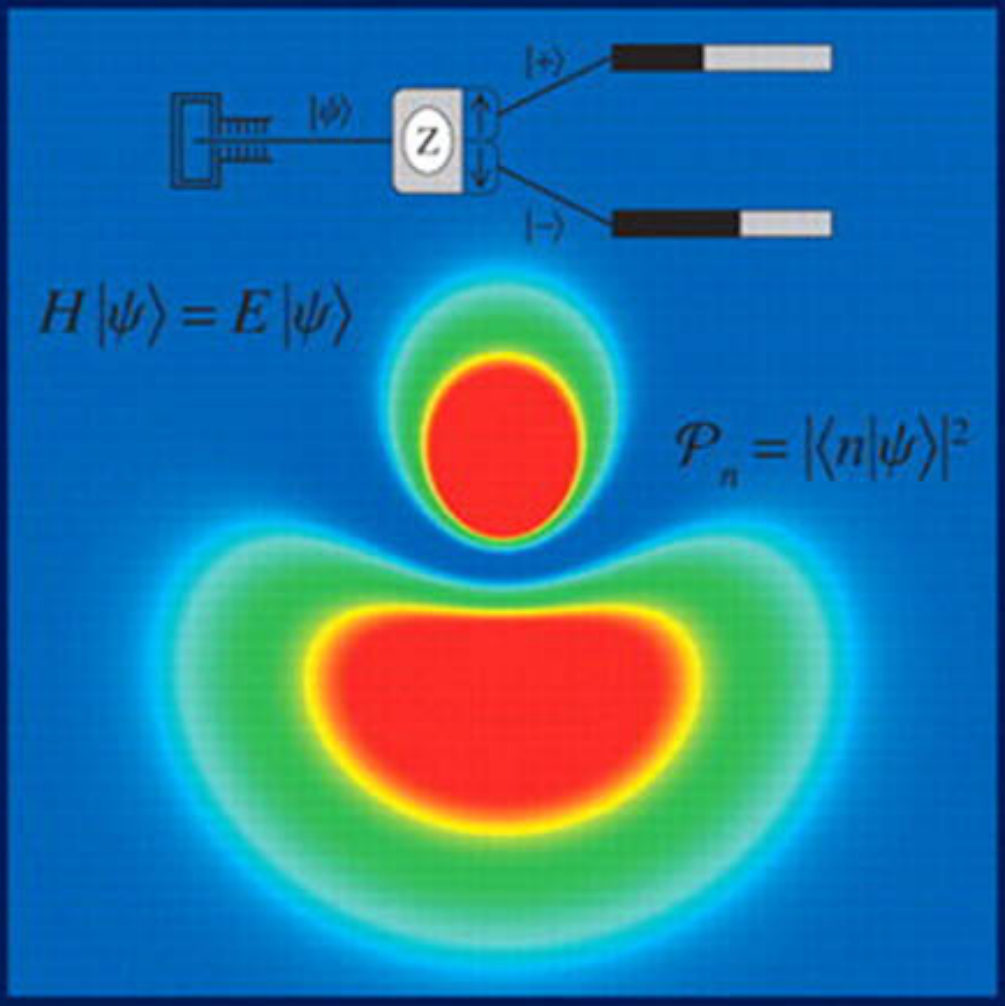
QUANTUM MECHANICS



DAVID H. MCINTYRE

QUANTUM MECHANICS
A Paradigms Approach

This page intentionally left blank

QUANTUM MECHANICS
A Paradigms Approach

David H. McIntyre
Oregon State University

with contributions from Corinne A. Manogue, Janet Tate
and the Paradigms in Physics group at Oregon State University

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Publisher: Jim Smith
Editorial Manager: Laura Kenney
Senior Project Editor: Katie Conley
Assistant Editors: Peter Alston and Steven Le
Senior Marketing Manager: Kerry McGinnis
Managing Editor: Corinne Benson
Production Project Manager: Mary O'Connell
Production Management and Composition: Element LLC
Cover Design: Mark Ong
Manufacturing Buyer: Kathy Sleys
Manager, Rights and Permissions: Zina Arabia
Manager, Cover Visual Research & Permissions: Karen Sanatar
Printer and Binder: Courier, Westford
Cover Printer: Courier, Westford
Cover Images: David H. McIntyre

Library of Congress Cataloging-in-Publication Data
McIntyre, David H.
Quantum mechanics : a paradigms approach / David H. McIntyre ; with contributions from Corinne A. Manogue, Janet Tate, and the Paradigms in Physics group at Oregon State University.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-321-76579-6
ISBN-10: 0-321-76579-6
1. Quantum theory. 2. Mechanics. I. Manogue, Corinne A. II. Tate, Janet. III. Oregon State University. IV. Title.
QC174.12.M3785 2012
530.12--dc23
2011039322

ISBN 10: 0-321-76579-6
ISBN 13: 978-0-321-76579-6

PEARSON

www.pearsonhighered.com 1 2 3 4 5 6 7 8 9 10—CRW—16 15 14 13 12 11

Brief Contents

This page intentionally left blank

Contents

This page intentionally left blank

Preface

This text is designed to introduce undergraduates at the junior and senior levels to quantum mechanics. The text is an outgrowth of the new physics major curriculum developed by the Paradigms in Physics program at Oregon State University. This new curriculum distributes material from the subdisciplines throughout the two upper-division years and provides students with a more gradual transition between introductory and advanced levels. We have also incorporated and developed modern pedagogical strategies to help improve student learning. This text covers the quantum mechanical aspects of our curriculum in a way that can also be used in traditional curricula, but that still preserves the advantages of the Paradigms approach to the ordering of materials and the use of student engagement activities.

PARADIGMS PROGRAM

The Paradigms project began in 1997, when the Department of Physics at Oregon State University began an extensive revision of the upper-division physics major. In an effort to encourage students to draw connections between the subdisciplines of physics, the structure of the Paradigms has been crafted to mimic the organization of expert physics knowledge. Students are presented with a model of how physicists organize their understanding of physical phenomena and problem solving. Each of the nine short junior-year Paradigms courses focuses on a specific paradigm or class of physics problems that serves as the centerpiece of the course and on which different tools and skills are built. In the senior year, students resume a more traditional curriculum, taking six capstone courses in the traditional disciplines. This curriculum incorporates a diverse set of student activities that allow students to stay actively engaged in the classroom and to work together in constructing their understanding of physics. Computer resources are used frequently to help students visualize the systems they are studying.

CONTENT AND APPROACH

Quantum mechanics is integrated into four of the junior-year Paradigms courses and one senior-year capstone course at Oregon State University. This text includes all the quantum mechanics topics covered in those five courses. We adopt a "spins-first" approach by introducing quantum mechanics through the analysis of sequential Stern-Gerlach spin measurements. This approach is based upon previous presentations of spin systems by Feynman, Leighton, and Sands; Cohen-Tannoudji, Diu, and Laloe; Sakurai; and Townsend. The aim of the spins-first approach is twofold: (1) To immediately immerse students in the inherently **quantum** mechanical aspects of physics by focusing on simple measurements that have no classical explanation, and (2) To give students early and extensive experience with the **mechanics** of quantum mechanics in the forms of Dirac and matrix notation.

The simplicity of the spin-1/2 and spin-1 systems allows the students to focus on these new features, which run counter to classical mechanics.

The first three chapters of this text deal exclusively with spin systems and extensions to general two- and three-state quantum mechanical systems. The basic postulates of quantum mechanics are illustrated through their manifestation in the Stern-Gerlach experiments. After these three chapters, students have the tools to tackle any quantum mechanical problem presented in Dirac or matrix notation. After a brief interlude into quantum spookiness (the EPR Paradox and Schrödinger's cat) in Chapter 4, we tackle the traditional wave function aspects of quantum mechanics. We present several quantum systems—a particle in a box, on a ring, on a sphere, the hydrogen atom, and the harmonic oscillator—and emphasize their common features and their connections to the basic postulates. The differential equations of angular momentum and the hydrogen atom radial problem are solved in detail to expose students to the rigor of series solutions, though we stress that these are again eigenvalue equations, no different in principle from the spin eigenvalue equations. Whenever possible, we continue the use of Dirac notation and matrix notation learned in the spin chapters, emphasizing the importance of fluency in multiple representations. We build upon the spins-first approach by using the spin-1/2 example to introduce perturbation theory, the addition of angular momentum, and identical particles.

USAGE

At Oregon State University, the content of this text is taught in five courses as shown below.

