}$ and thus find the states that are spin up and spin down along an arbitrary axis specified by the unit vector $\mathbf{n}$. We will restrict our attention to the case where $\mathbf{n} = \cos\phi\mathbf{i} + \sin\phi\mathbf{j}$ lies in the x-y plane, as indicated in Fig. 3.7. The choice $\phi = 0$ ($\phi = \pi/2$) will yield the eigenstates of $\hat{S}_x$ ($\hat{S}_y$) that we used extensively in Chapters 1 and 2. We will leave the more general case to the Problems (in particular, see Problem 3.2). We first express the eigenvalue equation in the form

$$\hat{S}_n|\mu\rangle = \mu\frac{\hbar}{2}|\mu\rangle \tag{3.91}$$

where, as we did earlier in our general discussion of angular momentum, we have included a factor of $\hbar$ so that μ is dimensionless. The factor of $\frac{1}{2}$ in the eigenvalue has been included to make things turn out nicely. After all, we know the eigenvalues already. Since the eigenvalues of $\hat{S}_z$ are $\pm\hbar/2$ and since our choice of the z axis is arbitrary, these must be the eigenvalues of $\hat{S}_n$ as well. Equation (3.91), however, does not presume particular eigenvalues, and we will see how solving the eigenvalue problem determines the allowed values of μ.

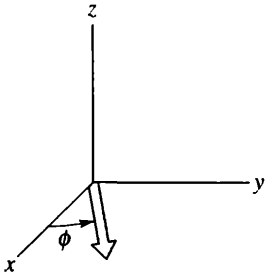


Figure 3.7 The spin-up-along-**n** state, where $\mathbf{n} = \cos\phi\mathbf{i} + \sin\phi\mathbf{j}$.

As in (2.63), we obtain two equations that can be expressed in matrix form by taking the inner product of (3.91) with the two bra vectors $\langle +\mathbf{z}|$ and $\langle -\mathbf{z}|$:

$$\frac{\hbar}{2}\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\cos\phi + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\sin\phi\right]\begin{pmatrix} \langle +\mathbf{z}|\mu\rangle \\ \langle -\mathbf{z}|\mu\rangle \end{pmatrix} = \mu\frac{\hbar}{2}\begin{pmatrix} \langle +\mathbf{z}|\mu\rangle \\ \langle -\mathbf{z}|\mu\rangle \end{pmatrix} \tag{3.92}$$

where the 2×2 matrix on the left-hand side is just the matrix representation of $\hat{S}_n = \hat{S}_x\cos\phi + \hat{S}_y\sin\phi$. Dividing out the common factor of $\hbar/2$, we can write this equation as

$$\begin{pmatrix} -\mu & e^{-i\phi} \\ e^{i\phi} & -\mu \end{pmatrix}\begin{pmatrix} \langle +\mathbf{z}|\mu\rangle \\ \langle -\mathbf{z}|\mu\rangle \end{pmatrix} = 0 \tag{3.93}$$

This is a homogeneous equation in the two unknowns $\langle +\mathbf{z}|\mu\rangle$ and $\langle -\mathbf{z}|\mu\rangle$. A non-trivial solution requires that the determinant of the coefficients vanishes. Otherwise, the 2×2 matrix in (3.93) has an inverse, and multiplying the equation by the inverse would leave just the column vector equal to zero, that is, the trivial solution. Thus

$$\begin{vmatrix} -\mu & e^{-i\phi} \\ e^{i\phi} & -\mu \end{vmatrix} = 0 \tag{3.94}$$

showing that $\mu^2 - e^{i\phi}e^{-i\phi} = \mu^2 - 1 = 0$, or $\mu = \pm 1$.

Now that we know the eigenvalues, we may determine the corresponding eigenstates. The state with $\mu = +1$ is an eigenstate of $\hat{S}_n$ with eigenvalue $\hbar/2$. Thus, in our earlier notation, it is the state $|+\mathbf{n}\rangle$, and we can relabel it accordingly: $|\mu = 1\rangle = |+\mathbf{n}\rangle$. Substituting $\mu = +1$ into (3.93), we find that

$$\langle -\mathbf{z}|+\mathbf{n}\rangle = e^{i\phi}\langle +\mathbf{z}|+\mathbf{n}\rangle \tag{3.95}$$

The requirement that the state be normalized ($\langle +\mathbf{n}|+\mathbf{n}\rangle = 1$) is satisfied provided that

$$|\langle +\mathbf{z}|+\mathbf{n}\rangle|^2 + |\langle -\mathbf{z}|+\mathbf{n}\rangle|^2 = 1 \tag{3.96}$$

Substituting (3.95) into (3.96), we find

$$2|\langle +\mathbf{z}|+\mathbf{n}\rangle|^2 = 1 \tag{3.97}$$

Thus, up to an overall phase, we may choose $\langle +\mathbf{z}|+\mathbf{n}\rangle = 1/\sqrt{2}$, which with (3.95) shows that $\langle -\mathbf{z}|+\mathbf{n}\rangle = e^{i\phi}/\sqrt{2}$, or

$$|+\mathbf{n}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\phi}}{\sqrt{2}}|-\mathbf{z}\rangle \tag{3.98}$$

Note how, up to an overall phase, this result agrees with (2.41), which we obtained by rotating the state $|+\mathbf{x}\rangle$ by an angle ϕ counterclockwise about the z axis, namely, $|+\mathbf{n}\rangle = \hat{R}(\phi\mathbf{k})|+\mathbf{x}\rangle$.

The state with $\mu = -1$ is an eigenstate of $\hat{S}_n$ with eigenvalue $-\hbar/2$. We can thus relabel this state $|\mu = -1\rangle = |-\mathbf{n}\rangle$. If we substitute the value $\mu = -1$ into (3.93), we find that

$$\langle -\mathbf{z}|-\mathbf{n}\rangle = -e^{i\phi}\langle +\mathbf{z}|-\mathbf{n}\rangle \tag{3.99}$$

Satisfying

$$|\langle +\mathbf{z}|-\mathbf{n}\rangle|^2 + |\langle -\mathbf{z}|-\mathbf{n}\rangle|^2 = 1 \tag{3.100}$$

we obtain

$$|-\mathbf{n}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle - \frac{e^{i\phi}}{\sqrt{2}}|-\mathbf{z}\rangle \tag{3.101}$$

These results are in agreement with our earlier forms for these states: setting $\phi = 0$ in (3.98) and (3.101) yields

$$|\pm\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle \pm \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle \tag{3.102}$$

while setting $\phi = \pi/2$ yields

$$|\pm\mathbf{y}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle \pm \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle \tag{3.103}$$

However, in deriving (3.102) and (3.103) here, we have not had to appeal to the results from the Stern–Gerlach experiments. We have relied on only the commutation relations of the generators of rotations and their identification with the angular momentum operators. In a similar fashion, we can work out the spin eigenstates of a particle with arbitrary intrinsic spin s. In this latter case, because there are $2s + 1$ spin states for a particle with intrinsic spin s, the corresponding eigenvalue problem will involve $(2s + 1) \times (2s + 1)$ matrices. The procedure for determining the eigenstates and corresponding eigenvalues is the same as we have used in this section, but the algebra becomes more involved as the dimensionality of the matrices increases.

EXAMPLE 3.4 Determine the matrix representation for $\hat{S}_x$ using the spin-$\frac{3}{2}$ states as a basis.

SOLUTION For $s = \frac{3}{2}$, there are four basis states, namely $|\frac{3}{2}, \frac{3}{2}\rangle$, $|\frac{3}{2}, \frac{1}{2}\rangle$, $|\frac{3}{2}, -\frac{1}{2}\rangle$, and $|\frac{3}{2}, -\frac{3}{2}\rangle$. These four states are eigenstates of $\hat{\mathbf{S}}^2$ with eigenvalue $\frac{3}{2}(\frac{3}{2}+1)\hbar^2$ as well as being eigenstates of $\hat{S}_z$ with eigenvalues $\frac{3}{2}\hbar$, $\frac{1}{2}\hbar$, $-\frac{1}{2}\hbar$, and $-\frac{3}{2}\hbar$, respectively.

Using

$$\hat{S}_+|s, m\rangle = \sqrt{s(s+1) - m(m+1)}\,\hbar|s, m+1\rangle$$

we see that

$$\hat{S}_+|\tfrac{3}{2}, \tfrac{1}{2}\rangle = \sqrt{3}\,\hbar|\tfrac{3}{2}, \tfrac{3}{2}\rangle$$

$$\hat{S}_+|\tfrac{3}{2} - \tfrac{1}{2}\rangle = 2\hbar|\tfrac{3}{2}, \tfrac{1}{2}\rangle$$

$$\hat{S}_+|\tfrac{3}{2}, -\tfrac{3}{2}\rangle = \sqrt{3}\,\hbar|\tfrac{3}{2}, -\tfrac{1}{2}\rangle$$

Thus the matrix representation for $\hat{S}_+$ is given by

$$\hat{S}_+ \to \hbar \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The matrix representation of $\hat{S}_-$ is the transpose, complex conjugate of this matrix, namely

$$\hat{S}_- \to \hbar \begin{pmatrix} 0 & 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}$$

Thus the matrix representation of $\hat{S}_x$ is given by

$$\hat{S}_x = \frac{1}{2}(\hat{S}_+ + \hat{S}_-) \to \frac{\hbar}{2} \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}$$

3.7 A Stern–Gerlach Experiment with Spin-1 Particles

Let's return to the sort of Stern–Gerlach experiments that we examined in Chapter 1, but this time let's perform one of these experiments with a beam of neutral spin-1 instead of spin-$\frac{1}{2}$ particles. Since the z component of the angular momentum of a spin-1 particle can take on the three values $\hbar$, 0, and $-\hbar$, an unpolarized beam passing through an SG**z** device splits into three different beams, with the particles deflected upward, not deflected at all, or deflected downward, depending on the value of S_z (see Fig. 3.8).

What happens if a beam of spin-1 particles passes through an SG**y** device? An unpolarized beam should split into three beams since S_y can also take on the three values $\hbar$, 0, and $-\hbar$. If we follow this SG**y** device with an SG**z** device, we can ask, for example, what fraction of the particles with $S_y = \hbar$ will be found to have $S_z = \hbar$ when they exit the SG**z** device (see Fig. 3.9)? Unlike the case of spin $\frac{1}{2}$, where it was "obvious" for two SG devices whose inhomogeneous magnetic fields were at right angles to each other that 50 percent of the particles would be spin up and 50 percent

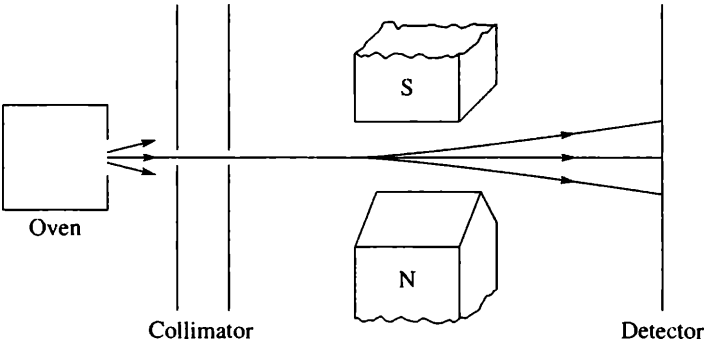


Figure 3.8 A schematic diagram indicating the paths that a spin-1 particle with S_z equal to $\hbar$, 0, or $-\hbar$ would follow in a Stern–Gerlach device.

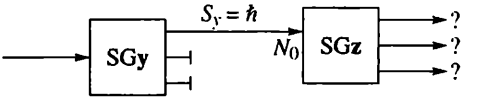


Figure 3.9 A block diagram for an experiment with spin-1 particles with two SG devices whose inhomogeneous magnetic fields are oriented at right angles to each other. What fraction of the particles exiting the SG**y** device with $S_y = \hbar$ exits the SG**z** device in each of the three channels?

would be spin down when they exited the last SG device, here the answer is not so clear. In fact, you might try guessing how the particles will be distributed before going on. To answer this question, we need to calculate the amplitude to find a particle with $S_y = \hbar$ in a state with $S_z = \hbar$, that is, to calculate the amplitude ${}_z\langle 1, 1|1, 1\rangle_y$, where we have put a subscript on the ket and bra indicating that they are eigenstates of $\hat{S}_y$ and $\hat{S}_z$, respectively. A natural way to determine the amplitude ${}_z\langle 1, 1|1, 1\rangle_y$ is to determine the eigenstates of $\hat{S}_y$ for a spin-1 particle in the S_z basis. We use the representation of $\hat{S}_y$ in the $\hat{S}_z$ basis from (3.28):

$$\hat{S}_y \xrightarrow[S_z \text{ basis}]{} \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \tag{3.104}$$

The eigenvalue equation

$$\hat{S}_y|1, \mu\rangle_y = \mu\hbar|1, \mu\rangle_y \tag{3.105}$$

becomes the matrix equation

$$\frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mu\hbar \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{3.106}$$

which can be expressed in the form

$$\begin{pmatrix} -\mu & -i/\sqrt{2} & 0 \\ i/\sqrt{2} & -\mu & -i/\sqrt{2} \\ 0 & i/\sqrt{2} & -\mu \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \tag{3.107}$$

Note that we have represented the eigenstate by the column vector

$$|1, \mu\rangle_y \to \begin{pmatrix} {}_z\langle 1, 1|1, \mu\rangle_y \\ {}_z\langle 1, 0|1, \mu\rangle_y \\ {}_z\langle 1, -1|1, \mu\rangle_y \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{3.108}$$

in the S_z basis, where we have used a, b, and c for the amplitudes for notational convenience. As we discussed in the preceding section, a nontrivial solution to (3.106) requires that the determinant of the coefficients in (3.107) must vanish:

$$\begin{vmatrix} -\mu & -i/\sqrt{2} & 0 \\ i/\sqrt{2} & -\mu & -i/\sqrt{2} \\ 0 & i/\sqrt{2} & -\mu \end{vmatrix} = 0 \tag{3.109}$$

showing that $-\mu(\mu^2 - \frac{1}{2}) + (i/\sqrt{2})(-i\mu/\sqrt{2}) = 0$, which can be written in the form $\mu(\mu^2 - 1) = 0$. Thus we see that the eigenvalues are indeed given by μ equals

1, 0, and -1, corresponding to eigenvalues $\hbar$, 0, and $-\hbar$ for $\hat{S}_y$, as expected. If we now, for example, substitute the eigenvalue $\mu = 1$ into (3.106), we obtain the equation

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{3.110}$$

indicating that for this eigenstate

$$-ib = \sqrt{2}a \qquad ia - ic = \sqrt{2}b \quad \text{and} \quad ib = \sqrt{2}c \tag{3.111}$$

From the first and last of these equations we see that $c = -a$. Since $b = i\sqrt{2}a$, the column vector in the S_z basis representing the eigenstate of $\hat{S}_y$ with eigenvalue $\hbar$ is given by

$$|1, 1\rangle_y \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} a \\ i\sqrt{2}a \\ -a \end{pmatrix} \tag{3.112}$$

The requirement that the state be normalized is

$$(a^*, -i\sqrt{2}a^*, -a^*)\begin{pmatrix} a \\ i\sqrt{2}a \\ -a \end{pmatrix} = 4|a|^2 = 1 \tag{3.113}$$

Thus, up to an overall phase, we can choose $a = \frac{1}{2}$, showing that

$$|1, 1\rangle_y \xrightarrow[S_z \text{ basis}]{} \frac{1}{2}\begin{pmatrix} 1 \\ i\sqrt{2} \\ -1 \end{pmatrix} \tag{3.114}$$

or, expressed in terms of kets,

$$|1, 1\rangle_y = \frac{1}{2}|1, 1\rangle + i\frac{\sqrt{2}}{2}|1, 0\rangle - \frac{1}{2}|1, -1\rangle \tag{3.115}$$

Note that we have not put subscripts on the kets on the right-hand side of (3.115) because, if there is no ambiguity, we will use the convention that without subscripts these are understood to be eigenkets of $\hat{S}_z$.

Based on our result, we can now ascertain how a beam of spin-1 particles exiting an SG**y** device in the state $|1, 1\rangle_y$, that is, with $S_y = \hbar$, will split when it passes through an SG**z** device. The probability of the particles exiting this SG**z** device with $S_z = \hbar$ is given by $|\langle 1, 1|1, 1\rangle_y|^2 = |\frac{1}{2}|^2 = \frac{1}{4}$; the probability of the particles

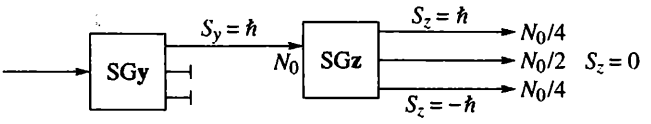


Figure 3.10 A block diagram showing the results of the Stern–Gerlach experiment with spin-1 particles.

exiting this SG**z** device with $S_z = 0$ is given by $|\langle 1, 0|1, 1\rangle_y|^2 = |i\sqrt{2}/2|^2 = \frac{1}{2}$; and the probability that the particles exit the SG**z** device with $S_z = -\hbar$ is given by $|\langle 1, -1|1, 1\rangle_y|^2 = |-\frac{1}{2}|^2 = \frac{1}{4}$. So when a beam of spin-1 "spin-up" particles from one SG device passes through another SG device whose inhomogeneous magnetic field is oriented at right angles to that of the initial device, 25 percent of the particles are deflected up, 50 percent of the particles are not deflected, and 25 percent of the particles are deflected down (see Fig. 3.10). This is to be compared with the 50 percent up and 50 percent down that we saw earlier for spin-$\frac{1}{2}$ particles in a similar experiment.

EXAMPLE 3.5 Determine the fraction of spin-1 particles exiting the SG**y** device with $S_y = 0$ that exits the SG**z** device in each of the three channels, namely with $S_z = \hbar$, $S_z = 0$, and $S_z = -\hbar$.

SOLUTION Return to (3.106) and put $\mu = 0$, which shows that $b = 0$ and $a = c$. Thus, the normalized eigenstate with $S_y = 0$ is

$$|1, 0\rangle_y \xrightarrow[S_z \text{ basis}]{} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

or, expressed in terms of kets,

$$|1, 0\rangle_y = \frac{1}{\sqrt{2}}|1, 1\rangle + \frac{1}{\sqrt{2}}|1, -1\rangle$$

Therefore $|\langle 1, 1|1, 0\rangle_y|^2 = |\langle 1, -1|1, 0\rangle_y|^2 = |1/\sqrt{2}|^2 = 1/2$. Thus 50 percent of the particles exit the SG**z** device with $S_z = \hbar$ and 50 percent exit with $S_z = -\hbar$.

The results of this chapter may convince you that it is not easy to predict the results of Stern–Gerlach experiments without a detailed calculation. If you need more evidence, try your hand at Problem 3.22 or Problem 3.25, where a beam of spin-$\frac{3}{2}$ particles is sent through a series of SG devices.

3.8 Summary

To a physicist, angular momentum along with linear momentum and energy constitute the "big three" space-time dynamical variables used to describe a system.[8] Angular momentum enters quantum mechanics in the form of three operators—$\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$—that generate rotations of states about the x, y, and z axes, respectively. Because finite rotations about different axes do not commute, the generators satisfy the commutation relations

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z \qquad [\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x \qquad [\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y \tag{3.116}$$

where the commutator of two operators $\hat{A}$ and $\hat{B}$ is defined by the relationship

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \tag{3.117}$$

Although the three generators $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$ do not commute with each other, they each commute with

$$\hat{\mathbf{J}}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \tag{3.118}$$

Thus, we can find simultaneous eigenstates of $\hat{\mathbf{J}}^2$ and one of the components, for example, $\hat{J}_z$. These eigenstates are denoted by the kets $|j, m\rangle$ where

$$\hat{\mathbf{J}}^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle \tag{3.119a}$$

$$\hat{J}_z|j, m\rangle = m\hbar|j, m\rangle \tag{3.119b}$$

Physically, we can see why $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ commute, since the eigenvalue for $\hat{\mathbf{J}}^2$ specifies the magnitude of the angular momentum for the state and the *magnitude* of the angular momentum, like the length of any vector, is not affected by a rotation.

The linear combination of the generators

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \tag{3.120}$$

is a raising operator:

$$\hat{J}_+|j, m\rangle = \sqrt{j(j+1) - m(m+1)}\,\hbar|j, m+1\rangle \tag{3.121}$$

whereas $\hat{J}_- = \hat{J}_x - i\hat{J}_y$ is a lowering operator:

$$\hat{J}_-|j, m\rangle = \sqrt{j(j+1) - m(m-1)}\,\hbar|j, m-1\rangle \tag{3.122}$$

[8] Relativistically, we could term them the big two, grouping linear momentum and energy together as an energy-momentum four-vector. The importance of these variables arises primarily because of the conservation laws that exist for angular momentum, linear momentum, and energy. In Chapter 4 we will begin to see how these conservation laws arise. Intrinsic spin angular momentum plays an unusually important role, which we will see when we consider systems of identical particles in Chapter 12.

Since the magnitude of the projection of the angular momentum on an axis for a state must be less than the magnitude of the angular momentum itself, there are limits on how far you can raise or lower the m values, which are sufficient to determine the allowed values of j and m:

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots \tag{3.123}$$

and for any particular j, m ranges from $+j$ to $-j$ in integral steps:

$$m = j, j-1, j-2, \ldots, -j+1, -j \tag{3.124}$$

The eigenstates of $\hat{J}_n = \hat{\mathbf{J}} \cdot \mathbf{n}$, the component of the angular momentum along an axis specified by the unit vector $\mathbf{n}$, can be determined by setting up the eigenvalue equation

$$\hat{J}_n |j, m\rangle_n = m\hbar |j, m\rangle_n \tag{3.125}$$

using the eigenstates of $\hat{J}_z$ as a basis. Since for a particular j, there are $2j+1$ different states $|j, m\rangle$, the eigenvalue equation (3.125) can be expressed as a matrix equation with the matrix representation of $\hat{J}_n = \hat{\mathbf{J}} \cdot \mathbf{n} = \hat{J}_x n_x + \hat{J}_y n_y + \hat{J}_z n_z$ following directly from (3.119b), (3.120), (3.121), and (3.122). As an important example, the matrix representations for spin $\frac{1}{2}$ are given by

$$\hat{\mathbf{S}} \xrightarrow[S_z \text{ basis}]{} \frac{\hbar}{2}\boldsymbol{\sigma} \tag{3.126}$$

with the Pauli spin matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.127}$$

In (3.126) we have labeled the angular momentum operators by $\hat{\mathbf{S}}$ instead of $\hat{\mathbf{J}}$, because when $j = \frac{1}{2}$ we know that we are dealing with intrinsic spin.

Finally, when two Hermitian operators do not commute,

$$[\hat{A}, \hat{B}] = iC \tag{3.128}$$

there is a fundamental uncertainty relation

$$\Delta A \Delta B \geq \frac{|\langle C \rangle|}{2} \tag{3.129}$$

From this result follows uncertainty relations for angular momentum such as

$$\Delta J_x \Delta J_y \geq \frac{\hbar}{2} |\langle J_z \rangle| \tag{3.130}$$

If the z component of the angular momentum has a definite nonzero value, making the right-hand side of (3.130) nonzero, then we cannot specify either the x or y component of the angular momentum with certainty, because this would require the

left-hand side of (3.130) to vanish, in contradiction to the inequality. This uncertainty relation is, of course, built into our results (3.123) and (3.124), which, like (3.130), follow directly from the commutation relations (3.116). Nonetheless, uncertainty relations such as (3.130) bring to the fore the sharp differences between the quantum and the classical worlds. In Chapter 6 we will see how (3.128) and (3.129) lead to the famous Heisenberg uncertainty relation $\Delta x \Delta p_x \geq \hbar/2$.

Problems

3.1. Verify for the operators $\hat{A}$, $\hat{B}$, and $\hat{C}$ that

(a) $[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}]$

(b) $[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C}$

Similarly, you can show that

(c) $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$

3.2. Using the $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ states of a spin-$\frac{1}{2}$ particle as a basis, set up and solve as a problem in matrix mechanics the eigenvalue problem for $\hat{S}_n = \hat{\mathbf{S}} \cdot \mathbf{n}$, where the spin operator $\hat{\mathbf{S}} = \hat{S}_x\mathbf{i} + \hat{S}_y\mathbf{j} + \hat{S}_z\mathbf{k}$ and $\mathbf{n} = \sin\theta\cos\phi\mathbf{i} + \sin\theta\sin\phi\mathbf{j} + \cos\theta\mathbf{k}$. Show that the eigenstates may be written as

$$|+\mathbf{n}\rangle = \cos\frac{\theta}{2}|+\mathbf{z}\rangle + e^{i\phi}\sin\frac{\theta}{2}|-\mathbf{z}\rangle$$

$$|-\mathbf{n}\rangle = \sin\frac{\theta}{2}|+\mathbf{z}\rangle - e^{i\phi}\cos\frac{\theta}{2}|-\mathbf{z}\rangle$$

Rather than simply verifying that these are eigenstates by substituting into the eigenvalue equation, obtain these states by directly solving the eigenvalue problem, as in Section 3.6.

3.3. Show that the Pauli spin matrices satisfy $\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}\,\mathbb{I}$, where i and j can take on the values 1, 2, and 3, with the understanding that $\sigma_1 = \sigma_x$, $\sigma_2 = \sigma_y$, and $\sigma_3 = \sigma_z$. Thus for $i = j$ show that $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbb{I}$, while for $i \neq j$ show that $\{\sigma_i, \sigma_j\} = 0$, where the curly brackets are called an **anticommutator,** which is defined by the relationship $\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$.

3.4. Verify that (a) $\boldsymbol{\sigma} \times \boldsymbol{\sigma} = 2i\boldsymbol{\sigma}$ and (b) $\boldsymbol{\sigma} \cdot \mathbf{a}\,\boldsymbol{\sigma} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b}\,\mathbb{I} + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b})$, where $\boldsymbol{\sigma} = \sigma_x\mathbf{i} + \sigma_y\mathbf{j} + \sigma_z\mathbf{k}$.

3.5. This problem demonstrates another way (also see Problem 3.2) to determine the eigenstates of $\hat{S}_n = \hat{\mathbf{S}} \cdot \mathbf{n}$. The operator

$$\hat{R}(\theta\mathbf{j}) = e^{-i\hat{S}_y\theta/\hbar}$$

rotates spin states by an angle θ counterclockwise about the y axis.

(a) Show that this rotation operator can be expressed in the form

$$\hat{R}(\theta \mathbf{j}) = \cos\frac{\theta}{2} - \frac{2i}{\hbar}\hat{S}_y \sin\frac{\theta}{2}$$

Suggestion: Use the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ as a basis. Express the operator $\hat{R}(\theta \mathbf{j})$ in matrix form by expanding $\hat{R}$ in a Taylor series. Examine the explicit form for the matrices representing $\hat{S}_y^2$, $\hat{S}_y^3$, and so on.

(b) Apply $\hat{R}$ in matrix form to the state $|+\mathbf{z}\rangle$ to obtain the state $|+\mathbf{n}\rangle$ given in Problem 3.2 with $\phi = 0$, that is, rotated by angle θ in the x-z plane. Show that $\hat{R}|-\mathbf{z}\rangle$ differs from $|-\mathbf{n}\rangle$ by an overall phase.

3.6. Derive (3.60).

3.7. Derive the Schwarz inequality

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2$$

Suggestion: Use the fact that

$$(\langle\alpha| + \lambda^*\langle\beta|)(|\alpha\rangle + \lambda|\beta\rangle) \geq 0$$

and determine the value of λ that minimizes the left-hand side of the equation.

3.8. Show that the operator $\hat{C}$ defined through $[\hat{A}, \hat{B}] = i\hat{C}$ is Hermitian, provided the operators $\hat{A}$ and $\hat{B}$ are Hermitian.

3.9. Calculate ΔS_x and ΔS_y for an eigenstate of $\hat{S}_z$ for a spin-$\frac{1}{2}$ particle. Check to see if the uncertainty relation $\Delta S_x \Delta S_y \geq \hbar|\langle S_z\rangle|/2$ is satisfied. Repeat your calculation for an eigenstate of $\hat{S}_x$.

3.10. Use the matrix representations of the spin-$\frac{1}{2}$ angular momentum operators $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ in the S_z basis to verify explicitly through matrix multiplication that

$$[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z$$

3.11. Determine the matrix representations of the spin-$\frac{1}{2}$ angular momentum operators $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ using the eigenstates of $\hat{S}_y$ as a basis.

3.12. Verify for a spin-$\frac{1}{2}$ particle that (a)

$$\hat{S}_z = (\hbar/2)|+\mathbf{z}\rangle\langle+\mathbf{z}| - (\hbar/2)|-\mathbf{z}\rangle\langle-\mathbf{z}|$$

and (b) the raising and lowering operators may be expressed as

$$\hat{S}_+ = \hbar|+\mathbf{z}\rangle\langle-\mathbf{z}| \quad \text{and} \quad \hat{S}_- = \hbar|-\mathbf{z}\rangle\langle+\mathbf{z}|$$

Note: It is sufficient to examine the action of these operators on the basis states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$, which of course form a complete set.