<table>
<tr><th colspan="4">Junior-Year Paradigms Courses</th></tr>
<tr><th>Spin and Quantum Measurement</th><th>Waves</th><th>Central Forces</th><th>Period Systems</th></tr>
<tr><td>1. Stern-Gerlach Experiments
2. Operators and Measurement
3. Schrödinger Time Evolution
4. Quantum Spookiness</td><td>Mechanical waves and EM waves
5. Quantized Energies: Particle in a Box
6. Unbound States</td><td>Planetary orbits
7. Angular Momentum
8. Hydrogen Atom</td><td>Coupled Oscillations
15. Periodic Systems</td></tr>
<tr><th colspan="4">Senior-Year Quantum Mechanics Capstone Course</th></tr>
<tr><td>9. Harmonic Oscillator
10. Perturbation Theory</td><td>11. Hyperfine Structure and the Addition of Angular Momentum
12. Perturbation of Hydrogen</td><td>13. Identical Particles
14. Time-Dependent Perturbation Theory</td><td>16. Modern Applications</td></tr>
</table>

For a traditional curriculum, the content of this text would cover a full-year course, either two semesters or three quarters. A proposed weekly outline for two 15-week semesters or three 10-week quarters is shown below.

Week	Chapter	Topics
1	1	Stern-Gerlach experiment, Quantum State Vectors, Bra-ket notation
2	1	Matrix notation, General Quantum Systems
3	2	Operators, Measurement, Commuting Observables
4	2	Uncertainty Principle, $\mathbf{S}^2$ Operator, Spin-1 System
5	3	Schrödinger Equation, Time Evolution
6	3	Spin Precession, Neutrino Oscillations, Magnetic Resonance
7	4	EPR Paradox, Bell's Inequalities, Schrödinger's Cat
8	5	Energy Eigenvalue Equation, Wave Function
9	5	One-Dimensional Potentials, Finite Well, Infinite Well
10	6	Free Particle, Wave Packets, Momentum Space
11	6	Uncertainty Principle, Barriers
12	7	Three-Dimensional Energy Eigenvalue Equation, Separation of Variables
13	7	Angular Momentum, Motion on a Ring and Sphere, Spherical Harmonics
14	8	Hydrogen Atom, Radial Equation, Energy Eigenvalues
15	8	Hydrogen Wave Functions, Spectroscopy
16	9	1-D Harmonic Oscillator, Operator Approach, Energy Spectrum
17	9	Harmonic Oscillator Wave Functions, Matrix Representation
18	9	Momentum Space Wave Functions, Time Dependence, Molecular Vibrations
19	10	Time-Independent Perturbation Theory: Nondegenerate, Degenerate
20	10	Perturbation Examples: Harmonic Oscillator, Stark Effect in Hydrogen
21	11	Hyperfine Structure, Coupled Basis
22	11	Addition of Angular Momenta, Clebsch-Gordan Coefficients
23	12	Hydrogen Atom: Fine Structure, Spin-Orbit, Zeeman Effect
24	13	Identical Particles, Symmetrization, Helium Atom
25	14	Time-Dependent Perturbation Theory, Harmonic Perturbation
26	14	Radiation, Selection Rules
27	15	Periodic Potentials, Bloch's Theorem
28	15	Dispersion Relation, Density of States, Semiconductors
29	16	Modern Applications of Quantum Mechanics, Laser Cooling and Trapping
30	16	Quantum Information Processing

AUDIENCE AND EXPECTED BACKGROUND

The intended audience is junior and senior physics majors, who are expected to have taken intermediate-level courses in modern physics and linear algebra. No other upper-level physics or mathematics courses are required. For our own students, we review matrix algebra in a seven contact hour "preface" course that precedes the Paradigms courses that teach quantum mechanics. The material for that preface course is in Appendix C. The material in Appendix B summarizes an earlier Paradigms course on oscillations, and the material in Appendix D summarizes the classical wave part of the Paradigms course on waves.

STUDENT ACTIVITIES AND WEBSITE

Student engagement activities are an integral part of the Paradigms curriculum. All of the activities that we have developed are freely available on our wiki website:

http://physics.oregonstate.edu/portfolioswiki

The wiki contains a wealth of information about the Paradigms project, the courses we teach, and the materials we have developed. Details about individual activities include descriptions, student handouts, instructor's guides, advice about how to use active engagement strategies, videos of classroom practice, narratives of classroom activities, and comments from users—both internal and external to Oregon State University. This is a dynamic website that is continually updated as we develop new activities and improve existing ones. We encourage you to visit the website and join the community. E-mail us with corrections, additions, and suggestions.

Each of the quantum mechanics activities that we use in our five courses is referenced in the resource section at the end of the appropriate chapter in the text. The quantum mechanics activities are collected within the wiki website with a direct link:

www.physics.oregonstate.edu/qmactivities

These activities include different types of activities such as computer-based activities, group activities, and class response activities. The most extensive activity is a computer simulation of Stern-Gerlach experiments. This SPINS software is a full-featured, menu-driven application that allows students to simulate successive Stern-Gerlach measurements and explore incompatible observables, eigenstate expansions, interference, and quantum dynamics. The use of the SPINS software facilitates our spins-first approach. The beauty of the simulation is that students steeped in classical physics perform a foundational quantum experiment and learn the most fascinating and counterintuitive aspects of quantum mechanics at an early stage.

ACKNOWLEDGMENTS

This work is the product of a broad and energetic community of educators and students within the Paradigms in Physics program. I thank all of our students for their hard work, insights, and innumerable suggestions. My colleagues Corinne Manogue and Janet Tate have developed some of the courses upon which this text is based. They have worked with me throughout the writing of this text and I am indebted to them for their valuable contributions. I gratefully acknowedge my fellow faculty who have developed and taught in the new curriculum: Dedra Demaree, Tevian Dray, Tomasz Giebultowicz, Elizabeth Gire, William Hetherington, Henri Jansen, Kenneth Krane, Yun-Shik Lee, Victor Madsen, Ethan Minot, Oksana Ostroverkhova, David Roundy, Philip Siemens, Albert Stetz, William

Warren, and Allen Wasserman. I would also like to acknowledge the important contributions of early teaching assistants Kerry Browne, Jason Janesky, Cheryl Klipp, Katherine Meyer, Steve Sahyun, and Emily Townsend—their expertise, dedication, and enthusiasm were above and beyond the call of duty. The many subsequent teaching assistants have also been enthusiastic and valued contributors. I also thank those who have contributed in various ways to the development of activities: Mario Belloni, Tim Budd, Wolfgang Christian, Paco Esquembre, Lichun Jia, and Shannon Mayer. I particularly thank Daniel Schroeder for sharing his original SPINS software. I acknowledge useful and constructive feedback from Jeffrey Dunham, Joshua Folk, Rubin Landau, Edward (Joe) Redish, Joseph Rothberg, Homeyra Sadaghiani, Daniel Schroeder, Chandralekha Singh, and Daniel Styer. The Paradigms advisory committee has also provided valuable feedback and I acknowledge David Griffiths, Bruce Mason, William McCallum, Harriett Platsek, and Michael Wittmann for their help. I am grateful to the successive Physics Department chairs, Kenneth Krane and Henri Jansen, and Deans Fred Horne and Sherman Bloomer at Oregon State University for their endorsement of the Paradigms project.

This material is based on work supported by the National Science Foundation under Grant Nos. 9653250, 0231194, and 0618877. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. I thank Duncan McBride and Jack Hehn for their encouragement and support of our endeavor.

Jim Smith at Addison Wesley has been enthusiastic about this project from the early stages. Peter Alston has navigated me through the editorial process with skill and patience. I am grateful to them and also to Katie Conley, Steven Le, and the rest of the staff at Addison Wesley for their work to produce this text.

David H. McIntyre
Corvallis, Oregon
November 2011

This page intentionally left blank

Prologue

It was a dark and stormy night. Erwin huddled under his covers as he had done numerous times that summer. As the wind and rain lashed at the window, he feared having to retreat to the storm cellar once again. The residents of Erwin's apartment building sought shelter whenever there were threats of tornadoes in the area. While it was safe down there, Erwin feared the ridicule he would face once again from the other school boys. In the rush to the cellar, Erwin seemed to always end up with a random pair of socks, and the other boys teased him about it mercilessly.

Not that Erwin hadn't tried hard to solve this problem. He had a very simple collection of socks—black or white, for either school or play; short or long, for either trousers or lederhosen. After the first few teasing episodes from the other boys, Erwin had sorted his socks into two separate drawers. He placed all the black socks in one drawer and all the white socks in another drawer. Erwin figured he could determine an individual sock's length in the dark of night simply by feeling it, but he had to have them presorted into white and black because the apartment generally lost power before the call to the shelter.

Unfortunately, Erwin found that this presorting of the socks by color was ineffective. Whenever he reached into the white sock drawer and chose two long socks, or two short socks, there was a 50% probability of any one sock being black or white. The results from the black sock drawer were the same. The socks seemed to have "forgotten" the color that Erwin had determined previously.

Erwin also tried sorting the socks into two drawers based upon their length, without regard to color. When he chose black or white socks from these long and short drawers, the socks had also "forgotten" whether they were long or short.