3.13. Repeat Problem 3.10 using the matrix representations (3.28) for a spin-1 particle in the J_z basis.

3.14. Use the spin-1 states $|1, 1\rangle$, $|1, 0\rangle$, and $|1, -1\rangle$ as a basis to form the matrix representations of the angular momentum operators and hence verify that the matrix representations (3.28) are correct.

3.15. Determine the eigenstates of $\hat{S}_x$ for a spin-1 particle in terms of the eigenstates $|1, 1\rangle$, $|1, 0\rangle$, and $|1, -1\rangle$ of $\hat{S}_z$.

3.16. A spin-1 particle exits an SG**z** device in a state with $S_z = \hbar$. The beam then enters an SG**x** device. What is the probability that the measurement of S_x yields the value 0?

3.17. A spin-1 particle is in the state

$$|\psi\rangle \xrightarrow[S_z \text{ basis}]{} \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3i \end{pmatrix}$$

(a) What are the probabilities that a measurement of S_z will yield the values $\hbar$, 0, or $-\hbar$ for this state? What is $\langle S_z \rangle$?

(b) What is $\langle S_x \rangle$ for this state? *Suggestion:* Use matrix mechanics to evaluate the expectation value.

(c) What is the probability that a measurement of S_x will yield the value $\hbar$ for this state?

3.18. Determine the eigenstates of $\hat{S}_n = \hat{\mathbf{S}} \cdot \mathbf{n}$ for a spin-1 particle, where the spin operator $\hat{\mathbf{S}} = \hat{S}_x\mathbf{i} + \hat{S}_y\mathbf{j} + \hat{S}_z\mathbf{k}$ and $\mathbf{n} = \sin\theta\cos\phi\,\mathbf{i} + \sin\theta\sin\phi\,\mathbf{j} + \cos\phi\,\mathbf{k}$. Use the matrix representation of the rotation operator in Problem 3.19 to check your result when $\phi = 0$.

3.19. Find the state with $S_n = \hbar$ of a spin-1 particle, where $\mathbf{n} = \sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{k}$, by rotating a state with $S_z = \hbar$ by angle θ counterclockwise about the y axis using the rotation operator $\hat{R}(\theta\mathbf{j}) = e^{-i\hat{S}_y\theta/\hbar}$. *Suggestion:* Use the matrix representation (3.104) for $\hat{S}_y$ in the S_z basis and expand the rotation operator in a Taylor series. Work out the matrices through the one representing $\hat{S}_y^3$ in order to see the pattern and show that

$$\hat{R}(\theta\mathbf{j}) \xrightarrow[S_z \text{ basis}]{} \begin{pmatrix} \frac{1+\cos\theta}{2} & -\frac{\sin\theta}{\sqrt{2}} & \frac{1-\cos\theta}{2} \\ \frac{\sin\theta}{\sqrt{2}} & \cos\theta & -\frac{\sin\theta}{\sqrt{2}} \\ \frac{1-\cos\theta}{2} & \frac{\sin\theta}{\sqrt{2}} & \frac{1+\cos\theta}{2} \end{pmatrix}$$

3.20. A beam of spin-1 particles is sent through a series of three Stern–Gerlach measuring devices (Fig. 3.11). The first SG**z** device transmits particles with $S_z = \hbar$

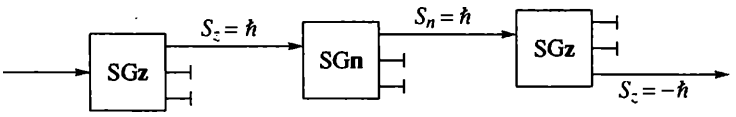


Figure 3.11 A Stern–Gerlach experiment with spin-1 particles.

and filters out particles with $S_z = 0$ and $S_z = -\hbar$. The second device, an SG**n** device, transmits particles with $S_n = \hbar$ and filters out particles with $S_n = 0$ and $S_n = -\hbar$, where the axis **n** makes an angle θ in the x-z plane with respect to the z axis. A last SG**z** device transmits particles with $S_z = -\hbar$ and filters out particles with $S_z = \hbar$ and $S_z = 0$.

(a) What fraction of the particles transmitted by the first SG**z** device will survive the third measurement? *Note:* The states with $S_n = \hbar$, $S_n = 0$, and $S_n = -\hbar$ in the S_z basis follow directly from applying the rotation operator given in Problem 3.19 to states with $S_z = \hbar$, $S_z = 0$, and $S_z = -\hbar$, respectively.

(b) How must the angle θ of the SG**n** device be oriented so as to maximize the number of particles that are transmitted by the final SG**z** device? What fraction of the particles survive the third measurement for this value of θ?

(c) What fraction of the particles survive the last measurement if the SG**n** device is removed from the experiment?

Repeat your calculation for parts (a), (b), and (c) if the last SG**z** device transmits particles with $S_z = 0$ only.

3.21. Introduce an angle θ defined by the relation $\cos\theta = J_z/|\mathbf{J}|$, reflecting the degree to which a particle's angular momentum lines up along the z axis. What is the smallest value of θ for (a) a spin-$\frac{1}{2}$ particle, (b) a spin-1 particle, and (c) a macroscopic spinning top?

3.22. Arsenic atoms in the ground state are spin-$\frac{3}{2}$ particles. A beam of arsenic atoms enters an SG**x** device, a Stern–Gerlach device with its inhomogeneous magnetic field oriented in the x direction. Atoms with $S_x = \frac{1}{2}\hbar$ then enter an SG**z** device. Determine the fraction of the atoms that exit the SG**z** device with $S_z = \frac{3}{2}\hbar$, $S_z = \frac{1}{2}\hbar$, $S_z = -\frac{1}{2}\hbar$, and $S_z = -\frac{3}{2}\hbar$.

3.23. For a spin-$\frac{3}{2}$ particle the matrix representation of the operator $\hat{S}_x$ in the S_z basis is given by

$$\hat{S}_x \rightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}$$

Pick one of the following states and verify that it is an eigenstate of $\hat{S}_x$ with the appropriate eigenvalue:

$$|\tfrac{3}{2}, \tfrac{3}{2}\rangle_x \rightarrow \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 \\ \sqrt{3} \\ \sqrt{3} \\ 1 \end{pmatrix} \qquad |\tfrac{3}{2}, \tfrac{1}{2}\rangle_x \rightarrow \frac{1}{2\sqrt{2}} \begin{pmatrix} \sqrt{3} \\ 1 \\ -1 \\ -\sqrt{3} \end{pmatrix}$$

$$|\tfrac{3}{2}, -\tfrac{1}{2}\rangle_x \rightarrow \frac{1}{2\sqrt{2}} \begin{pmatrix} \sqrt{3} \\ -1 \\ -1 \\ \sqrt{3} \end{pmatrix} \qquad |\tfrac{3}{2}, -\tfrac{3}{2}\rangle_x \rightarrow \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 \\ -\sqrt{3} \\ \sqrt{3} \\ -1 \end{pmatrix}$$

Do you notice any property of these representations that is at least consistent with the other states being correct?

3.24. A spin-$\frac{3}{2}$ particle is in the state

$$|\psi\rangle \xrightarrow[S_z \text{ basis}]{} N \begin{pmatrix} i \\ 2 \\ 3 \\ 4i \end{pmatrix}$$

(a) Determine a value for N so that $|\psi\rangle$ is appropriately normalized.

(b) What is $\langle S_x \rangle$ for this state? *Suggestion:* The matrix representation of $\hat{S}_x$ is given in Example 3.4.

(c) What is the probability that a measurement of S_x will yield the value $\hbar/2$ for this state? *Suggestion:* See Problem 3.23.

3.25.

(a) Determine the matrix representation for $\hat{S}_y$ for a spin-$\frac{3}{2}$ particle.

(b) Determine the normalized eigenstate of $\hat{S}_y$ with eigenvalue $\frac{3}{2}\hbar$.

(c) As noted in Problem 3.22, arsenic atoms in the ground state are spin-$\frac{3}{2}$ particles. A beam of arsenic atoms with $S_y = \frac{3}{2}\hbar$ enters an SG**z** device. Determine the fraction of the atoms that exit the SG**z** device with $S_z = \frac{3}{2}\hbar$, $S_z = \frac{1}{2}\hbar$, $S_z = -\frac{1}{2}\hbar$, and $S_z = -\frac{3}{2}\hbar$.

3.26. Show that if the two Hermitian operators $\hat{A}$ and $\hat{B}$ have a *complete* set of eigenstates in common, the operators commute.

3.27. Show that

$$e^{\hat{A}+\hat{B}} \neq e^{\hat{A}} e^{\hat{B}}$$

unless the operators $\hat{A}$ and $\hat{B}$ commute. Problem 7.19 shows what happens if $\hat{A}$ and $\hat{B}$ do not commute but each commutes with their commutator $[\hat{A}, \hat{B}]$.

CHAPTER 4

Time Evolution

Most of the interesting questions in physics, as in life, concern how things change with time. Just as we have introduced angular momentum operators to generate rotations, we will introduce an operator called the Hamiltonian to generate time translations of our quantum systems. After obtaining the fundamental equation of motion in quantum mechanics, the Schrödinger equation, we will examine the time evolution of a number of two-state systems, including spin precession and magnetic resonance of a spin-$\frac{1}{2}$ particle in an external magnetic field and the ammonia molecule.

4.1 The Hamiltonian and the Schrödinger Equation

We begin our discussion of time development in quantum mechanics with the time-evolution operator $\hat{U}(t)$ that translates a ket vector forward in time:

$$\hat{U}(t)|\psi(0)\rangle = |\psi(t)\rangle \tag{4.1}$$

where $|\psi(0)\rangle$ is the initial state of the system at time $t = 0$ and $|\psi(t)\rangle$ is the state of the system at time t. In order to conserve probability,[1] time evolution should not affect the normalization of the state:

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|\hat{U}^{\dagger}(t)\hat{U}(t)|\psi(0)\rangle = \langle\psi(0)|\psi(0)\rangle = 1 \tag{4.2}$$

[1] In most applications of nonrelativistic quantum mechanics, the total probability of finding the particle doesn't vary in time. However, an electron could disappear, for example, by meeting up with its antiparticle, the positron, and being annihilated. Processes such as particle creation and annihilation require relativistic quantum field theory for their description.

which requires

$$\hat{U}^{\dagger}(t)\hat{U}(t) = 1 \tag{4.3}$$

Thus the time-evolution operator must be unitary.

Just as we introduced the generator of rotations in (2.29) by considering an infinitesimal rotation, here we consider an infinitesimal time translation:

$$\hat{U}(dt) = 1 - \frac{i}{\hbar}\hat{H}\,dt \tag{4.4}$$

where the operator $\hat{H}$ is the **generator of time translations**. Clearly, we need an operator in order to change the initial ket into a different ket at a later time. This is the role played by $\hat{H}$. Unitarity of the time-evolution operator dictates that $\hat{H}$ is a Hermitian operator (see Problem 4.1).

We can now show that $\hat{U}$ satisfies a first-order differential equation in time. Since

$$\hat{U}(t+dt) = \hat{U}(dt)\hat{U}(t) = \left(1 - \frac{i}{\hbar}\hat{H}\,dt\right)\hat{U}(t) \tag{4.5}$$

then

$$\hat{U}(t+dt) - \hat{U}(t) = \left(-\frac{i}{\hbar}\hat{H}\,dt\right)\hat{U}(t) \tag{4.6}$$

indicating that the time-evolution operator satisfies[2]

$$i\hbar\frac{d}{dt}\hat{U} = \hat{H}\hat{U}(t) \tag{4.7}$$

We can also apply the operator equation (4.6) to the initial state $|\psi(0)\rangle$ to obtain

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \tag{4.8}$$

This equation, known as the **Schrödinger equation**, is the fundamental equation of motion that determines how states evolve in time in quantum mechanics. Schrödinger first proposed the equation in 1926, although not as an equation involving ket vectors but rather as a wave equation that follows from the position-space representation of (4.8), as we will see in Chapter 6.

If $\hat{H}$ is time independent, we can obtain a closed-form expression for $\hat{U}$ from a series of infinitesimal time translations:

---

[2] The derivative of an operator is defined in the usual way, that is,

$$\frac{d\hat{U}}{dt} = \lim_{\Delta t \to 0}\frac{\hat{U}(t+\Delta t) - \hat{U}(t)}{\Delta t}$$

$$\hat{U}(t) = \lim_{N\to\infty}\left[1 - \frac{i}{\hbar}\hat{H}\left(\frac{t}{N}\right)\right]^N = e^{-i\hat{H}t/\hbar} \tag{4.9}$$

where we have taken advantage of Problem 2.1. Then

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi(0)\rangle \tag{4.10}$$

Thus in order to solve the equation of motion in quantum mechanics when $\hat{H}$ is time independent, all we need is to know the initial state of the system $|\psi(0)\rangle$ *and* to be able to work out the action of the operator (4.9) on this state.

What is the physical significance of the operator $\hat{H}$? Like the generator of rotations, $\hat{H}$ is a Hermitian operator. From (4.4) we see that the dimensions of $\hat{H}$ are those of Planck's constant divided by time—namely, energy. In addition, when $\hat{H}$ itself is time independent, the expectation value of the observable to which the operator $\hat{H}$ corresponds is also independent of time:

$$\langle\psi(t)|\hat{H}|\psi(t)\rangle = \langle\psi(0)|\hat{U}^\dagger(t)\hat{H}\hat{U}(t)|\psi(0)\rangle = \langle\psi(0)|\hat{H}|\psi(0)\rangle \tag{4.11}$$

since $\hat{H}$ commutes with $\hat{U}$.[3] All of these things suggest that we identify $\hat{H}$ as the energy operator, known as the **Hamiltonian**. Therefore

$$\langle E\rangle = \langle\psi|\hat{H}|\psi\rangle \tag{4.12}$$

The eigenstates of the Hamiltonian, which are the energy eigenstates satisfying

$$\hat{H}|E\rangle = E|E\rangle \tag{4.13}$$

play a special role in quantum mechanics. The action of the time-evolution operator $\hat{U}(t)$ on these states is easy to determine using the Taylor series for the exponential:

$$\begin{aligned} e^{-i\hat{H}t/\hbar}|E\rangle &= \left[1 - \frac{i\hat{H}t}{\hbar} + \frac{1}{2!}\left(-\frac{i\hat{H}t}{\hbar}\right)^2 + \cdots\right]|E\rangle \\ &= \left[1 - \frac{iEt}{\hbar} + \frac{1}{2!}\left(-\frac{iEt}{\hbar}\right)^2 + \cdots\right]|E\rangle = e^{-iEt/\hbar}|E\rangle \end{aligned} \tag{4.14}$$

The operator $\hat{H}$ in the exponent can simply be replaced by the energy eigenvalue *when* the time-evolution operator acts on an eigenstate of the Hamiltonian. Thus if the initial state of the system is an energy eigenstate, $|\psi(0)\rangle = |E\rangle$, then

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|E\rangle = e^{-iEt/\hbar}|E\rangle \tag{4.15}$$

[3] To establish that $\hat{H}$ commutes with $\hat{U}$, use the Taylor-series expansion for $\hat{U}$, as in (4.14).

The state just picks up an overall phase as time progresses; thus, the physical state of the system does not change with time. We often call such an energy eigenstate a **stationary state** to emphasize this lack of time dependence.

You might worry that physics could turn out to be boring with a lot of emphasis on stationary states. However, if the initial state $|\psi(0)\rangle$ is a *superposition* of energy eigenstates with different energies, the *relative* phases between these energy eigenstates will change with time. Such a state is not a stationary state and the time-evolution operator will generate interesting time behavior. All we need to do to determine this time dependence is to express this initial state as a superposition of energy eigenstates, since we now know the action of the time-evolution operator on each of these states. We will see examples in Sections 4.3 and 4.5.

4.2 Time Dependence of Expectation Values

The Schrödinger equation permits us to determine in general which variables exhibit time dependence for their expectation values. If we consider an observable A, then

$$\begin{aligned}
\frac{d}{dt}\langle A\rangle &= \frac{d}{dt}\langle\psi(t)|\hat{A}|\psi(t)\rangle \\
&= \left(\frac{d}{dt}\langle\psi(t)|\right)\hat{A}|\psi(t)\rangle + \langle\psi(t)|\hat{A}\left(\frac{d}{dt}|\psi(t)\rangle\right) + \langle\psi(t)|\frac{\partial\hat{A}}{\partial t}|\psi(t)\rangle \\
&= \left(\frac{1}{-i\hbar}\langle\psi(t)|\hat{H}\right)\hat{A}|\psi(t)\rangle + \langle\psi(t)|\hat{A}\left(\frac{1}{i\hbar}\hat{H}|\psi(t)\rangle\right) + \langle\psi(t)|\frac{\partial\hat{A}}{\partial t}|\psi(t)\rangle \\
&= \frac{i}{\hbar}\langle\psi(t)|[\hat{H},\hat{A}]|\psi(t)\rangle + \langle\psi(t)|\frac{\partial\hat{A}}{\partial t}|\psi(t)\rangle \qquad (4.16)
\end{aligned}$$

The appearance of the last term involving $\partial\hat{A}/\partial t$ in this equation allows for the possibility that the operator depends *explicitly* on time. Equation (4.16) shows that provided the operator corresponding to a variable does not have any explicit time dependence ($\partial\hat{A}/\partial t = 0$), the expectation value of that variable will be a **constant of the motion** whenever the operator commutes with the Hamiltonian.

What do we mean by explicit time dependence in the operator? Our examples in Sections 4.3 and 4.4 will probably illustrate this best. The Hamiltonian for a spin-$\frac{1}{2}$ particle in a constant magnetic field is given in (4.17). There is no explicit t dependence in $\hat{H}$; therefore substituting $\hat{H}$ for the operator $\hat{A}$ in (4.16) indicates that energy is conserved, since $\hat{H}$ of course commutes with itself. However, if we examine the Hamiltonian (4.34) for a spin-$\frac{1}{2}$ particle in a time-dependent magnetic field, we see explicit time dependence within the Hamiltonian in the factor $\cos\omega t$. Such a Hamiltonian does not lead to an expectation value for the energy of the spin

system that is independent of time because $\partial \hat{H}/\partial t = 0$. There is clearly an external system that is pumping electromagnetic energy into and out of the spin system.

4.3 Precession of a Spin-$\frac{1}{2}$ Particle in a Magnetic Field

As our first example of quantum dynamics, let's consider the time evolution of the spin state of a spin-$\frac{1}{2}$ particle in a constant magnetic field. We will choose the z axis to be in the direction of the magnetic field, $\mathbf{B} = B_0\mathbf{k}$, and take the charge of the spin-$\frac{1}{2}$ particle to be $q = -e$, that is, to have the same charge as an electron. The energy operator, or Hamiltonian, is given by

$$\hat{H} = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B} = -\frac{gq}{2mc}\hat{\mathbf{S}} \cdot \mathbf{B} = \frac{ge}{2mc}\hat{S}_z B_0 = \omega_0 \hat{S}_z \tag{4.17}$$

where we have used (1.3) to relate the magnetic moment operator $\hat{\boldsymbol{\mu}}$ and the intrinsic spin operator $\hat{\mathbf{S}}$. We have also defined $\omega_0 = geB_0/2mc$. The eigenstates of $\hat{H}$ are the eigenstates of $\hat{S}_z$:

$$\hat{H}|+\mathbf{z}\rangle = \omega_0 \hat{S}_z|+\mathbf{z}\rangle = \frac{\hbar\omega_0}{2}|+\mathbf{z}\rangle = E_+|+\mathbf{z}\rangle \tag{4.18a}$$

$$\hat{H}|-\mathbf{z}\rangle = \omega_0 \hat{S}_z|-\mathbf{z}\rangle = -\frac{\hbar\omega_0}{2}|-\mathbf{z}\rangle = E_-|-\mathbf{z}\rangle \tag{4.18b}$$

where we have denoted the energy eigenvalues of the spin-up and spin-down states by E_+ and E_-, respectively.

What happens as time progresses? Since the Hamiltonian is time independent, we can take advantage of (4.9):

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar} = e^{-i\omega_0\hat{S}_z t/\hbar} = e^{-i\hat{S}_z\phi/\hbar} = \hat{R}(\phi\mathbf{k}) \tag{4.19}$$

where in the last two steps we have expressed the time-development operator as the rotation operator that rotates states about the z axis by angle $\phi = \omega_0 t$. Thus we see that placing the particle in a magnetic field in the z direction rotates the spin of the particle about the z axis as time progresses, with a period $T = 2\pi/\omega_0$. Using the terminology of classical physics, we say that the particle's spin is precessing about the z axis, as depicted in Fig. 4.1. However, we should be careful not to carry over too completely the classical picture of a magnetic moment precessing in a magnetic field since in the quantum system the angular momentum—and hence the magnetic moment—of the particle cannot actually be pointing in a specific direction because of the uncertainty relations such as (3.75).

In order to see how we work out the details of quantum dynamics, let's take a specific example. With $\mathbf{B} = B_0\mathbf{k}$, we choose $|\psi(0)\rangle = |+\mathbf{x}\rangle$. The state $|+\mathbf{x}\rangle$ is a

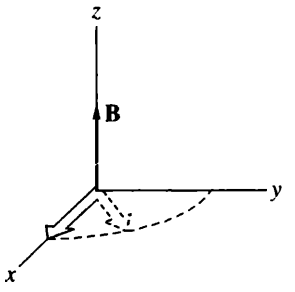


Figure 4.1 A spin-$\frac{1}{2}$ particle, initially in the state $|+\mathbf{x}\rangle$, precesses about the magnetic field, which points in the z direction.

superposition of eigenstates of $\hat{S}_z$, and therefore from (4.18) it is a superposition of energy eigenstates with different energies. The state at time t is given by

$$\begin{aligned}|\psi(t)\rangle &= e^{-i\hat{H}t/\hbar}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle\right)\\ &= \frac{e^{-iE_+t/\hbar}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{-iE_-t/\hbar}}{\sqrt{2}}|-\mathbf{z}\rangle\\ &= \frac{e^{-i\omega_0 t/2}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\omega_0 t/2}}{\sqrt{2}}|-\mathbf{z}\rangle \end{aligned} \tag{4.20}$$

This state does not simply pick up an overall phase as time progresses; it is not a stationary state. Equation (4.20) can also be written as

$$|\psi(t)\rangle = e^{-i\omega_0 t/2}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\omega_0 t}}{\sqrt{2}}|-\mathbf{z}\rangle\right) \tag{4.21}$$

which is just an overall phase factor times the spin-up state $|+\mathbf{n}\rangle$ that we found in (3.98), provided we choose the azimuthal angle $\phi = \omega_0 t$.

Let's investigate how the probabilities of being in various spin states and the spin expectation values evolve in time. We use the expression (4.20) for $|\psi(t)\rangle$. Note that

$$|\langle+\mathbf{z}|\psi(t)\rangle|^2 = \left|\frac{e^{-i\omega_0 t/2}}{\sqrt{2}}\right|^2 = \frac{1}{2} \tag{4.22a}$$

$$|\langle-\mathbf{z}|\psi(t)\rangle|^2 = \left|\frac{e^{i\omega_0 t/2}}{\sqrt{2}}\right|^2 = \frac{1}{2} \tag{4.22b}$$

are independent of time, and therefore

$$\langle S_z\rangle = \frac{1}{2}\left(\frac{\hbar}{2}\right) + \frac{1}{2}\left(-\frac{\hbar}{2}\right) = 0 \tag{4.23}$$

is also a constant of the motion.

When we examine the components of the intrinsic spin in the x-y plane, we do see explicit time dependence. Since

$$\langle +\mathbf{x}|\psi(t)\rangle = \left(\frac{1}{\sqrt{2}}\langle +\mathbf{z}| + \frac{1}{\sqrt{2}}\langle -\mathbf{z}|\right)\left(\frac{e^{-i\omega_0 t/2}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\omega_0 t/2}}{\sqrt{2}}|-\mathbf{z}\rangle\right)$$
$$= \frac{1}{\sqrt{2}}(1,\ 1)\frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\omega_0 t/2} \\ e^{i\omega_0 t/2}\end{pmatrix} = \cos\frac{\omega_0 t}{2} \tag{4.24}$$

where in the second line we have used the matrix representations for the states in the S_z basis, then

$$|\langle +\mathbf{x}|\psi(t)\rangle|^2 = \cos^2\frac{\omega_0 t}{2} \tag{4.25}$$

As a check, note that the probability of the particle being spin up along the x axis is one at time $t = 0$, as required by the initial condition. Similarly,

$$\langle -\mathbf{x}|\psi(t)\rangle = \left(\frac{1}{\sqrt{2}}\langle +\mathbf{z}| - \frac{1}{\sqrt{2}}\langle -\mathbf{z}|\right)\left(\frac{e^{-i\omega_0 t/2}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\omega_0 t/2}}{\sqrt{2}}|-\mathbf{z}\rangle\right)$$
$$= \frac{1}{\sqrt{2}}(1,\ -1)\frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\omega_0 t/2} \\ e^{i\omega_0 t/2}\end{pmatrix} = -i\,\sin\frac{\omega_0 t}{2} \tag{4.26}$$

and

$$|\langle -\mathbf{x}|\psi(t)\rangle|^2 = \sin^2\frac{\omega_0 t}{2} \tag{4.27}$$

The sum of the probabilities to be spin up or spin down along x is one for all times, since these two states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ form a complete set and probability is conserved. We can determine the average value of S_x either as the sum of the eigenvalues multiplied by the probabilities of obtaining each of these eigenvalues,

$$\langle S_x\rangle = \cos^2\frac{\omega_0 t}{2}\left(\frac{\hbar}{2}\right) + \sin^2\frac{\omega_0 t}{2}\left(-\frac{\hbar}{2}\right) = \frac{\hbar}{2}\cos\omega_0 t \tag{4.28a}$$

or from

$$\langle S_x\rangle = \langle\psi(t)|\hat{S}_x|\psi(t)\rangle$$
$$= \frac{1}{\sqrt{2}}\left(e^{i\omega_0 t/2},\ e^{-i\omega_0 t/2}\right)\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\omega_0 t/2} \\ e^{i\omega_0 t/2}\end{pmatrix}$$
$$= \frac{\hbar}{2}\cos\omega_0 t \tag{4.28b}$$

where we have used the representation for the bra and the ket vectors and the operator in the S_z basis.

A similar calculation yields

$$|\langle +\mathbf{y}|\psi(t)\rangle|^2 = \frac{1+\sin\omega_0 t}{2} \tag{4.29a}$$

$$|\langle -\mathbf{y}|\psi(t)\rangle|^2 = \frac{1-\sin\omega_0 t}{2} \tag{4.29b}$$

and

$$\langle S_y\rangle = \frac{\hbar}{2}\sin\omega_0 t \tag{4.30}$$

All of these results are consistent with the spin precessing counterclockwise around the z axis with a period $T = 2\pi/\omega_0$, in agreement with our analysis using the explicit form (4.19) of the time-evolution operator as a rotation operator. If the charge q of the particle is taken to be positive rather than negative, ω_0 is negative, and the spin precesses in a clockwise direction.

Before going on to examine some examples of spin precession, it is worthwhile commenting on the time dependence of the expectation values (4.23), (4.28), and (4.30). First, note from (4.16) that

$$\frac{d}{dt}\langle S_z\rangle = \frac{i}{\hbar}\langle\psi|[\hat{H}, \hat{S}_z]|\psi\rangle \tag{4.31}$$

We can see from the explicit form of the Hamiltonian (4.17), which is just a constant multiple of $\hat{S}_z$, that $\hat{H}$ commutes with $\hat{S}_z$ and therefore $\langle S_z\rangle$ is time independent [as (4.23) shows]. It is interesting to consider this result from the perspective of rotational invariance. In particular, with the *external* magnetic field in the z direction, rotations about the z axis leave the spin Hamiltonian unchanged. Thus the generator $\hat{S}_z$ of these rotations must commute with $\hat{H}$, and consequently from (4.31) $\langle S_z\rangle$ is a constant of the motion. The advantage of thinking in terms of symmetry (a **symmetry operation** is one that leaves the system invariant) is that we can use symmetry to determine the constants of the motion *before* we actually carry out the calculations. We can also know in advance that $\langle S_x\rangle$ and $\langle S_y\rangle$ should vary with time. After all, since $\hat{S}_x$ and $\hat{S}_y$ generate rotations about the x and y axes, respectively, and the Hamiltonian is not invariant under rotations about these axes, $\hat{H}$ does not commute with these generators.