After these fruitless attempts to solve his problem through experiments, Erwin decided to save himself the fashion embarrassment, and he replaced his sock collection with a set of medium length brown socks. However, he continued to ponder the mysteries of the socks throughout his childhood.

After many years of daydreaming about the mystery socks, Erwin Schrödinger proposed his theory of "Quantum Socks" and become famous. And that is the beginning of the story of the quantum socks.

The End.

Farfetched?? You bet. But Erwin's adventure with his socks is the way quantum mechanics works. Read on.

This page intentionally left blank

CHAPTER

1 Stern-Gerlach Experiments

It was not a dark and stormy night when Otto Stern and Walther Gerlach performed their now famous experiment in 1922. The Stern-Gerlach experiment demonstrated that measurements on microscopic or quantum particles are not always as certain as we might expect. Quantum particles behave as mysteriously as Erwin's socks—sometimes forgetting what we have already measured. Erwin's adventure with the mystery socks is farfetched because you know that everyday objects do not behave like his socks. If you observe a sock to be black, it remains black no matter what other properties of the sock you observe. However, the Stern-Gerlach experiment goes against these ideas. Microscopic or quantum particles do not behave like the classical objects of your everyday experience. The act of observing a quantum particle affects its measurable properties in a way that is foreign to our classical experience.

In these first three chapters, we focus on the Stern-Gerlach experiment because it is a conceptually simple experiment that demonstrates many basic principles of quantum mechanics. We discuss a variety of experimental results and the quantum theory that has been developed to predict those results. The mathematical formalism of quantum mechanics is based upon six postulates that we will introduce as we develop the theoretical framework. (A complete list of these postulates is in Section 1.5.) We use the Stern-Gerlach experiment to learn about quantum mechanics theory for two primary reasons: (1) It demonstrates how quantum mechanics works in principle by illustrating the postulates of quantum mechanics, and (2) it demonstrates how quantum mechanics works in practice through the use of Dirac notation and matrix mechanics to solve problems. By using a simple example, we can focus on the principles and the new mathematics, rather than having the complexity of the physics obscure these new aspects.

1.1 ■ STERN-GERLACH EXPERIMENT

In 1922 Otto Stern and Walther Gerlach performed a seminal experiment in the history of quantum mechanics. In its simplest form, the experiment consisted of an oven that produced a beam of neutral atoms, a region of space with an inhomogeneous magnetic field, and a detector for the atoms, as depicted in Fig. 1.1. Stern and Gerlach used a beam of silver atoms and found that the beam was split into two in its passage through the magnetic field. One beam was deflected upwards and one downwards in relation to the direction of the magnetic field gradient.

To understand why this result is so at odds with our classical expectations, we must first analyze the experiment classically. The results of the experiment suggest an interaction between a neutral particle and a magnetic field. We expect such an interaction if the particle possesses a magnetic moment $\boldsymbol{\mu}$. The potential energy of this interaction is $E = -\boldsymbol{\mu}\cdot\mathbf{B}$, which results in a force $\mathbf{F} = \nabla(\boldsymbol{\mu}\cdot\mathbf{B})$. In the

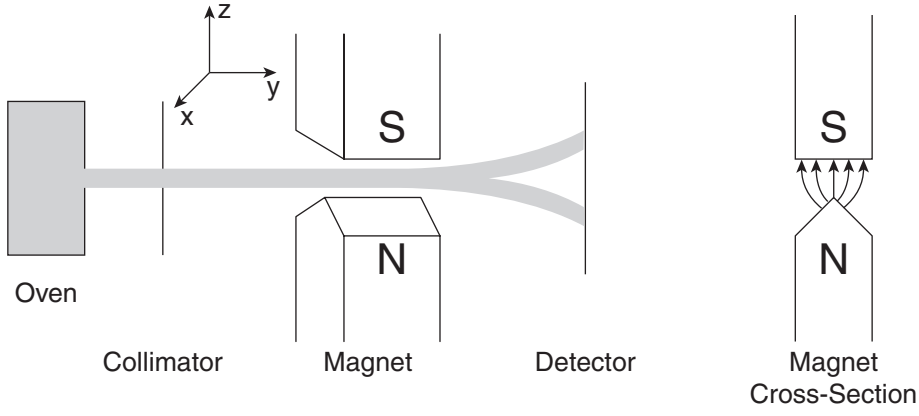


FIGURE 1.1 Stern-Gerlach experiment to measure the spin component of neutral particles along the z-axis. The magnet cross section at right shows the inhomogeneous field used in the experiment.

Stern-Gerlach experiment, the magnetic field gradient is primarily in the z-direction, and the resulting z-component of the force is

$$\begin{aligned} F_z &= \frac{\partial}{\partial z}(\boldsymbol{\mu}\cdot\mathbf{B}) \\ &\cong \mu_z \frac{\partial B_z}{\partial z}. \end{aligned} \tag{1.1}$$

This force is perpendicular to the direction of motion and deflects the beam in proportion to the component of the magnetic moment in the direction of the magnetic field gradient.

Now consider how to understand the origin of the atom's magnetic moment from a classical viewpoint. The atom consists of charged particles, which, if in motion, can produce loops of current that give rise to magnetic moments. A loop of area A and current I produces a magnetic moment

$$\mu = IA \tag{1.2}$$

in MKS units. If this loop of current arises from a charge q traveling at speed v in a circle of radius r, then

$$\begin{aligned} \mu &= \frac{q}{2\pi r/v}\pi r^2 \\ &= \frac{qrv}{2} \\ &= \frac{q}{2m}L, \end{aligned} \tag{1.3}$$

where $L = mrv$ is the orbital angular momentum of the particle. In the same way that the earth revolves around the sun and rotates around its own axis, we can also imagine a charged particle in an atom having **orbital angular momentum L** and a new property, the **intrinsic angular momentum**, which we label **S** and call **spin**. The intrinsic angular momentum also creates current loops, so we expect a similar relation between the magnetic moment $\boldsymbol{\mu}$ and **S**. The exact calculation

involves an integral over the charge distribution, which we will not do. We simply assume that we can relate the magnetic moment to the intrinsic angular momentum in the same fashion as Eq. (1.3), giving

$$\boldsymbol{\mu} = g\frac{q}{2m}\mathbf{S}, \tag{1.4}$$

where the dimensionless **gyroscopic ratio** g contains the details of that integral.

A silver atom has 47 electrons, 47 protons, and 60 or 62 neutrons (for the most common isotopes). The magnetic moments depend on the inverse of the particle mass, so we expect the heavy protons and neutrons ($\approx 2000\, m_e$) to have little effect on the magnetic moment of the atom and so we neglect them. From your study of the periodic table in chemistry, you recall that silver has an electronic configuration $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 4d^{10} 5s^1$, which means that there is only the lone $5s$ electron outside of the closed shells. The electrons in the closed shells can be represented by a spherically symmetric cloud with no orbital or intrinsic angular momentum (unfortunately we are injecting some quantum mechanical knowledge of atomic physics into this classical discussion). That leaves the lone $5s$ electron as a contributor to the magnetic moment of the atom as a whole. An electron in an s state has no orbital angular momentum, but it does have spin. Hence the magnetic moment of this electron, and therefore of the entire neutral silver atom, is

$$\boldsymbol{\mu} = -g\frac{e}{2m_e}\mathbf{S}, \tag{1.5}$$

where e is the magnitude of the electron charge. The classical force on the atom can now be written as

$$F_z \cong -g\frac{e}{2m_e}S_z\frac{\partial B_z}{\partial z}. \tag{1.6}$$

The deflection of the beam in the Stern-Gerlach experiment is thus a measure of the component (or projection) S_z of the spin along the z-axis, which is the orientation of the magnetic field gradient.

If we assume that the $5s$ electron of each atom has the same magnitude $|\mathbf{S}|$ of the intrinsic angular momentum or spin, then classically we would write the z-component as $S_z = |\mathbf{S}|\cos\theta$, where θ is the angle between the z-axis and the direction of the spin $\mathbf{S}$. In the thermal environment of the oven, we expect a random distribution of spin directions and hence all possible angles θ. Thus we expect some continuous distribution (the details are not important) of spin components from $S_z = -|\mathbf{S}|$ to $S_z = +|\mathbf{S}|$, which would yield a continuous spread in deflections of the silver atomic beam. Rather, the experimental result that Stern and Gerlach observed was that there are only two deflections, indicating that there are only two possible values of the z-component of the electron spin. The magnitudes of these deflections are consistent with values of the spin component of

$$S_z = \pm\frac{\hbar}{2}, \tag{1.7}$$

where $\hbar$ is Planck's constant h divided by 2π and has the numerical value

$$\begin{aligned}\hbar &= 1.0546 \times 10^{-34}\ \text{J•s} \\ &= 6.5821 \times 10^{-16}\ \text{eV•s}.\end{aligned} \tag{1.8}$$

This result of the Stern-Gerlach experiment is evidence of the **quantization** of the electron's spin angular momentum component along an axis. This quantization is at odds with our classical

expectations for this measurement. The factor of 1/2 in Eq. (1.7) leads us to refer to this as a **spin-1/2** system.