EXAMPLE 4.1 Verify that the expectation values (4.28) and (4.30) satisfy

$$\frac{d}{dt}\langle S_x\rangle = \frac{i}{\hbar}\langle\psi|[\hat{H}, \hat{S}_x]|\psi\rangle$$

SOLUTION Since

$$\hat{H} = \omega_0\hat{S}_z$$

we want to see if

$$\frac{d}{dt}\langle S_x\rangle = \frac{i}{\hbar}\langle\psi|[\omega_0\hat{S}_z, \hat{S}_x]|\psi\rangle = -\omega_0\langle S_y\rangle$$

where we have used

$$[\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y$$

one of the fundamental commutation relations of the angular momentum operators. Substituting in the expectation values (4.28) and (4.30), we see that indeed

$$\begin{aligned}\frac{d}{dt}\langle S_x\rangle &= \frac{d}{dt}\left(\frac{\hbar}{2}\cos\omega_0 t\right)\\ &= -\frac{\hbar\omega_0}{2}\sin\omega_0 t\\ &= -\omega_0\langle S_y\rangle\end{aligned}$$

THE g FACTOR OF THE MUON

An interesting application of spin precession is the determination of the g factor of the muon. The pion is a spin-0 particle that decays into a muon and a neutrino. The primary decay mode, for example, of the positively charged pion is $\pi^+ \rightarrow \mu^+ + \nu_\mu$, where the subscript on the neutrino indicates that it is a type of neutrino associated with the muon. Unlike photons, which are both right- and left-circularly polarized, neutrinos are essentially left handed.[4] For a spin-$\frac{1}{2}$ particle like the neutrino this means that the projection of the angular momentum along the direction of motion of the neutrino is only $-\hbar/2$. There is no $+\hbar/2$ projection. Conservation of angular momentum in the decay of a pion at rest requires that the muon produced in this decay, which is also a spin-$\frac{1}{2}$ particle, be left handed as well (see Fig. 4.2). The muon is unstable and decays via $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$, with a lifetime of approximately 2.2 microseconds in the muon's rest frame. As a consequence of the weak interactions responsible for the decay, the positron is preferentially emitted in a direction opposite to the spin direction of the muon, and therefore monitoring the decay of the muon gives us information about its spin orientation. If the muon is brought to rest, say in graphite, and placed in a magnetic field of magnitude B_0 along the z direction with the initial spin state spin up along the x axis as in our earlier discussion, the spin of the muon will precess. A detector located along the x axis to detect the positrons that are produced in the decay should yield a counting rate proportional

[4] The existence of neutrino oscillations indicates that neutrinos have a very small mass. If the neutrino mass were exactly zero, neutrinos would be purely left handed.

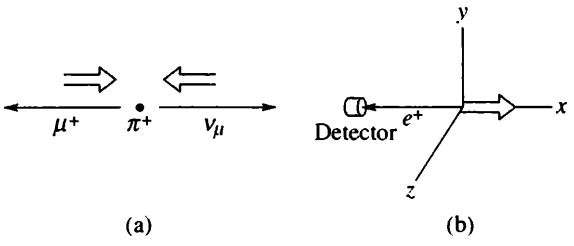


Figure 4.2 (a) Conservation of linear and angular momentum requires that the decay of the spin-0 pion in its rest frame produces a left-handed μ^+, since the ν_μ is essentially a left-handed particle. (b) The μ^+ is brought to rest with its spin up along the x axis and allowed to precess in a magnetic field in the z direction. The positrons from the μ^+ decay are emitted preferentially in the opposite direction to the spin of the μ^+.

to (4.25) as the muon's spin precesses in the magnetic field. Figure 4.3 shows the data from a typical experiment that we can use to obtain a value for the g factor (see Problem 4.7). The first measurements of this sort were carried out by Garwin et al.,[5] who found $g = 2.00 \pm 0.10$. The best experimental value for $g - 2$ of the muon, good to six significant figures, comes from a spin-precession experiment carried out at Brookhaven National Laboratory.[6] There is much interest in measuring the g factor of the muon because its accurate determination can provide information about the strong and electro-weak interactions at short distances, as well as a detailed test of quantum electrodynamics.

2π ROTATIONS OF A SPIN-$\frac{1}{2}$ PARTICLE

As a second illustration of spin precession, let's consider a beautiful experiment that demonstrates that rotating a spin-$\frac{1}{2}$ particle through 2π radians causes the state of the particle to change sign, as shown in (2.43). At first thought, it might not seem feasible to test this prediction since the state of the particle picks up an overall phase as the result of such a rotation. However, as we saw in our discussion of Experiment 4 in Chapter 1, a single particle can have amplitudes to take two separate paths and how these amplitudes add up, or interfere, depends on their *relative* phases. Werner et al.[7] used neutrons as the spin-$\frac{1}{2}$ particles and constructed an interferometer of the

[5] R. L. Garwin, L. M. Lederman, and M. Weinrich, *Phys. Rev.* **105**, 1415 (1957).

[6] This measurement [G. W. Bennett et al., *Phys. Rev. Lett.* **92**, 1618102 (2004)] takes advantage of the fact that the difference between the frequency at which the muon circles in a constant magnetic field (its cyclotron frequency) and the frequency of spin precession for a muon initially polarized parallel or antiparallel to its direction of motion is proportional to $g - 2$.

[7] S. A. Werner, R. Colella, A. W. Overhauser, and C. F. Eagen, *Phys. Rev. Lett.* **35**, 1053 (1975).

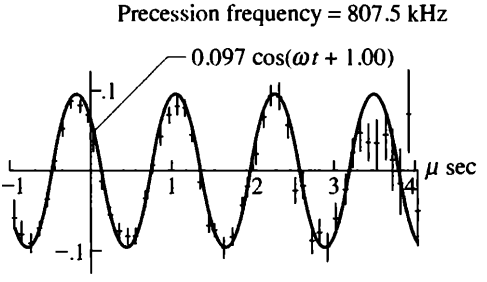


Figure 4.3 Data on the precession of a muon in a magnetic field of magnitude 60 gauss. Adapted from J. Sandweiss et al., *Phys. Rev. Lett.* **30**, 1002 (1973).

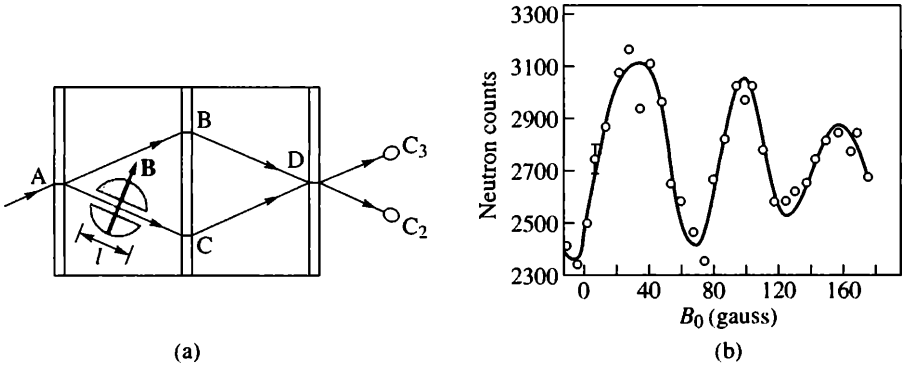


Figure 4.4 (a) A schematic diagram of the neutron interferometer and (b) the difference in counts between the counters C_3 and C_2 as a function of the magnetic field strength. Adapted from Werner et al., *Phys. Rev. Lett.* **35**, 1053 (1975).

type first developed for X-rays. Their schematic of the interferometer is shown in Fig. 4.4a. A monoenergetic beam of thermal neutrons is split by Bragg reflection from a crystal of silicon into two beams at A, one of which traverses path ABD and the other path ACD. A silicon crystal is used to deflect the beams at B and C, as well as to recombine them at D. As in a typical interferometer, there will be constructive or destructive interference depending on the path difference between the two legs ABD and ACD. The relative phase of the two beams can be altered, however, by allowing one of the beams to pass through a uniform magnetic field. As indicated by (4.21), there will be an additional phase difference of

$$\phi = \omega_0 T = \frac{geB_0}{2Mc} T \tag{4.32}$$

introduced, where M is the mass of the nucleon, B_0 is the magnitude of the uniform field on the path AC, and T is the amount of time the beam spends in the magnetic

field.[8] In the experiment, the magnitude of the magnetic field strength could be varied between 0 and 500 gauss. The difference in B_0, which we call ΔB, needed to produce successive maxima is determined by the requirement that

$$\frac{ge\Delta B}{2Mc}T = 4\pi \tag{4.33}$$

Notice that we have used the fact that a rotation by 4π radians is required to return the overall phase of spin-$\frac{1}{2}$ ket to its original value. As shown in Fig. 4.4b, Werner et al. found $\Delta B = 62 \pm 2$ gauss in their experiment. If rotating a ket by 2π radians were sufficient to keep the phase of the ket the same, the observed value of ΔB would have been one half as large as that found in the experiment. Thus the experimental results give an unambiguous confirmation of the unusual prediction (2.43) of quantum mechanics for spin-$\frac{1}{2}$ particles.

EXAMPLE 4.2 The Hamiltonian for a spin-$\frac{1}{2}$ particle in a magnetic field $\mathbf{B} = B_0\mathbf{i}$ is given by

$$\hat{H} = \omega_0 \hat{S}_x$$

where $\omega_0 = geB_0/2mc$. If initially the particle is in the state

$$|\psi(0)\rangle = |+\mathbf{z}\rangle$$

determine $|\psi(t)\rangle$, the state of the particle at time t.

SOLUTION The time development operator is given by

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar} = e^{-i\omega_0\hat{S}_x t/\hbar}$$

Since

$$e^{-i\omega_0\hat{S}_x t/\hbar} = e^{-i\hat{S}_x\phi/\hbar} = \hat{R}(\phi\mathbf{i})$$

where $\phi = \omega_0 t$, the Hamiltonian causes the spin to rotate, or precess, about the x axis in this case. In order to work out the action of the time development

[8] Three comments about this expression are in order. (1) Since a neutron is a neutral particle, it might seem strange for it to have a magnetic moment at all. That $g/2 = -1.91$ is an indication that the neutron is not itself a fundamental particle, but rather is composed of more fundamental charged constituents called quarks. (2) In nuclear physics, magnetic moments are generally expressed in terms of the nuclear magneton where the mass M in (4.32) is really the mass of the proton. Since the mass of the proton differs from the mass of the neutron by less than 0.2 percent, we can ignore this distinction unless we are interested in results to this accuracy. (3) The time T can be expressed as $T = lM/p$, where p is the momentum of the neutron and l is the path length in the magnetic field region. We can then use the de Broglie relation $p = h/\lambda$ [see (6.56)] to express this time in terms of the wavelength of the neutron. It is actually λ that is determined when selecting the energy of the neutron beam using the techniques of crystal diffraction.

operator on the state $|\psi(0)\rangle$, we need to express $|\psi(0)\rangle$ in terms of the eigenstates $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ of the Hamiltonian. Note that

$$\begin{aligned}|\psi(0)\rangle &= |+\mathbf{z}\rangle \\ &= |+\mathbf{x}\rangle\langle+\mathbf{x}|+\mathbf{z}\rangle + |-\mathbf{x}\rangle\langle-\mathbf{x}|+\mathbf{z}\rangle \\ &= \frac{1}{\sqrt{2}}|+\mathbf{x}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{x}\rangle\end{aligned}$$

Thus

$$\begin{aligned}|\psi(t)\rangle &= e^{-i\hat{H}t/\hbar}\left(\frac{1}{\sqrt{2}}|+\mathbf{x}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{x}\rangle\right) \\ &= \frac{e^{-i\omega_0 t/2}}{\sqrt{2}}|+\mathbf{x}\rangle + \frac{e^{i\omega_0 t}}{\sqrt{2}}|-\mathbf{x}\rangle\end{aligned}$$

Expressing the states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$ in terms of the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$, we see that

$$\begin{aligned}|\psi(t)\rangle &= \frac{e^{-i\omega_0 t/2}}{2}\left(|+\mathbf{z}\rangle + |-\mathbf{z}\rangle\right) + \frac{e^{i\omega_0 t/2}}{2}\left(|+\mathbf{z}\rangle - |-\mathbf{z}\rangle\right) \\ &= \frac{1}{2}\left(e^{i\omega_0 t/2} + e^{-i\omega_0 t/2}\right)|+\mathbf{z}\rangle - \frac{1}{2}\left(e^{i\omega_0 t/2} - e^{-i\omega_0 t/2}\right)|-\mathbf{z}\rangle \\ &= \cos\frac{\omega_0 t}{2}|+\mathbf{z}\rangle - i\sin\frac{\omega_0 t}{2}|-\mathbf{z}\rangle\end{aligned}$$

We can use this result to calculate, for example, $\langle S_z\rangle$:

$$\begin{aligned}\langle S_z\rangle &= \cos^2\frac{\omega_0 t}{2}\left(\frac{\hbar}{2}\right) + \sin^2\frac{\omega_0 t}{2}\left(-\frac{\hbar}{2}\right) \\ &= \left(\frac{\hbar}{2}\right)\left(\cos^2\frac{\omega_0 t}{2} - \sin^2\frac{\omega_0 t}{2}\right) \\ &= \frac{\hbar}{2}\cos\omega_0 t\end{aligned}$$

This is the same result that we obtained for $\langle S_x\rangle$ in (4.28). After all, although in this example the magnetic field pointed in the x direction and the particle's state was initially spin up along the z axis, you could have chosen to label these axes the z and x axes, respectively, making this example problem exactly the same as the example worked out at the beginning of this section. The main reason for including this example problem here is to emphasize the strategy for working out time dependence when the initial state is *not* an eigenstate of the Hamiltonian, namely, write the initial state as a superposition of the eigenstates of the Hamiltonian and then apply the time development operator to this superposition.

4.4 Magnetic Resonance

When a spin-$\frac{1}{2}$ particle precesses in a magnetic field in the z direction, the probability of the particle being spin up or spin down along z doesn't vary with time, as shown in (4.22). After all, the states $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ are stationary states of the Hamiltonian (4.17). However, if we alter the Hamiltonian by applying in addition an oscillating magnetic field transverse to the z axis, we can induce transitions between these two states by properly adjusting the frequency of this transverse field. The energy difference $E_+ - E_- = \hbar\omega_0$ can then be measured with high accuracy. This magnetic resonance gives us an excellent way of determining ω_0. Initially, physicists used magnetic resonance techniques to make accurate determinations of g factors and thus gain fundamental information about the nature of these particles. On the other hand, with known values for g, one can use the technique to make accurate determinations of the magnetic field B_0 in which the spin is precessing. For electrons or nuclei in atoms or molecules, this magnetic field is a combination of the known externally applied field and the local magnetic field at the site of the electron or nucleus. This local field provides valuable information about the nature of the bonds that electrons in the atom make with neighboring atoms in a molecule. More recently, magnetic resonance imaging (MRI) has become an important diagnostic tool in medicine.

The spin Hamiltonian for magnetic resonance is given by

$$\hat{H} = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B} = -\frac{gq}{2mc}\hat{\mathbf{S}} \cdot \mathbf{B} = -\frac{gq}{2mc}\hat{\mathbf{S}} \cdot (B_1 \cos\omega t\, \mathbf{i} + B_0\mathbf{k}) \tag{4.34}$$

where the magnetic field includes a constant magnetic field in the z direction and an oscillating magnetic field in the x direction. As we did for spin precession, we choose $q = -e$ and set $egB_0/2mc = \omega_0$. We also define $egB_1/2mc = \omega_1$. The Hamiltonian can now be written as

$$\hat{H} = \omega_0\hat{S}_z + \omega_1(\cos\omega t)\hat{S}_x \tag{4.35}$$

This Hamiltonian is time dependent, so we cannot use the expression (4.9) for the time-evolution operator.[9]

To determine how spin states evolve in time, we return to the Schrödinger equation (4.8). Let's take the state of the particle at time $t = 0$ to be $|+\mathbf{z}\rangle$. We will work in the S_z basis and express $|\psi(t)\rangle$ in this basis by

[9] If we were to choose our total system to be sufficiently large, including, in this example, the energy of both the spin system *and* the electromagnetic field, we would find that the total energy is conserved. Here we are treating the magnetic field as an external field acting on the quantum spin system.

$$|\psi(t)\rangle \rightarrow \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \tag{4.36}$$

with the initial condition

$$|\psi(0)\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{4.37}$$

In this basis, the time-development equation $\hat{H}|\psi(t)\rangle = i\hbar d|\psi(t)\rangle/dt$ is given by

$$\frac{\hbar}{2}\begin{pmatrix} \omega_0 & \omega_1 \cos \omega t \\ \omega_1 \cos \omega t & -\omega_0 \end{pmatrix}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = i\hbar \begin{pmatrix} \dot{a}(t) \\ \dot{b}(t) \end{pmatrix} \tag{4.38}$$

where $\dot{a}(t) = da/dt$ and $\dot{b}(t) = db/dt$. This coupled set of first-order differential equations cannot be solved exactly. In practice, however, the transverse field B_1 is significantly weaker than the field B_0 in the z direction and therefore the frequency ω_1 is considerably smaller than ω_0. We can take advantage of this fact to obtain an approximate solution to (4.38).

First, note that if $\omega_1 = 0$, the solution to (4.38) is

$$a(t) = a(0)e^{-i\omega_0 t/2} \quad \text{and} \quad b(t) = b(0)e^{i\omega_0 t/2} \tag{4.39}$$

in agreement with the time dependence of our earlier results (4.20). This suggests that we try writing

$$\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} c(t)e^{-i\omega_0 t/2} \\ d(t)e^{i\omega_0 t/2} \end{pmatrix} \tag{4.40}$$

where we expect that we have included the major part of the time dependence in the exponentials. If we substitute (4.40) into (4.38), we obtain

$$\begin{aligned} i\begin{pmatrix} \dot{c}(t) \\ \dot{d}(t) \end{pmatrix} &= \frac{\omega_1}{2}\cos \omega t \begin{pmatrix} d(t)e^{i\omega_0 t} \\ c(t)e^{-i\omega_0 t} \end{pmatrix} \\ &= \frac{\omega_1}{4}\begin{pmatrix} \left(e^{i(\omega_0+\omega)t} + e^{i(\omega_0-\omega)t}\right) d(t) \\ \left(e^{i(\omega-\omega_0)t} + e^{-i(\omega_0+\omega)t}\right) c(t) \end{pmatrix} \end{aligned} \tag{4.41}$$

Unless ω is chosen to be very near to ω_0, both the exponentials in the second line of (4.41) are rapidly oscillating functions that when multiplied by a more slowly oscillating function such as $c(t)$ or $d(t)$, whose time scale is set by ω_1, will cause the right-hand side of (4.41) to average to zero.[10] However, if ω is near ω_0, the terms oscillating at $\omega_0 + \omega$ can be neglected with respect to those oscillating at $\omega_0 - \omega$, and these latter terms are now oscillating sufficiently slowly that c and d vary with

[10] In a typical electron spin resonance (ESR) experiment in a field of 10^4 gauss, $\omega_0 \sim 10^{11}$ Hz, while for nuclear magnetic resonance (NMR) with protons in a comparable field, $\omega_0 \sim 10^8$ Hz.

time. Here we will solve for this time dependence when ω is equal to ω_0, the resonant condition, and leave the more general case as a problem.

Setting $\omega = \omega_0$ and neglecting the exponentials oscillating at $2\omega_0$, we obtain

$$i\begin{pmatrix}\dot{c}(t)\\ \dot{d}(t)\end{pmatrix} = \frac{\omega_1}{4}\begin{pmatrix}d(t)\\ c(t)\end{pmatrix} \tag{4.42}$$

If we take the time derivative of these two coupled equations and then use (4.42) to eliminate the terms involving a single derivative, we obtain the uncoupled second-order differential equations

$$\begin{pmatrix}\ddot{c}(t)\\ \ddot{d}(t)\end{pmatrix} = -\left(\frac{\omega_1}{4}\right)^2\begin{pmatrix}c(t)\\ d(t)\end{pmatrix} \tag{4.43}$$

The solution to (4.43) satisfying the initial condition $c(0) = 1$ and $d(0) = 0$ [see (4.42)] is $c(t) = \cos(\omega_1 t/4)$ and $d(t) = -i\sin(\omega_1 t/4)$. Thus the probability of finding the particle in the state $|-\mathbf{z}\rangle$ at time t is given by

$$|\langle -\mathbf{z}|\psi(t)\rangle|^2 = b^*(t)b(t) = d^*(t)d(t) = \sin^2\frac{\omega_1 t}{4} \tag{4.44a}$$

for a spin-$\frac{1}{2}$ particle that initially resides in the state $|+\mathbf{z}\rangle$ at $t = 0$. Similarly, the probability of finding the particle in the state $|+\mathbf{z}\rangle$ is given by

$$|\langle -\mathbf{z}|\psi(t)\rangle|^2 = a^*(t)a(t) = c^*(t)c(t) = \cos^2\frac{\omega_1 t}{4} \tag{4.44b}$$

Of course, these two probabilities sum to one, since these two states form a complete set and probability is conserved in time. If a particle initially in the state $|+\mathbf{z}\rangle$ makes a transition to the state $|-\mathbf{z}\rangle$, the energy of the spin system is reduced by $E_+ - E_- = \hbar\omega_0$, assuming $\omega_0 > 0$. This energy is added to the electromagnetic energy of the oscillating field that is *stimulating* the transition. For t between zero and $2\pi/\omega_1$, the probability of making a transition to the lower energy state grows until $b^*(t)b(t) = 1$ and $a^*(t)a(t) = 0$. Then the particle is in the state $|-\mathbf{z}\rangle$. Next for t between $2\pi/\omega_1$ and $4\pi/\omega_1$, the probability of being in the lower energy state decreases and the probability of being in the higher energy state grows as the system absorbs energy back from the electromagnetic field. This cycle of emission and absorption continues indefinitely (see Fig. 4.5).

As noted earlier, there is a probability of inducing a transition between the two spin states even when the frequency ω is not equal to ω_0. If the system is initially in the spin-up state, the probability of being in the lower energy spin-down state at time t is given by Rabi's formula (see Problem 4.9),

$$|\langle -\mathbf{z}|\psi(t)\rangle|^2 = \frac{\omega_1^2/4}{(\omega_0-\omega)^2 + \omega_1^2/4}\sin^2\frac{\sqrt{(\omega_0-\omega)^2+\omega_1^2/4}}{2}t \tag{4.45}$$

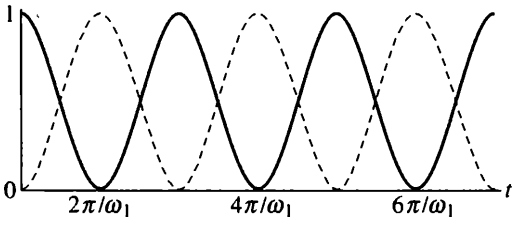


Figure 4.5 The probabilities $|\langle +\mathbf{z}|\psi(t)\rangle|^2$ (solid line) and $|\langle -\mathbf{z}|\psi(t)\rangle|^2$ (dashed line) for a spin-$\frac{1}{2}$ particle that is in the state $|+\mathbf{z}\rangle$ at $t = 0$ when the time-dependent magnetic field in the x direction is tuned to be resonant frequency.

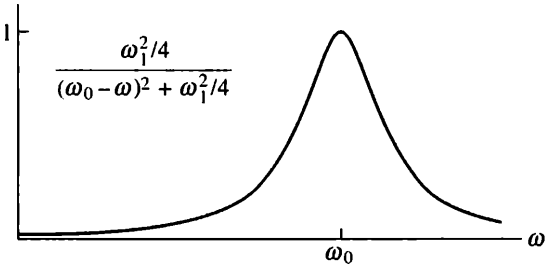


Figure 4.6 A sketch of the magnetic-resonance transition probability as a function of the frequency ω of the time-dependent magnetic field.

The maximum probability of transition is plotted as a function of ω in Fig. 4.6. Monitoring the losses and gains in energy to the oscillating field as a function of ω gives us a nice handle on whether the frequency of this field is indeed the resonant frequency of the spin system. Notice in (4.45) that making B_1 smaller makes ω_1 smaller and the curve in Fig. 4.6 narrower, permitting a more accurate determination of ω_0.

In practice, the physical spin system consists of a large number of particles, either electrons or nuclei, that are in thermal equilibrium at some temperature T. The relative number of particles in the two energy states is given by the Boltzmann distribution, so slightly more of the particles are in the lower energy state. There will be a net absorption of energy proportional to the *difference* in populations of the two levels, since the magnetic field induces transitions in both directions. Of course, if we just sit at the resonant frequency, the populations will equalize quickly and there will be no more absorption. Thus, in practice, it is necessary to move the system away from resonance, often by varying slightly the field B_0, thus permitting thermal equilibrium to be reestablished. In the case of nuclear magnetic resonance, the nuclear magnetic moments are located at the center of the atoms, surrounded by

electrons, and are relatively isolated thermally from their surroundings. Therefore, it can be difficult to get the nuclear spins to "relax" back to thermal equilibrium, even when the resonance condition no longer persists. In this case thermal contact can be increased by doping the sample with paramagnetic ions.

4.5 The Ammonia Molecule and the Ammonia Maser

As our last example in this chapter of a two-state system, we consider the ammonia molecule.[11] At first glance, the ammonia molecule does not seem a promising hunting ground for a two-state system. After all, NH_3 is a complicated system of four nuclei and ten electrons interacting with each other to form bonds between the atoms, making the stable state of the molecule a pyramid with three hydrogen atoms forming the base and a nitrogen atom at the apex (see Fig. 4.7). Here we won't worry about all of this internal dynamics, nor will we concern ourselves with how the molecule as a whole is rotating or translating. Rather, we will take the molecule to be in a fixed state as regards all of these degrees of freedom and focus on the location of the nitrogen atom; namely, is the nitrogen atom above or below the plane formed by three hydrogen atoms? The existence of a reasonably well-defined location for the nitrogen atom indicates that there is a potential well in which the nitrogen atom finds it energetically advantageous to reside. However, the geometry of the molecule tells us that if there is a potential well above the plane, there must be a similar well below the plane. Which state does the nitrogen atom choose? Nature likes to find the lowest energy state, so we are led to solve the energy eigenvalue problem to determine the allowed states and energies of the system.

We introduce two kets:

$$|1\rangle = |N \text{ above the plane}\rangle \quad \text{and} \quad |2\rangle = |N \text{ below the plane}\rangle \tag{4.46}$$

and construct the matrix representation of the Hamiltonian using these two states, depicted in Fig. 4.7, as basis states. The symmetry of the two physical configurations suggests that the expectation value of the Hamiltonian in these states, an energy that we denote by E_0, should be the same for the two states. Thus

$$\langle 1|\hat{H}|1\rangle = \langle 2|\hat{H}|2\rangle = E_0 \tag{4.47}$$

where $\hat{H}$ is the Hamiltonian of the system. What about the off-diagonal matrix elements? If we look back to our discussion of time evolution of the spin system in magnetic resonance, we see that when we set the off-diagonal matrix elements of the Hamiltonian in (4.38) equal to zero, the spin-up and spin-down states were stationary states; if the system were in one of these states initially, it remained in

[11] Our discussion of the ammonia molecule as a two-state system is inspired by the treatment in vol. 3 of *The Feynman Lectures on Physics*.

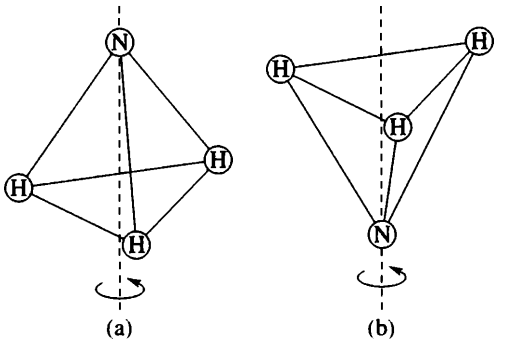


Figure 4.7 The two states of the ammonia molecule with (a) the nitrogen atom above the plane in state $|1\rangle$ and (b) the nitrogen atom below the plane in state $|2\rangle$.

that state forever, as (4.39) shows. For the ammonia molecule, the vanishing of the off-diagonal matrix elements, such as $\langle 2|\hat{H}|1\rangle$, would mean that a molecule initially, for example, in the state $|1\rangle$, with the N atom above the plane, would remain in that state. Now, if the potential barrier between the two wells were infinitely high, there would be no chance that a nitrogen atom above the plane in state $|1\rangle$ would be found below the plane in state $|2\rangle$. However, although the energy barrier formed by the three hydrogen atoms is large, it is not infinite, and there is a small amplitude for a nitrogen atom to tunnel between the two states. This means that the off-diagonal matrix element $\langle 2|\hat{H}|1\rangle$ is nonzero. We will take its value to be $-A$. Thus in the $|1\rangle$-$|2\rangle$ basis

$$\hat{H} \rightarrow \begin{pmatrix} \langle 1|\hat{H}|1\rangle & \langle 1|\hat{H}|2\rangle \\ \langle 2|\hat{H}|1\rangle & \langle 2|\hat{H}|2\rangle \end{pmatrix} = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \tag{4.48}$$

where A is a positive constant. We will see that this sign for A is required to get the correct disposition of the energy levels. Note that if, as we have presumed, the off-diagonal matrix elements are real, Hermiticity of $\hat{H}$, as well as the symmetry of the situation, requires that they be equal. In principle, if we were really adept at carrying out quantum mechanics calculations for molecules, we would be able to calculate the value of A from first principles. We can think we understand all the physics of the electromagnetic interactions responsible for holding the molecule together, but NH_3 is composed of a large number of particles and no one is able to work out all the details. We can think of (4.48) as a phenomenological Hamiltonian where the value for a constant such as A must be determined experimentally.

We are now ready to determine the energy eigenstates and eigenvalues of $\hat{H}$. The energy eigenvalue equation

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{4.49}$$

in the $|1\rangle$-$|2\rangle$ basis is given by

$$\begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \end{pmatrix} = E \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \end{pmatrix} \tag{4.50}$$

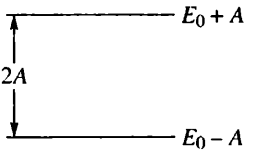


Figure 4.8 The two energy levels of the ammonia molecule.