In this example, we have chosen the z-axis along which to measure the spin component, but there is nothing special about this direction in space. We could have chosen any other axis and we would have obtained the same results.

Now that we know the fine details of the Stern-Gerlach experiment, we simplify the experiment for the rest of our discussions by focusing on the essential features. A simplified schematic representation of the experiment is shown in Fig. 1.2, which depicts an oven that produces the beam of atoms, a Stern-Gerlach device with two output ports for the two possible values of the spin component, and two counters to detect the atoms leaving the output ports of the Stern-Gerlach device. The Stern-Gerlach device is labeled with the axis along which the magnetic field is oriented. The up and down arrows indicate the two possible measurement results for the device; they correspond respectively to the results $S_z = \pm\hbar/2$ in the case where the field is oriented along the z-axis. There are only two possible results in this case, so they are generally referred to as **spin up** and **spin down**. The physical quantity that is measured, S_z in this case, is called an **observable**. In our detailed discussion of the experiment above, we chose the field gradient in such a manner that the spin up states were deflected upwards. In this new simplification, the deflection itself is not an important issue. We simply label the output port with the desired state and count the particles leaving that port. The Stern-Gerlach device sorts (or filters, selects or analyzes) the incoming particles into the two possible outputs $S_z = \pm\hbar/2$ in the same way that Erwin sorted his socks according to color or length. We follow convention and refer to a Stern-Gerlach device as an **analyzer**.

In Fig. 1.2, the input and output beams are labeled with a new symbol called a **ket**. We use the ket $|+\rangle$ as a mathematical representation of the quantum state of the atoms that exit the upper port corresponding to $S_z = +\hbar/2$. The lower output beam is labeled with the ket $|-\rangle$, which corresponds to $S_z = -\hbar/2$, and the input beam is labeled with the more generic ket $|\psi\rangle$. The kets are representations of the quantum states. They are used in mathematical expressions and they represent all the information that we can know about the state. This ket notation was developed by Paul A. M. Dirac and is central to the approach to quantum mechanics that we take in this text. We will discuss the mathematics of these kets in full detail later. With regard to notation, you will find many different ways of writing the same ket. The symbol within the ket brackets is any simple label to distinguish the ket from other different kets. For example, the kets $|+\rangle$, $|+\hbar/2\rangle$, $|S_z = +\hbar/2\rangle$, $|+\hat{\mathbf{z}}\rangle$, and $|\uparrow\rangle$ are all equivalent ways of writing the same thing, which in this case signifies that we have measured the z-component of the spin and found it to be $+\hbar/2$ or spin up. Though we may label these kets in different ways, they all refer to the same physical state and so they all behave the same mathematically. The symbol $|\pm\rangle$ refers to both the $|+\rangle$ and $|-\rangle$ kets. The first postulate of quantum mechanics tells us that kets in general describe the quantum state mathematically and that they contain all the information that we can know about the state. We denote a general ket as $|\psi\rangle$.

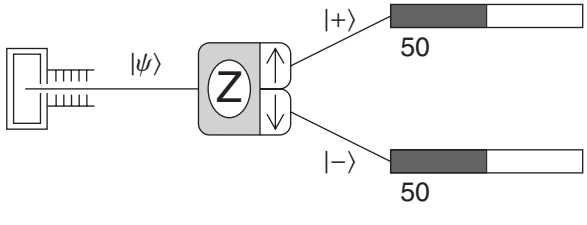


FIGURE 1.2 Simplified schematic of the Stern-Gerlach experiment, depicting a source of atoms, a Stern-Gerlach analyzer, and two counters.

Postulate 1

The state of a quantum mechanical system, including all the information you can know about it, is represented mathematically by a normalized ket $|\psi\rangle$.

We have chosen the particular simplified schematic representation of the Stern-Gerlach experiment shown in Fig. 1.2, because it is the same representation used in the SPINS software program that you may use to simulate these experiments. The SPINS program allows you to perform all the experiments described in this text. This software is freely available, as detailed in Resources at the end of the chapter. In the SPINS program, the components are connected with simple lines to represent the paths the atoms take. The directions and magnitudes of deflections of the beams in the program are not relevant. That is, whether the spin up output beam is drawn as deflected upwards, downwards, or not at all, is not relevant. The labeling on the output port is enough to tell us what that state is. Thus the extra ket label $|+\rangle$ on the spin up output beam in Fig. 1.2 is redundant and will be dropped soon.

The SPINS program permits alignment of Stern-Gerlach analyzing devices along all three axes and also at any angle ϕ measured from the x-axis in the x-y plane. This would appear to be difficult, if not impossible, given that the atomic beam in Fig. 1.1 is directed along the y-axis, making it unclear how to align the magnet in the y-direction and measure a deflection. In our depiction and discussion of Stern-Gerlach experiments, we ignore this technical complication.

In the SPINS program, as in real Stern-Gerlach experiments, the numbers of atoms detected in particular states can be predicted by probability rules that we will discuss later. To simplify our schematic depictions of Stern-Gerlach experiments, the numbers shown for detected atoms are those obtained by using the calculated probabilities without any regard to possible statistical uncertainties. That is, if the theoretically predicted probabilities of two measurement possibilities are each 50%, then our schematics will display equal numbers for those two possibilities, whereas in a real experiment, statistical uncertainties might yield a 55%/45% split in one experiment and a 47%/53% split in another, etc. The SPINS program simulations are designed to give statistical uncertainties, so you will need to perform enough experiments to convince yourself that you have a sufficiently good estimate of the probability (see SPINS Lab 1 for more information on statistics).

Now let's consider a series of simple Stern-Gerlach experiments with slight variations that help to illustrate the main features of quantum mechanics. We first describe the experiments and their results and draw some qualitative conclusions about the nature of quantum mechanics. Then we introduce the formal mathematics of the ket notation and show how it can be used to predict the results of each of the experiments.

1.1.1 ■ Experiment 1

The first experiment is shown in Fig. 1.3 and consists of a source of atoms, two Stern-Gerlach analyzers both aligned along the z-axis, and counters for the output ports of the analyzers. The atomic beam coming into the first Stern-Gerlach analyzer is split into two beams at the output, just like the original experiment. Now instead of counting the atoms in the upper output beam, the spin component is measured again by directing those atoms into the second Stern-Gerlach analyzer. The result of this experiment is that no atoms are ever detected coming out of the lower output port of the second Stern-Gerlach analyzer. All atoms that are output from the upper port of the first analyzer also pass

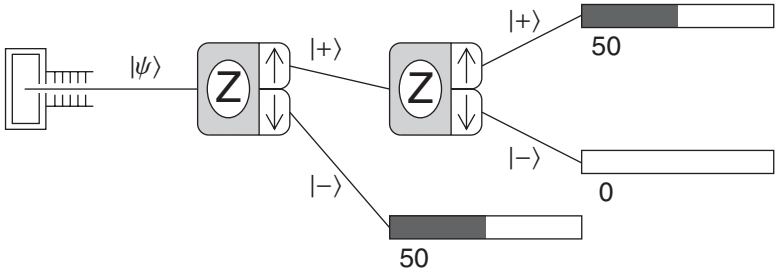


FIGURE 1.3 Experiment 1 measures the spin component along the z-axis twice in succession.

through the upper port of the second analyzer. Thus we say that when the first Stern-Gerlach analyzer measures an atom to have a z-component of spin $S_z = +\hbar/2$, then the second analyzer also measures $S_z = +\hbar/2$ for that atom. This result is not surprising, but it sets the stage for results of experiments to follow.

Though both Stern-Gerlach analyzers in Experiment 1 are identical, they play different roles in this experiment. The first analyzer *prepares* the beam in a particular quantum state $(|+\rangle)$ and the second analyzer *measures* the resultant beam, so we often refer to the first analyzer as a **state preparation device**. By preparing the state with the first analyzer, the details of the source of atoms can be ignored. Thus our main focus in Experiment 1 is what happens at the second analyzer because we know that any atom entering the second analyzer is represented by the $|+\rangle$ ket prepared by the first analyzer. All the experiments we will describe employ a first analyzer as a state preparation device, though the SPINS program has a feature where the state of the atoms coming from the oven is determined but unknown, and the user can perform experiments to determine the unknown state using only one analyzer in the experiment.