The eigenvalues are determined by requiring

$$\begin{vmatrix} E_0 - E & -A \\ -A & E_0 - E \end{vmatrix} = 0 \tag{4.51}$$

which yields $E = E_0 \pm A$. We will denote the energy eigenstate with energy $E_I = E_0 - A$ by $|I\rangle$. Substituting the eigenvalue into (4.50) shows that $\langle 1|I\rangle = \langle 2|I\rangle$, so that we may write[12]

$$|I\rangle = \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle \tag{4.52}$$

Energy eigenstate $|II\rangle$ with energy $E_{II} = E_0 + A$ satisfies $\langle 1|II\rangle = -\langle 2|II\rangle$ and thus may be written as

$$|II\rangle = \frac{1}{\sqrt{2}}|1\rangle - \frac{1}{\sqrt{2}}|2\rangle \tag{4.53}$$

The existence of tunneling between the states $|1\rangle$ and $|2\rangle$ has split the energy states of the molecule into two states with different energies, one with energy $E_0 - A$ and the other with energy $E_0 + A$, as shown in Fig. 4.8. The wavelength of the electromagnetic radiation emitted when the molecule makes a transition between these two energy states is observed to be $1\frac{1}{4}$ cm, corresponding to an energy separation $E_{II} - E_I = h\nu = hc/\lambda$ of 10^{-4} eV. This small energy separation is to be compared with a typical spacing of atomic energy levels that is on the order of electron volts, requiring optical or uv photons to excite the atom. Molecules also have vibrational and rotational energy levels, but these modes are excited by photons in the infrared or far infrared, respectively. Exciting an ammonia molecule from state $|I\rangle$ to state $|II\rangle$ requires electromagnetic radiation of an even longer wavelength, in the microwave part of the spectrum. The smallness of this energy difference $E_{II} - E_I = 2A$ is a reflection of the smallness of the amplitude for tunneling from state $|1\rangle$ to $|2\rangle$.

Notice that neither in energy eigenstate $|I\rangle$ nor $|II\rangle$ is the nitrogen atom located above or below the plane formed by the three hydrogen atoms. Under the transformation $|1\rangle \leftrightarrow |2\rangle$ that flips the position of the nitrogen atom, the state $|I\rangle$ is symmetric, that is, $|I\rangle \rightarrow |I\rangle$, while the state $|II\rangle$ is antisymmetric, that is, $|II\rangle \rightarrow -|II\rangle$. We can,

[12] In the normalization of the state, we have neglected the nonzero amplitude $\langle 2|1\rangle$ because of its small magnitude.

however, localize the nitrogen atom above the plane, for example, by superposing the energy eigenstates:

$$|1\rangle = \frac{1}{\sqrt{2}}|I\rangle + \frac{1}{\sqrt{2}}|II\rangle \tag{4.54}$$

If $|\psi(0)\rangle = |1\rangle$, then

$$\begin{aligned} |\psi(t)\rangle &= e^{-i\hat{H}t/\hbar}\left(\frac{1}{\sqrt{2}}|I\rangle + \frac{1}{\sqrt{2}}|II\rangle\right) \\ &= \frac{e^{-i(E_0-A)t/\hbar}}{\sqrt{2}}|I\rangle + \frac{e^{-i(E_0+A)t/\hbar}}{\sqrt{2}}|II\rangle \\ &= e^{-i(E_0-A)t/\hbar}\left(\frac{1}{\sqrt{2}}|I\rangle + \frac{e^{-2iAt/\hbar}}{\sqrt{2}}|II\rangle\right) \end{aligned} \tag{4.55}$$

where in the last step we have pulled an overall phase factor out in front of the ket. Since the initial state of the molecule is a superposition of energy states with different energies, the molecule is not in a stationary state. We see that the *relative* phase between the two energy eigenstates changes with time, and thus the state of the molecule is really varying in time. The motion is periodic with a period T determined from $2AT/\hbar = 2\pi$. What is the nature of the motion? When $t = T/2$, the relative phase is π and

$$|\psi(T/2)\rangle = (\text{overall phase})\left(\frac{1}{\sqrt{2}}|I\rangle - \frac{1}{\sqrt{2}}|II\rangle\right) = (\text{overall phase})\,|2\rangle \tag{4.56}$$

The nitrogen atom is located below the plane. Thus the nitrogen atom oscillates back and forth above and below the plane with a frequency $\nu = 1/T = A/\pi\hbar = 2A/h$. This frequency, which equals 24 GHz, is the same as the frequency of the electromagnetic radiation emitted when the molecule makes a transition between states $|II\rangle$ and $|I\rangle$.

THE MOLECULE IN A STATIC EXTERNAL ELECTRIC FIELD

Since the valence electrons in the ammonia molecule tend to reside somewhat closer to the nitrogen atom, the nitrogen atom is somewhat negative and the hydrogen atoms are somewhat positive. Thus the molecule has an *electric* dipole moment $\boldsymbol{\mu}_e$ directed away from the nitrogen atom toward the plane formed by the hydrogen atoms. Just as the *magnetic* dipole moment associated with its spin angular momentum allowed us to interact with a spin-$\frac{1}{2}$ particle in Stern–Gerlach or spin-precession experiments by inserting it in a magnetic field, we can interact with the ammonia molecule by placing it in an external electric field $\mathbf{E}$, as indicated in Fig. 4.9. There is an energy of interaction with the electric field of the form $-\boldsymbol{\mu}_e \cdot \mathbf{E}$ that will differ depending on whether the nitrogen atom is above the plane in state $|1\rangle$ or below the plane in state $|2\rangle$. The presence of this electric field modifies the matrix representation of the

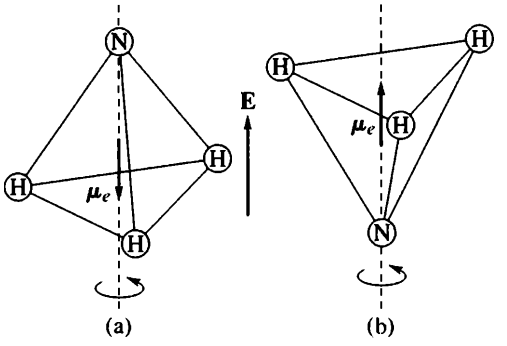


Figure 4.9 The electric dipole moment $\boldsymbol{\mu}_e$ of the ammonia molecule in (a) state $|1\rangle$ and (b) state $|2\rangle$. In the presence of an external electric field $\mathbf{E}$, the two states acquire different energies, as indicated in (4.57).

Hamiltonian in the $|1\rangle$-$|2\rangle$ basis:[13]

$$\hat{H} \to \begin{pmatrix} \langle 1|\hat{H}|1\rangle & \langle 1|\hat{H}|2\rangle \\ \langle 2|\hat{H}|1\rangle & \langle 2|\hat{H}|2\rangle \end{pmatrix} = \begin{pmatrix} E_0 + \mu_e|\mathbf{E}| & -A \\ -A & E_0 - \mu_e|\mathbf{E}| \end{pmatrix} \tag{4.57}$$

where we assume the external field is sufficiently weak that it does not affect the amplitude for the nitrogen atom to tunnel through the barrier. The eigenvalues are determined by the requirement that

$$\begin{vmatrix} E_0 + \mu_e|\mathbf{E}| - E & -A \\ -A & E_0 - \mu_e|\mathbf{E}| - E \end{vmatrix} = 0 \tag{4.58a}$$

or

$$E = E_0 \pm \sqrt{(\mu_e|\mathbf{E}|)^2 + A^2} \tag{4.58b}$$

See Fig. 4.10. Most external electric fields satisfy $\mu_e|\mathbf{E}| \ll A$, so we can expand the square root in a Taylor series or a binomial expansion to obtain

$$E \cong E_0 \pm A \pm \frac{1}{2}\frac{(\mu_e|\mathbf{E}|)^2}{A} \tag{4.59}$$

As in the Stern–Gerlach experiments where we used an inhomogeneous magnetic field to make measurements of the intrinsic spin and select spin-up and spin-down states, here we use an inhomogeneous electric field to separate NH_3 molecules into those in states $|I\rangle$ and $|II\rangle$. If we call the direction in which the electric field increases the z direction, then the force in that direction is given by

$$F_z \cong -\frac{\partial}{\partial z}\left[\pm\frac{(\mu_e|\mathbf{E}|)^2}{2A}\right] \tag{4.60}$$

[13] It is customary to use $\boldsymbol{\mu}_e$ for the electric dipole moment to avoid confusion with the symbol for momentum. We also use $|\mathbf{E}|$ for the magnitude of the electric field to avoid confusion with the symbol for energy.

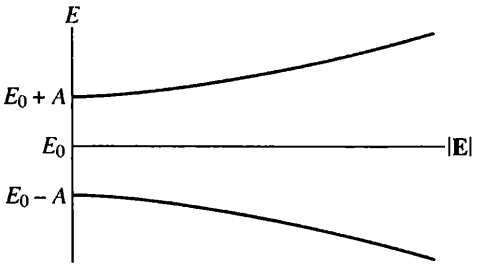


Figure 4.10 The energy levels of the ammonia molecule in an external electric field.

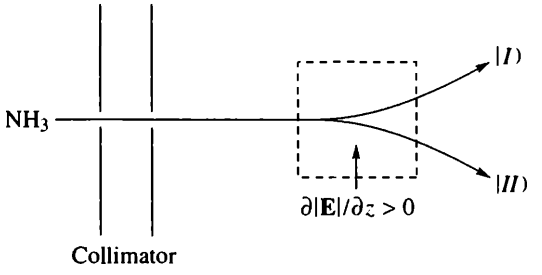


Figure 4.11 A beam of ammonia molecules passing through a region in which there is a strong electric field gradient separates into two beams, one with the molecules in state $|I\rangle$ and the other with the molecules in the state $|II\rangle$.

Notice that the minus sign in (4.59) corresponds to the state with energy $E_0 - A$ in the absence of the external electric field. Hence a molecule in state $|I\rangle$ will be deflected in the positive z direction, while a molecule in the state $|II\rangle$ will be deflected in the negative z direction, as shown in Fig. 4.11. Because of the small value of A in the denominator in (4.60), it is relatively easy to separate a beam of ammonia molecules in, for example, a gas jet by sending them through a region in which there is a large gradient in the electric field.

THE MOLECULE IN A TIME-DEPENDENT ELECTRIC FIELD

We are now ready to induce transitions between states $|I\rangle$ and $|II\rangle$ by applying a time-dependent electric field of the form $\mathbf{E} = \mathbf{E}_0 \cos \omega t$. There will be a resonant absorption or emission of electromagnetic energy, provided that $\hbar\omega$ is equal to the energy difference $2A$ between the two states. This sounds similar to the magnetic resonance effects that we treated in the previous section, and in fact the mathematics describing the two problems is essentially identical. To see this, consider the Hamiltonian in the $|1\rangle$-$|2\rangle$ basis as given in (4.57) with a time-dependent electric field. If we transform to the $|I\rangle$-$|II\rangle$ basis, we obtain (see Problem 4.10)

$$\hat{H} \rightarrow \begin{pmatrix} \langle I|\hat{H}|I\rangle & \langle I|\hat{H}|II\rangle \\ \langle II|\hat{H}|I\rangle & \langle II|\hat{H}|II\rangle \end{pmatrix} = \begin{pmatrix} E_0 - A & \mu_e|\mathbf{E}_0| \cos \omega t \\ \mu_e|\mathbf{E}_0| \cos \omega t & E_0 + A \end{pmatrix} \tag{4.61}$$

Comparing this matrix with that for the Hamiltonian in (4.38) of a spin-$\frac{1}{2}$ particle in an oscillating transverse magnetic field, we see that it is possible to draw a one-to-one correspondence between each term in the two matrices: $E_+ = \hbar\omega_0/2 \rightarrow E_0 + A$,

$E_- = -\hbar\omega_0/2 \to E_0 - A$, and $\hbar\omega_1/2 \to \mu_e|\mathbf{E}_0|$. Thus one can follow the steps leading to the probability of making a transition between the spin-up and spin-down states in (4.44) and apply them to this new problem to obtain the probability of making a transition between states $|I\rangle$ and $|II\rangle$. Therefore, at resonance the probability of finding the ammonia molecule in state $|I\rangle$ for a molecule initially in the state $|II\rangle$ at time $t = 0$ is

$$|\langle I|\psi(t)\rangle|^2 = \sin^2 \frac{\mu_e|\mathbf{E}_0|t}{2\hbar} \tag{4.62}$$

We can combine the results of this section and the preceding one to provide a description of a simple ammonia maser (Microwave Amplification by Stimulated Emission of Radiation). First we use an inhomogeneous electric field to select a beam of ammonia molecules that are in the upper energy state $|II\rangle$; then we send this beam into a microwave cavity whose resonant frequency is tuned to 24 GHz, the resonant frequency of the ammonia molecule. If the molecules spend a time T in the cavity such that $\mu_e|\mathbf{E}_0|T/2\hbar = \pi/2$, then according to (4.62) they will all make transitions from state $|II\rangle$ to state $|I\rangle$. The molecular energy released in this transition is fed into the cavity, where it can be used as microwave radiation.[14]

4.6 The Energy-Time Uncertainty Relation

As our last topic on time evolution, let's consider the energy-time uncertainty relation

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{4.63}$$

The uncertainty relation is somewhat of a misnomer; unlike our previous uncertainty relations such as (3.74), only ΔE in (4.63) is a legitimate uncertainty. It reflects the spread in energy characterizing a particular state. To see the meaning of Δt, consider an example. Let's return to the ammonia molecule that is initially in state $|1\rangle$, with the nitrogen atom above the plane. As (4.55) shows, this state is not an energy eigenstate but a superposition of two energy eigenstates with different energies. The uncertainty in the energy of a molecule in this state is given by

$$\begin{aligned}\Delta E &= \left(\langle E^2\rangle - \langle E\rangle^2\right)^{1/2} \\ &= \left\{\tfrac{1}{2}(E_0 + A)^2 + \tfrac{1}{2}(E_0 - A)^2 - \left[\tfrac{1}{2}(E_0 + A) + \tfrac{1}{2}(E_0 - A)\right]^2\right\}^{1/2} \\ &= A\end{aligned} \tag{4.64}$$

[14] The key element missing from our discussion of the maser is the coherent nature of the radiation that it produces. So far we have treated the electromagnetic field as a classical field and have not taken into account its quantum properties, that is, that it is really composed of photons. We will examine this issue in more detail in Chapter 14.