1.1.2 ■ Experiment 2

The second experiment is shown in Fig. 1.4 and is identical to Experiment 1 except that the second Stern-Gerlach analyzer has been rotated by 90° to be aligned with the x-axis. Now the second analyzer measures the spin component along the x-axis rather the z-axis. Atoms input to the second analyzer are still represented by the ket $|+\rangle$ because the first analyzer is unchanged. The result of this experiment is that atoms appear at both possible output ports of the second analyzer. Atoms leaving the upper port of the second analyzer have been measured to have $S_x = +\hbar/2$, and atoms leaving

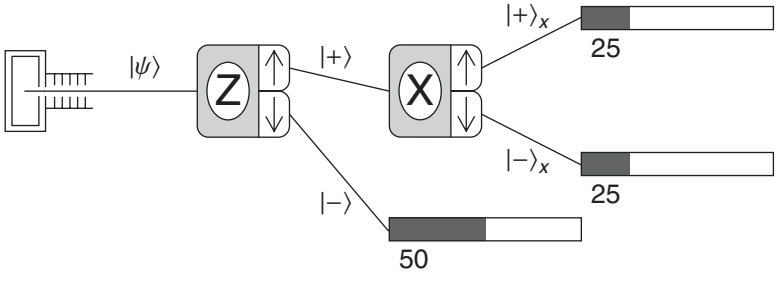


FIGURE 1.4 Experiment 2 measures the spin component along the z-axis and then along the x-axis.

the lower port have $S_x = -\hbar/2$. On average, each of these ports has 50% of the atoms that left the upper port of the first analyzer. As shown in Fig. 1.4, the output states of the second analyzer have new labels $|+\rangle_x$ and $|-\rangle_x$, where the x subscript denotes that the spin component has been measured along the x-axis. We assume that if no subscript is present on the quantum ket (e.g., $|+\rangle$), then the spin component is along the z-axis. This use of the z-axis as the default is a common convention throughout our work and also in much of physics.

A few items are noteworthy about this experiment. First, we notice that there are still only two possible outputs of the second Stern-Gerlach analyzer. The fact that it is aligned along a different axis doesn't affect the fact that we get only two possible results for the case of a spin-1/2 particle. Second, it turns out that the results of this experiment would be unchanged if we used the lower port of the first analyzer. That is, atoms entering the second analyzer in state $|-\rangle$ would also result in half the atoms in each of the $|\pm\rangle_x$ output ports. Finally, we cannot predict which of the second analyzer output ports any particular atom will come out. This can be demonstrated in actual experiments by recording the individual counts out of each port. The arrival sequences at any counter are completely random. We can say only that there is a 50% probability that an atom from the second analyzer will exit the upper analyzer port and a 50% probability that it will exit the lower port. The random arrival of atoms at the detectors can be seen clearly in the SPINS program simulations.

This probabilistic nature is at the heart of quantum mechanics. One might be tempted to say that we just don't know enough about the system to predict which port the atom will exit. That is to say, there may be some other variables, of which we are ignorant, that would allow us to predict the results. Such a viewpoint is known as a **local hidden variable theory**. John Bell proved that such theories are not compatible with the experimental results of quantum mechanics. The conclusion to draw from this is that even though quantum mechanics is a probabilistic theory, it is a complete description of reality. We will have more to say about this in Chapter 4.

Note that the 50% probability referred to above is the probability that an atom input to the second analyzer exits one particular output port. It is not the probability for an atom to pass through the whole system of Stern-Gerlach analyzers. It turns out that the results of this experiment (the 50/50 split at the second analyzer) are the same for any combination of two orthogonal axes of the first and second analyzers.

1.1.3 ■ Experiment 3

Experiment 3, shown in Fig. 1.5, extends Experiment 2 by adding a third Stern-Gerlach analyzer aligned along the z-axis. Atoms entering the third analyzer have been measured by the first Stern-Gerlach analyzer to have spin component up along the z-axis, and by the second analyzer to have spin component up along the x-axis. The third analyzer then measures how many atoms have spin component up or down

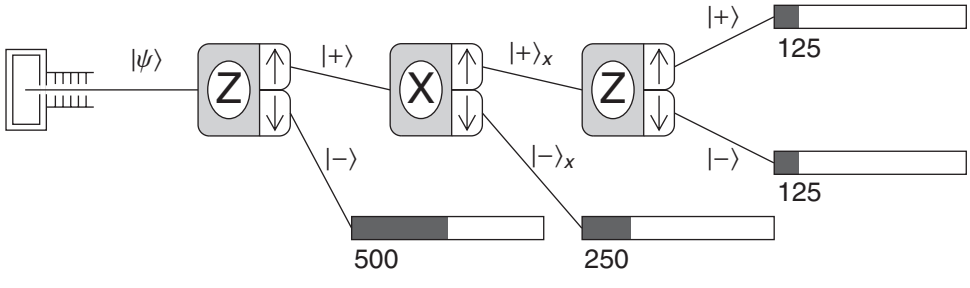


FIGURE 1.5 Experiment 3 measures the spin component three times in succession.

along the z-axis. Classically, one would expect that the final measurement would yield the result spin up along the z-axis, because that was measured at the first analyzer. That is to say: classically the first two analyzers tell us that the atoms have $S_z = +\hbar/2$ and $S_x = +\hbar/2$, so the third measurement must yield $S_z = +\hbar/2$. But that doesn't happen, as Erwin learned with his quantum socks in the Prologue. The quantum mechanical result is that the atoms are split with 50% probability into each output port at the third analyzer. Thus the last two analyzers behave like the two analyzers of Experiment 2 (except with the order reversed), and the fact that there was an initial measurement that yielded $S_z = +\hbar/2$ is somehow forgotten or erased.

This result demonstrates another key feature of quantum mechanics: a measurement disturbs the system. The second analyzer has disturbed the system such that the spin component along the z-axis does not have a unique value, even though we measured it with the first analyzer. Erwin saw this when he sorted, or measured, his socks by color and then by length. When he looked, or measured, a third time, he found that the color he had measured originally was now random—the socks had forgotten about the first measurement. One might ask: Can I be more clever in designing the experiment such that I don't disturb the system? The short answer is no. There is a fundamental incompatibility in trying to measure the spin component of the atom along two different directions. So we say that S_x and S_z are **incompatible observables**. We cannot know the measured values of both simultaneously. The state of the system can be represented by the ket $|+\rangle = |S_z = +\hbar/2\rangle$ or by the ket $|+\rangle_x = |S_x = +\hbar/2\rangle$, but it cannot be represented by a ket $|S_z = +\hbar/2, S_x = +\hbar/2\rangle$ that specifies values of both components. Having said this, it should be said that not all pairs of quantum mechanical observables are incompatible. It is possible to do some experiments without disturbing some of the other aspects of the system. We will see in Section 2.4 that whether two observables are compatible or not is very important in how we analyze a quantum mechanical system.

Not being able to measure both the S_z and S_x spin components is clearly distinct from the classical case where we can measure all three components of the spin vector, which tells us which direction the spin is pointing. In quantum mechanics, the incompatibility of the spin components means that we cannot know which direction the spin is pointing. So when we say "the spin is up," we really mean only that the spin component along that one axis is up (vs. down). The quantum mechanical spin vector cannot be said to be pointing in any given direction. As is often the case, we must check our classical intuition at the door of quantum mechanics.

1.1.4 ■ Experiment 4

Experiment 4 is depicted in Fig. 1.6 and is a slight variation on Experiment 3. Before we get into the details, note a few changes in the schematic drawings. As promised, we have dropped the ket labels on the beams because they are redundant. We have deleted the counters on all but the last analyzer and instead simply blocked the unwanted beams and given the average number of atoms passing from one analyzer to the next. The beam blocks are shown explicitly in Fig. 1.6 but will not be shown after this to be consistent with the SPINS program. Note also that in Experiment 4c two output beams are combined as input to the following analyzer. This is simple in principle and in the SPINS program but can be difficult in practice. The recombination of the beams must be done properly so as to avoid "disturbing" the beams. If you care to read more about this problem, see Feynman's *Lectures on Physics*, volume 3. We will have more to say about the "disturbance" later in Section 2.2. For now we simply assume that the beams can be recombined in the proper manner.