We can express the time evolution of the state (4.55) in terms of the uncertainty ΔE as

$$|\psi(t)\rangle = (\text{overall phase})\left(\frac{1}{\sqrt{2}}|I\rangle + \frac{e^{-2i\Delta E t/\hbar}}{\sqrt{2}}|II\rangle\right) \tag{4.65}$$

How long do we have to wait before the state of the molecule changes? The answer to this question is the quantity we call Δt. To be sure the state (4.65) has changed, we need to be sure the relative phase between the energy eigenstates $|I\rangle$ and $|II\rangle$ has changed significantly from its value of zero at $t = 0$ to something of order unity. This requires that the time interval Δt satisfy $2\Delta E \Delta t/\hbar \gtrsim 1$, which is in accord with (4.63).[15] In (4.55) we saw that the time required for the nitrogen atom to appear below the plane in state $|2\rangle$ is determined by the requirement that the relative phase change by π. Thus the time interval Δt determined by (4.63) is roughly one-third of the time required for the nitrogen atom to oscillate from above to below the plane.

Notice that if, instead of being in a superposition of energy eigenstates with different energies, the state of the molecule had been an energy eigenstate, there would be a definite value for the energy of the molecule, and hence $\Delta E = 0$. But in this case, the ket would pick up only an overall phase as time evolved, and the time interval Δt required for the state to change would be infinite. An energy eigenstate is really a stationary state.

Our discussion and example should make clear that Δt is not an uncertainty at all. Time in nonrelativistic quantum mechanics is just a *parameter* and not a dynamical variable like energy, angular momentum, position, or momentum with which there may be an uncertainty depending on the state of the system. When we discuss the state of the system at time t, there is no inherent limit on how accurately we can specify this time.

In the example we chose a particular initial state $|\psi\rangle$ and then examined the length of the **evolutionary time** Δt for that state to change. Now that we understand the meaning of the uncertainty relation (4.63), we can turn this around slightly. An atom (or an ammonia molecule) in an excited-energy state will not remain in this state indefinitely, even if undisturbed by any outside influence. It will decay to lower energy states with some lifetime τ. In Chapter 14 we will see how to calculate the lifetime for excited states of the hydrogen atom using the Hamiltonian arising from the interactions of charged particles with the electromagnetic field. Thus an excited state is not a stationary state, and the lifetime τ sets a natural evolutionary time for that state. Therefore, from (4.63) there must be an uncertainty in the energy of the excited state given by $\Delta E \sim \hbar/\tau$. Photons emitted in this transition will have not

[15] We have taken the lower limit in this example as an *approximate* equality since we have somewhat arbitrarily chosen to say that the system has changed when the phase in (4.65) reaches one. A more formal derivation of (4.63) and corresponding specification of Δt are given in Example 4.3.

a definite energy but rather a spread in energies. This is the origin of the natural linewidth (see Problem 4.16).

EXAMPLE 4.3 Consider any observable A associated with the state of the system in quantum mechanics. Show that there is an uncertainty relation of the form

$$\Delta E \left(\frac{\Delta A}{|d\langle A\rangle/dt|} \right) \geq \frac{\hbar}{2}$$

provided the operator $\hat{A}$ does not depend explicitly on time. The quantity $\Delta A/|d\langle A\rangle/dt|$ is a time that we may call Δt. What is the physical significance of Δt?

SOLUTION Recall that $[\hat{A}, \hat{B}] = i\hat{C}$ implies that $\Delta A \Delta B \geq |\langle C\rangle|/2$. Start with the commutator $[\hat{A}, \hat{H}]$; then

$$\Delta A \Delta E \geq \frac{1}{2} |\langle \psi | [\hat{A}, \hat{H}] | \psi \rangle|$$

But since

$$\frac{d\langle A\rangle}{dt} = \frac{i}{\hbar} \langle \psi | [\hat{H}, \hat{A}] | \psi \rangle$$

then

$$\Delta A \Delta E \geq \frac{\hbar}{2} \left| \frac{d\langle A\rangle}{dt} \right|$$

or

$$\Delta E \left(\frac{\Delta A}{|d\langle A\rangle/dt|} \right) \geq \frac{\hbar}{2}$$

If we define

$$\Delta t = \frac{\Delta A}{|d\langle A\rangle/dt|}$$

then

$$\Delta A \Delta t \geq \frac{\hbar}{2}$$

The time Δt is the time necessary for the expectation value to change by an amount on the order of the uncertainty. Thus it is the time you need to wait to be sure that the results of measuring A have really changed. For example, for position, if $\Delta x = 1\,\text{cm}$ and $d\langle x\rangle/dt = 1\,\text{mm/s}$, then $\Delta x/|d\langle x\rangle/dt| = 10\,\text{s}$, which is the time necessary for $\langle x\rangle$ to shift by an amount Δx.

4.7 Summary

Time development is where much of the action occurs in quantum mechanics. To move states forward in time, we introduce a time-evolution operator $\hat{U}(t)$ so that

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle \tag{4.66}$$

In order for probability to be conserved as time evolves,

$$\begin{aligned}\langle\psi(t)|\psi(t)\rangle &= \langle\psi(0)|\hat{U}^\dagger(t)\hat{U}(t)|\psi(0)\rangle \\ &= \langle\psi(0)|\psi(0)\rangle\end{aligned} \tag{4.67}$$

and consequently the operator $\hat{U}(t)$ must be unitary:

$$\hat{U}^\dagger(t)\hat{U}(t) = 1 \tag{4.68}$$

The Hamiltonian $\hat{H}$, the energy operator, enters as the generator of time translations through the infinitesimal time-evolution operator:

$$\hat{U}(dt) = 1 - \frac{i}{\hbar}\hat{H}dt \tag{4.69}$$

The unitarity requirement (4.68) then dictates that the Hamiltonian is Hermitian.

The time-evolution operator obeys the differential equation

$$\hat{H}\hat{U}(t) = i\hbar\frac{d}{dt}\hat{U}(t) \tag{4.70}$$

leading to the Schrödinger equation:

$$\hat{H}|\psi(t)\rangle = i\hbar\frac{d}{dt}|\psi(t)\rangle \tag{4.71}$$

A particularly useful solution to (4.70) occurs when the Hamiltonian is independent of time, in which case the time-development operator is given by

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar} \tag{4.72}$$

The action of the time-development operator (4.72) on an energy eigenstate $|E\rangle$ is given by

$$e^{-i\hat{H}t/\hbar}|E\rangle = e^{-iEt/\hbar}|E\rangle \tag{4.73}$$

showing that a single energy eigenstate just picks up an overall phase as time evolves and is therefore a stationary state. Time evolution for a state $|\psi\rangle$ that can be expressed as a superposition of energy eigenstates as

$$|\psi(0)\rangle = \sum_n |E_n\rangle\langle E_n|\psi(0)\rangle \tag{4.74}$$

is given by

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}\sum_n |E_n\rangle\langle E_n|\psi(0)\rangle$$

$$= \sum_n e^{-iE_nt/\hbar}|E_n\rangle\langle E_n|\psi(0)\rangle \tag{4.75}$$

When the superposition (4.74) involves states with different energies, the relative phase between the energy eigenstates changes with time. The time Δt (the evolutionary time) necessary for the system to change with time in this case satisfies

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{4.76}$$

where ΔE is the usual uncertainty in energy for the state $|\psi\rangle$.

Expectation values satisfy

$$\frac{d}{dt}\langle A\rangle = \frac{i}{\hbar}\langle\psi(t)|[\hat{H}, \hat{A}]|\psi(t)\rangle + \langle\psi(t)|\frac{\partial\hat{A}}{\partial t}|\psi(t)\rangle \tag{4.77}$$

which tells us that observables that do not explicitly depend on time will be constants of the motion when they commute with the Hamiltonian.

Although this chapter is devoted to time evolution, the similarity between the operators that generate rotations [see (3.10)] and the operator that generates time translations [see (4.72)] is striking. Or compare the form for an infinitesimal rotation operator $\hat{R}(d\phi\mathbf{n}) = 1 - i\hat{J}_n d\phi/\hbar$ for rotations by angle $d\phi$ about the axis specified by the unit vector $\mathbf{n}$ with the infinitesimal time translation operator (4.69). We can actually tie the rotation operator and the time-evolution operator together with a common thread—namely, symmetry. A symmetry operation is one that leaves the physical system unchanged, or invariant. For example, if the Hamiltonian is invariant under rotations about an axis, the generator of rotations about that axis must commute with the Hamiltonian. But (4.77) then tells us that the component of the angular momentum along this axis is conserved, since its expectation value doesn't vary in time. Also, if the Hamiltonian is invariant under time translations, which simply means that $\hat{H}$ is independent of time, then of course energy is conserved. We will have more to say about symmetry, especially in Chapter 9, but this is our first indication of the important connection between symmetries of a physical system and conservation laws.

Problems

4.1. Show that unitarity of the infinitesimal time-evolution operator (4.4) requires that the Hamiltonian $\hat{H}$ be Hermitian.

4.2. Show that if the Hamiltonian depends on time and $[\hat{H}(t_1), \hat{H}(t_2)] = 0$, the time-development operator is given by

$$\hat{U}(t) = \exp\left[-\frac{i}{\hbar}\int_0^t dt'\hat{H}(t')\right]$$

4.3. Use (4.16) to verify that the expectation value of an observable A does not change with time if the system is in an energy eigenstate (a stationary state) and $\hat{A}$ does not depend explicitly on time.

4.4. A beam of spin-$\frac{1}{2}$ particles with speed v_0 passes through a series of two SG**z** devices. The first SG**z** device transmits particles with $S_z = \hbar/2$ and filters out particles with $S_z = -\hbar/2$. The second SG**z** device transmits particles with $S_z = -\hbar/2$ and filters out particles with $S_z = \hbar/2$. Between the two devices is a region of length l_0 in which there is a uniform magnetic field B_0 pointing in the x direction. Determine the smallest value of l_0 such that exactly 25 percent of the particles transmitted by the first SG**z** device are transmitted by the second device. Express your result in terms of $\omega_0 = egB_0/2mc$ and v_0.

4.5. A beam of spin-$\frac{1}{2}$ particles in the $|+\mathbf{z}\rangle$ state enters a uniform magnetic field B_0 in the x-z plane oriented at an angle θ with respect to the z axis. At time T later, the particles enter an SG**y** device. What is the probability the particles will be found with $S_y = \hbar/2$? Check your result by evaluating the special cases $\theta = 0$ and $\theta = \pi/2$.

4.6. Verify that the expectation values (4.23), (4.28), and (4.30) for a spin-$\frac{1}{2}$ particle precessing in a uniform magnetic field B_0 in the z direction satisfy (4.16).

4.7. Use the data given in Fig. 4.3 to determine the g factor of the muon.

4.8. A spin-$\frac{1}{2}$ particle, initially in a state with $S_n = \hbar/2$ with $\mathbf{n} = \sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{k}$, is in a constant magnetic field B_0 in the z direction. Determine the state of the particle at time t and determine how $\langle S_x\rangle$, $\langle S_y\rangle$, and $\langle S_z\rangle$ vary with time.

4.9. Derive Rabi's formula (4.45).

4.10. Express the Hamiltonian (4.57) for the ammonia molecule in the $|I\rangle$-$|II\rangle$ basis to obtain (4.61). Assume the electric field $\mathbf{E} = \mathbf{E}_0\cos\omega t$. Compare this Hamiltonian with that for a spin-$\frac{1}{2}$ particle in a time-dependent magnetic field that appears in (4.38) and deduce the form for the probability of finding the molecule in state $|I\rangle$ at time t if it is initially placed in the state $|II\rangle$; that is, what is the analogue of Rabi's formula (4.45) for the ammonia molecule?

4.11. A spin-1 particle with a magnetic moment $\boldsymbol{\mu} = (gq/2mc)\mathbf{S}$ is situated in a magnetic field $\mathbf{B} = B_0\mathbf{k}$ in the z direction. At time $t = 0$ the particle is in a state with $S_y = \hbar$ [see (3.115)]. Determine the state of the particle at time t. Calculate how the expectation values $\langle S_x\rangle$, $\langle S_y\rangle$, and $\langle S_z\rangle$ vary in time.

4.12. A particle with intrinsic spin one is placed in a constant external magnetic field B_0 in the x direction. The initial spin state of the particle is $|\psi(0)\rangle = |1, 1\rangle$, that is,

a state with $S_z = \hbar$. Take the spin Hamiltonian to be

$$\hat{H} = \omega_0 \hat{S}_x$$

and determine the probability that the particle is in the state $|1, -1\rangle$ at time t. *Suggestion:* If you haven't already done so, you should first work out Problem 3.15 to determine the eigenstates of $\hat{S}_x$ for a spin-1 particle in terms of the eigenstates of $\hat{S}_z$.

4.13. Let

$$\begin{pmatrix} E_0 & 0 & A \\ 0 & E_1 & 0 \\ A & 0 & E_0 \end{pmatrix}$$

be the matrix representation of the Hamiltonian for a three-state system with basis states $|1\rangle$, $|2\rangle$, and $|3\rangle$.

(a) If the state of the system at time $t = 0$ is $|\psi(0)\rangle = |2\rangle$, what is $|\psi(t)\rangle$?

(b) If the state of the system at time $t = 0$ is $|\psi(0)\rangle = |3\rangle$, what is $|\psi(t)\rangle$?

4.14. The matrix representation of the Hamiltonian for a photon propagating *along* the optic axis (taken to be the z axis) of a quartz crystal using the linear polarization states $|x\rangle$ and $|y\rangle$ as basis is given by

$$\hat{H} \xrightarrow[|x\rangle\text{-}|y\rangle \text{ basis}]{} \begin{pmatrix} 0 & -iE_0 \\ iE_0 & 0 \end{pmatrix}$$

(a) What are the eigenstates and eigenvalues of the Hamiltonian?

(b) A photon enters the crystal linearly polarized in the x direction, that is $|\psi(0)\rangle = |x\rangle$. What is the $|\psi(t)\rangle$, the state of the photon at time t? Express your answer in the $|x\rangle$-$|y\rangle$ basis. Show that the photon remains linearly polarized as it travels through the crystal. Explain what is happening to the polarization of the photon as time increases.

4.15. If the Hamiltonian for a spin-$\frac{3}{2}$ particle is given by

$$\hat{H} = \omega_0 \hat{S}_x$$

and at time $t = 0$ $|\psi(0)\rangle = |\frac{3}{2}, \frac{3}{2}\rangle$, determine the probability that the particle is in the state $|\frac{3}{2}, -\frac{3}{2}\rangle$ at time t. Evaluate this probability when $t = \pi/\omega_0$ and explain your result. *Suggestion:* See Problem 3.23 for the eigenstates of $\hat{S}_x$.

4.16. The lifetime of hydrogen in the $2p$ state to decay to the $1s$ ground state is 1.6×10^{-9} s [see (14.169)]. Estimate the uncertainty ΔE in energy of this excited state. What is the corresponding linewidth in angstroms?