Experiment 4a is identical to Experiment 3. In Experiment 4b, the upper beam of the second analyzer is blocked and the lower beam is sent to the third analyzer. In Experiment 4c, both beams are combined with our new method and sent to the third analyzer. It should be clear from our previous experiments that Experiment 4b has the same results as Experiment 4a. We now ask about the results of

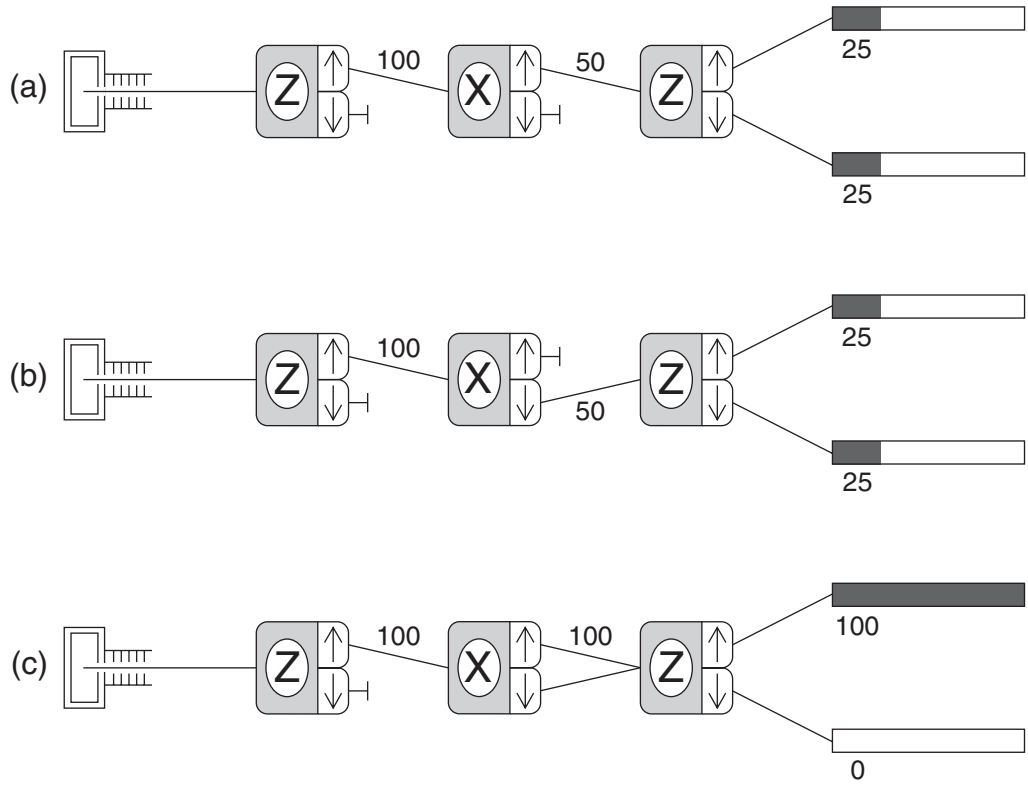


FIGURE 1.6 Experiment 4 measures the spin component three times in succession and uses (a and b) one or (c) two beams from the second analyzer.

Experiment 4c. If we were to use classical probability analysis, then Experiment 4a would indicate that the probability for an atom leaving the first analyzer to take the upper path through the second analyzer and then exit through the upper port of the third analyzer is 25%, where we are now referring to the total probability for those two steps. Likewise, Experiment 4b would indicate that the total probability to take the lower path through the second analyzer and exit through the upper port of the third analyzer is also 25%. Hence the total probability to exit from the upper port of the third analyzer when both paths are available, which is Experiment 4c, would be 50%, and likewise for the exit from the lower port.

However, the quantum mechanical result in Experiment 4c is that all the atoms exit the upper port of the third analyzer and none exits the lower port. The atoms now appear to "remember" that they were initially measured to have spin up along the z-axis. By combining the two beams from the second analyzer, we have avoided the quantum mechanical disturbance that was evident in Experiments 3, 4a, and 4b. The result is now the same as Experiment 1, which means it is as if the second analyzer is not there.

To see how odd this is, look carefully at what happens at the lower port of the third analyzer. In this discussion, we refer to percentages of atoms leaving the first analyzer, because that analyzer is the same in all three experiments. In Experiments 4a and 4b, 50% of the atoms are blocked after the middle analyzer and 25% of the atoms exit the lower port of the third analyzer. In Experiment 4c, 100% of the atoms pass from the second analyzer to the third analyzer, yet fewer atoms come out of the lower port. In fact, no atoms make it through the lower port! So we have a situation where

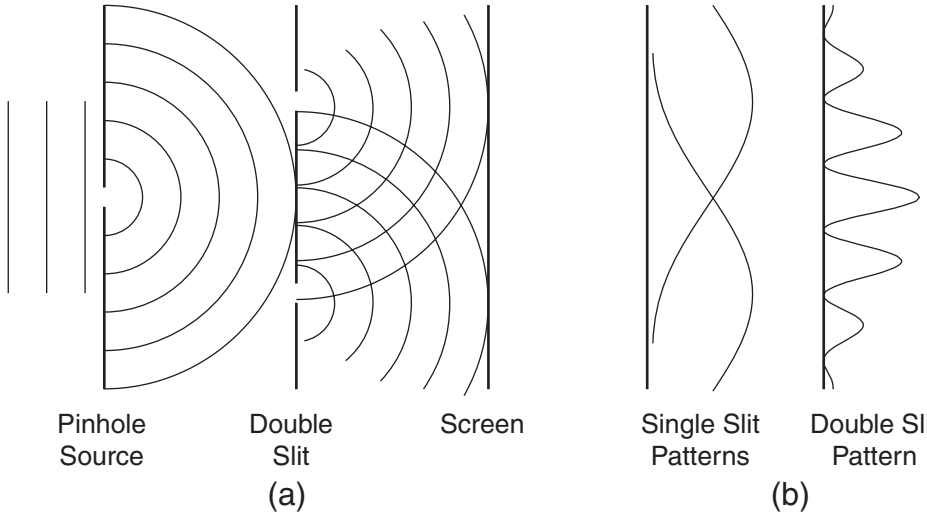


FIGURE 1.7 (a) Young's double-slit interference experiment and (b) resultant intensity patterns observed on the screen, demonstrating single-slit diffraction and double-slit interference.

allowing more ways or paths to reach a counter results in fewer counts. Classical probability theory cannot explain this aspect of quantum mechanics. It is as if you opened a second window in a room to get more sunlight and the room went dark!

However, you may already know of a way to explain this effect. Imagine a procedure whereby combining two effects leads to cancellation rather than enhancement. The concept of wave interference, especially in optics, comes to mind. In the Young's double-slit experiment, light waves pass through two narrow slits and create an interference pattern on a distant screen, as shown in Fig. 1.7. Either slit by itself produces a nearly uniform illumination of the screen, but the two slits combined produce bright and dark interference fringes, as shown in Fig. 1.7(b). We explain this by adding together the electric field vectors of the light from the two slits, then squaring the resultant vector to find the light intensity. We say that we add the amplitudes and then square the total amplitude to find the resultant intensity. See Section 6.6 or an optics textbook for more details about this experiment.

We follow a similar prescription in quantum mechanics. We add together amplitudes and then take the square to find the resultant probability, which opens the door to interference effects. Before we discuss quantum mechanical interference, we must explain what we mean by an amplitude in quantum mechanics and how we calculate it.

1.2 ■ QUANTUM STATE VECTORS

Postulate 1 of quantum mechanics stipulates that kets are to be used for a mathematical description of a quantum mechanical system. These kets are abstract entities that obey many of the rules you know about ordinary spatial vectors. Hence they are called **quantum state vectors**. As we will show in Example 1.3, these vectors must employ complex numbers in order to properly describe quantum mechanical systems. Quantum state vectors are part of a vector space that we call a **Hilbert space**. The dimensionality of the Hilbert space is determined by the physics of the system at hand. In the Stern-Gerlach example, the two possible results for a spin component measurement dictate that the vector space has only two

dimensions. That makes this problem mathematically as simple as it can be, which is why we have chosen to study it. Because the quantum state vectors are abstract, it is hard to say much about what they are, other than how they behave mathematically and how they lead to physical predictions.

In the two-dimensional vector space of a spin-1/2 system, the two kets $|\pm\rangle$ form a basis, just like the **unit vectors** $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ form a basis for describing vectors in three-dimensional space. However, the analogy we want to make with these spatial vectors is only mathematical, not physical. The spatial unit vectors have three important mathematical properties that are characteristic of a basis: the basis vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are **normalized**, **orthogonal**, and **complete**. Spatial vectors are normalized if their magnitudes are unity, and they are orthogonal if they are geometrically perpendicular to each other. The basis is complete if any general vector in the space can be written as a linear superposition of the basis vectors. These properties of spatial basis vectors can be summarized as follows:

$$\begin{aligned} \hat{\mathbf{i}}\cdot\hat{\mathbf{i}} = \hat{\mathbf{j}}\cdot\hat{\mathbf{j}} = \hat{\mathbf{k}}\cdot\hat{\mathbf{k}} &= 1 \qquad \textit{normalization} \\ \hat{\mathbf{i}}\cdot\hat{\mathbf{j}} = \hat{\mathbf{i}}\cdot\hat{\mathbf{k}} = \hat{\mathbf{j}}\cdot\hat{\mathbf{k}} &= 0 \qquad \textit{orthogonality} \\ \mathbf{A} &= a_x\hat{\mathbf{i}} + a_y\hat{\mathbf{j}} + a_z\hat{\mathbf{k}} \qquad \textit{completeness}, \end{aligned} \tag{1.9}$$

where $\mathbf{A}$ is a general vector. Note that the **dot product**, also called the **scalar product**, is central to the description of these properties.

Continuing the mathematical analogy between spatial vectors and abstract vectors, we require that these same properties (at least conceptually) apply to quantum mechanical basis vectors. For the S_z measurement, there are only two possible results, corresponding to the states $|+\rangle$ and $|-\rangle$, so these two states comprise a complete set of basis vectors. This basis is known as the $\boldsymbol{S_z}$ **basis**. We focus on this basis for now and refer to other possible basis sets later. The completeness of the basis kets $|\pm\rangle$ implies that a general quantum state vector $|\psi\rangle$ is a linear combination of the two basis kets:

$$|\psi\rangle = a|+\rangle + b|-\rangle, \tag{1.10}$$

where a and b are complex scalar numbers multiplying each ket. This addition of two kets yields another ket in the same abstract space. The complex scalar can appear either before or after the ket without affecting the mathematical properties of the ket (i.e., $a|+\rangle = |+\rangle a$). It is customary to use the Greek letter ψ (psi) for a general quantum state. You may have seen $\psi(x)$ used before as a quantum mechanical wave function. However, the state vector or ket $|\psi\rangle$ is not a wave function. Kets do not have any spatial dependence as wave functions do. We will study wave functions in Chapter 5.

To discuss orthogonality and normalization (known together as **orthonormality**) we must first define scalar products as they apply to these new kets. As we said above, the machinery of quantum mechanics requires the use of complex numbers. You may have seen other fields of physics use complex numbers. For example, sinusoidal oscillations can be described using the complex exponential $e^{i\omega t}$ rather than $\cos(\omega t)$. However, in such cases, the complex numbers are not required, but are rather a convenience to make the mathematics easier. When using complex notation to describe classical vectors like electric and magnetic fields, the definition of the dot product is generalized slightly, such that one of the vectors is complex conjugated. A similar approach is taken in quantum mechanics. The analog to the complex conjugated vector of classical physics is called a **bra** in the Dirac notation of quantum mechanics. Thus corresponding to a general ket $|\psi\rangle$, there is a bra, or bra vector, which is written as $\langle\psi|$. If a general ket $|\psi\rangle$ is specified as $|\psi\rangle = a|+\rangle + b|-\rangle$, then the corresponding bra $\langle\psi|$ is defined as

$$\langle\psi| = a^*\langle+| + b^*\langle-|, \tag{1.11}$$

where the basis bras $\langle +|$ and $\langle -|$ correspond to the basis kets $|+\rangle$ and $|-\rangle$, respectively, and the coefficients a and b have been complex conjugated.

The scalar product in quantum mechanics is defined as the product of a bra and a ket taken in the proper order—bra first, then ket second:

$$(\langle bra|)(|ket\rangle). \tag{1.12}$$

When the bra and ket are combined together in this manner, we get a bracket (*bra ket*)—*a little physics humor*—that is written in shorthand as

$$\langle bra|ket\rangle. \tag{1.13}$$

Thus, given the basis kets $|+\rangle$ and $|-\rangle$, one inner product, for example, is written as

$$(\langle +|)(|-\rangle) = \langle +|-\rangle \tag{1.14}$$

and so on. Note that we have eliminated the extra vertical bar in the middle. The scalar product in quantum mechanics is generally referred to as an **inner product** or a **projection**.

So how do we calculate the inner product $\langle +|+\rangle$? We do it the same way we calculate the dot product $\hat{\mathbf{i}}\cdot\hat{\mathbf{i}}$. We define it to be unity because we like basis vectors to be unit vectors. There is a little more to it than that, because in quantum mechanics (as we will see shortly) using normalized basis vectors is more rooted in physics than in our personal preferences for mathematical cleanliness. But for all practical purposes, if someone presents a set of basis vectors to you, you can probably assume that they are normalized. So the normalization of the spin-1/2 basis vectors is expressed in this new notation as $\langle +|+\rangle = 1$ and $\langle -|-\rangle = 1$.

Now, what about orthogonality? The spatial unit vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ used for spatial vectors are orthogonal to each other because they are at 90° with respect to each other. That orthogonality is expressed mathematically in the dot products $\hat{\mathbf{i}}\cdot\hat{\mathbf{j}} = \hat{\mathbf{i}}\cdot\hat{\mathbf{k}} = \hat{\mathbf{j}}\cdot\hat{\mathbf{k}} = 0$. For the spin basis kets $|+\rangle$ and $|-\rangle$, there is no spatial geometry involved. Rather, the spin basis kets $|+\rangle$ and $|-\rangle$ are orthogonal in the mathematical sense, which we express with the inner product as $\langle +|-\rangle = 0$. Again, we do not prove to you that these basis vectors are orthogonal, but we assume that a well-behaved basis set obeys orthogonality. Though there is no geometry in this property for quantum mechanical basis vectors, the fundamental idea of orthogonality is the same, so we use the same language—if a general vector "points" in the direction of a basis vector, then there is no component in the "direction" of the other unit vectors.

In summary, the properties of normalization, orthogonality, and completeness can be expressed in the case of a two-state spin-1/2 quantum system as:

$$\boxed{\begin{array}{ll} \left.\begin{array}{l}\langle +|+\rangle = 1\\ \langle -|-\rangle = 1\end{array}\right\} & \textit{normalization}\\ \left.\begin{array}{l}\langle +|-\rangle = 0\\ \langle -|+\rangle = 0\end{array}\right\} & \textit{orthogonality}\\ |\psi\rangle = a|+\rangle + b|-\rangle & \textit{completeness}\end{array}}. \tag{1.15}$$

Note that a product of kets (e.g., $|+\rangle|+\rangle$) or a similar product of bras (e.g., $\langle +|\langle +|$) is meaningless in this new notation, while a product of a ket and a bra in the "wrong" order (e.g., $|+\rangle\langle +|$) has a meaning that we will define in Section 2.2.3. Equations (1.15) are sufficient to define how the basis

kets behave mathematically. Note that the inner product is defined using a bra and a ket, though it is common to refer to the inner product of two kets, where it is understood that one is converted to a bra first. The order does matter, as we will see shortly.

Using this new notation, we can learn a little more about general quantum states and derive some expressions that will be useful later. Consider the general state vector $|\psi\rangle = a|+\rangle + b|-\rangle$. Take the inner product of this ket with the bra $\langle +|$ and obtain

$$\begin{aligned}\langle +|\psi\rangle &= \langle +|\left(a|+\rangle + b|-\rangle\right)\\ &= \langle +|a|+\rangle + \langle +|b|-\rangle\\ &= a\langle +|+\rangle + b\langle +|-\rangle\\ &= a,\end{aligned} \tag{1.16}$$

using the properties that inner products are distributive and that scalars can be moved freely through bras or kets. Likewise, you can show that $\langle -|\psi\rangle = b$. Hence the coefficients multiplying the basis kets are simply the inner products or projections of the general state $|\psi\rangle$ along each basis ket, albeit in an abstract complex vector space rather than the concrete three-dimensional space of normal vectors. Using these results, we rewrite the general state as

$$\begin{aligned}|\psi\rangle &= a|+\rangle + b|-\rangle\\ &= |+\rangle a + |-\rangle b\\ &= |+\rangle\{\langle +|\psi\rangle\} + |-\rangle\{\langle -|\psi\rangle\},\end{aligned} \tag{1.17}$$

where the rearrangement of the second equation again uses the property that scalars (e.g., $a = \langle +|\psi\rangle$) can be moved through bras or kets.

For a general state vector $|\psi\rangle = a|+\rangle + b|-\rangle$, we defined the corresponding bra to be $\langle\psi| = a^*\langle +| + b^*\langle -|$. Thus, the inner product of the state $|\psi\rangle$ with the basis ket $|+\rangle$ taken in the reverse order compared to Eq. (1.16) yields

$$\begin{aligned}\langle\psi|+\rangle &= \langle +|a^*|+\rangle + \langle -|b^*|+\rangle\\ &= a^*\langle +|+\rangle + b^*\langle -|+\rangle\\ &= a^*.\end{aligned} \tag{1.18}$$

Thus, we see that an inner product with the states reversed results in a complex conjugation of the inner product:

$$\langle +|\psi\rangle = \langle\psi|+\rangle^*. \tag{1.19}$$

This important property holds for any inner product. For example, the inner product of two general states is

$$\boxed{\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*}. \tag{1.20}$$

Now we come to a new mathematical aspect of quantum vectors that differs from the use of vectors in classical mechanics. The rules of quantum mechanics (postulate 1) require that all state vectors describing a quantum system be normalized, not just the basis kets. This is clearly different from ordinary spatial vectors, where the length or magnitude of a vector means something and only the unit vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are normalized to unity. This new rule means that in the quantum mechanical state

space only the direction—in an abstract sense—is important. If we apply this normalization requirement to a general state $|\psi\rangle$, then we obtain

$$\begin{aligned}
&\langle\psi|\psi\rangle = \{a^*\langle+| + b^*\langle-|\}\{a|+\rangle + b|-\rangle\} = 1 \\
&\Rightarrow a^*a\langle+|+\rangle + a^*b\langle+|-\rangle + b^*a\langle-|+\rangle + b^*b\langle-|-\rangle = 1 \\
&\Rightarrow a^*a + b^*b = 1 \\
&\Rightarrow |a|^2 + |b|^2 = 1,
\end{aligned} \tag{1.21}$$

or using the expressions for the coefficients obtained above,

$$|\langle+|\psi\rangle|^2 + |\langle-|\psi\rangle|^2 = 1. \tag{1.22}$$

Example 1.1 Normalize the vector $|\psi\rangle = C(1|+\rangle + 2i|-\rangle)$. The complex constant C is often referred to as the **normalization constant**.

To normalize $|\psi\rangle$, we set the inner product of the vector with itself equal to unity and then solve for C—note the requisite complex conjugations

$$\begin{aligned}
1 &= \langle\psi|\psi\rangle \\
&= C^*\{1\langle+| - 2i\langle-|\}C\{1|+\rangle + 2i|-\rangle\} \\
&= C^*C\{1\langle+|+\rangle + 2i\langle+|-\rangle - 2i\langle-|+\rangle + 4\langle-|-\rangle\} \\
&= 5|C|^2 \\
&\Rightarrow |C| = \frac{1}{\sqrt{5}}.
\end{aligned} \tag{1.23}$$

The overall phase of the normalization constant is not physically meaningful (Problem 1.3), so we follow the standard convention and choose it to be real and positive. This yields $C = 1/\sqrt{5}$. The normalized quantum state vector is then

$$|\psi\rangle = \frac{1}{\sqrt{5}}(1|+\rangle + 2i|-\rangle). \tag{1.24}$$

Now comes the crucial element of quantum mechanics. We postulate that each term in the sum of Eq. (1.22) is equal to the **probability** that the quantum state described by the ket $|\psi\rangle$ is measured to be in the corresponding basis state. Thus

$$\mathcal{P}_{S_z=+\hbar/2} = |\langle+|\psi\rangle|^2 \tag{1.25}$$

is the probability that the state $|\psi\rangle$ is found to be in the state $|+\rangle$ when a measurement of S_z is made, meaning that the result $S_z = +\hbar/2$ is obtained. Likewise,

$$\mathcal{P}_{S_z=-\hbar/2} = |\langle-|\psi\rangle|^2 \tag{1.26}$$

is the probability that the measurement yields the result $S_z = -\hbar/2$. The subscript on the probability indicates the measured value. For the spin component measurements, we will usually abbreviate this to, for example, $\mathcal{P}_+$ for an $S_z = +\hbar/2$ result or $\mathcal{P}_{-y}$ for an $S_y = -\hbar/2$ measurement.

We now have a prescription for predicting the outcomes of the experiments we have been discussing. For example, the experiment shown in Fig. 1.8 has the state $|\psi\rangle = |+\rangle$ prepared by the first Stern-Gerlach device and then input to the second Stern-Gerlach device aligned along the z-axis. Therefore the probabilities of measuring the input state $|\psi\rangle = |+\rangle$ to have the two output values are as shown. Because the spin-1/2 system has only two possible measurement results, these two probabilities must sum to unity—there is a 100% probability of recording some value in the experiment. This basic rule of probabilities is why the rules of quantum mechanics require that all state vectors be properly normalized before they are used in any calculation of probabilities. The experimental predictions shown in Fig. 1.8 are an example of the fourth postulate of quantum mechanics, which is presented below.

Postulate 4 (Spin-1/2 system)

The probability of obtaining the value $\pm\hbar/2$ in a measurement of the observable S_z on a system in the state $|\psi\rangle$ is

$$\mathcal{P}_{\pm} = |\langle \pm|\psi\rangle|^2,$$

where $|\pm\rangle$ is the basis ket of S_z corresponding to the result $\pm\hbar/2$.

This is labeled as the fourth postulate because we have written this postulate using the language of the spin-1/2 system, while the general statement of the fourth postulate presented in Section 1.5 requires the second and third postulates of Section 2.1. A general spin component measurement is shown in Fig. 1.9, along with a histogram that compactly summarizes the measurement results.

Because the quantum mechanical probability is found by squaring an inner product, we refer to an inner product, $\langle +|\psi\rangle$ for example, as a **probability amplitude** or sometimes just an **amplitude**; much like a classical wave intensity is found by squaring the wave amplitude. Note that the convention is to put the input or initial state on the right and the output or final state on the left: $\langle out|in\rangle$, so one would read from right to left in describing a problem. Because the probability involves the complex square of the amplitude, and $\langle out|in\rangle = \langle in|out\rangle^*$, this convention is not critical for calculating probabilities. Nonetheless, it is the accepted practice and is important in situations where several amplitudes are combined.

Armed with these new quantum mechanical rules and tools, let's continue to analyze the experiments discussed earlier. Using the experimental results and the new rules we have introduced, we can learn more about the mathematical behavior of the kets and the relationships among them. We will focus on the first two experiments for now and return to the others in the next chapter.

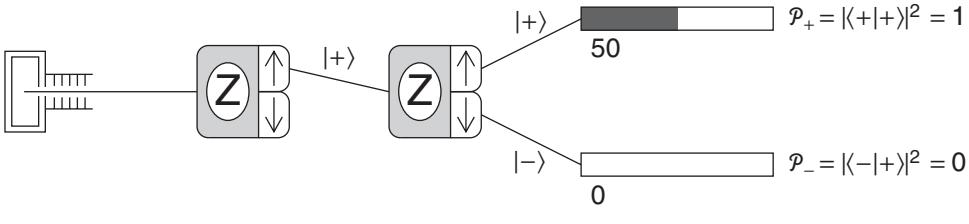


FIGURE 1.8 Probabilities of spin component measurements.

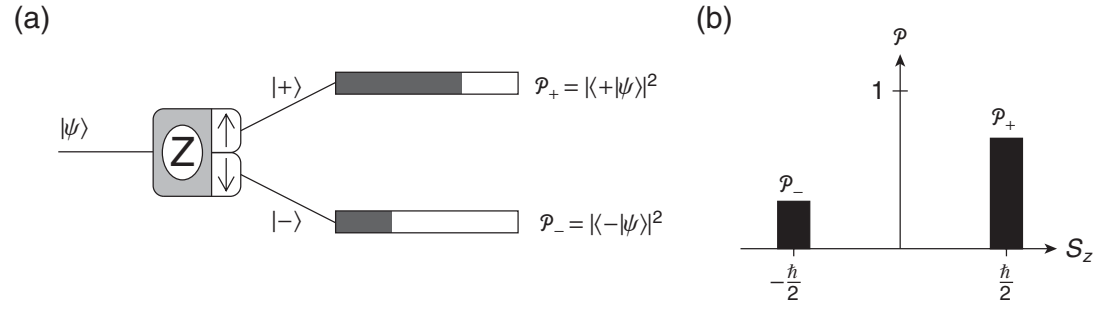


FIGURE 1.9 (a) Spin component measurement for a general input state and (b) histogram of measurement results.

1.2.1 ■ Analysis of Experiment 1

In Experiment 1, the first Stern-Gerlach analyzer prepared the system in the $|+\rangle$ state and the second analyzer later measured this state to be in the $|+\rangle$ state and not in the $|-\rangle$ state. The results of the experiment are summarized in the histogram in Fig. 1.10. We can use the fourth postulate to predict the results of this experiment. We take the inner product of the input state $|+\rangle$ with each of the possible output basis states $|+\rangle$ and $|-\rangle$. Because we know that the basis states are normalized and orthogonal, we calculate the probabilities to be

$$\begin{aligned} \mathcal{P}_+ &= |\langle + | + \rangle|^2 = 1 \\ \mathcal{P}_- &= |\langle - | + \rangle|^2 = 0. \end{aligned} \tag{1.27}$$

These predictions agree exactly with the histogram of experimental results shown in Fig. 1.10. A $|+\rangle$ state is always measured to have $S_z = +\hbar/2$.

1.2.2 ■ Analysis of Experiment 2

In Experiment 2, the first Stern-Gerlach analyzer prepared the system in the $|+\rangle$ state and the second analyzer performed a measurement of the spin component along the x-axis, finding 50% probabilities for each of the two possible states $|+\rangle_x$ and $|-\rangle_x$, as shown in the histogram in Fig. 1.11(a). For this experiment, we cannot predict the results of the measurements, because we do not yet have

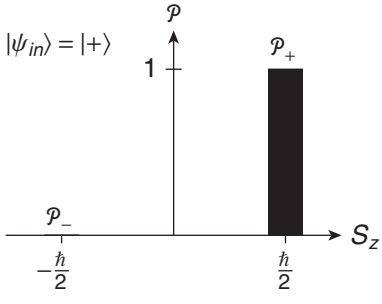


FIGURE 1.10 Histogram of S_z spin component measurements for Experiment 1 with $|\psi_{in}\rangle = |+\rangle$.

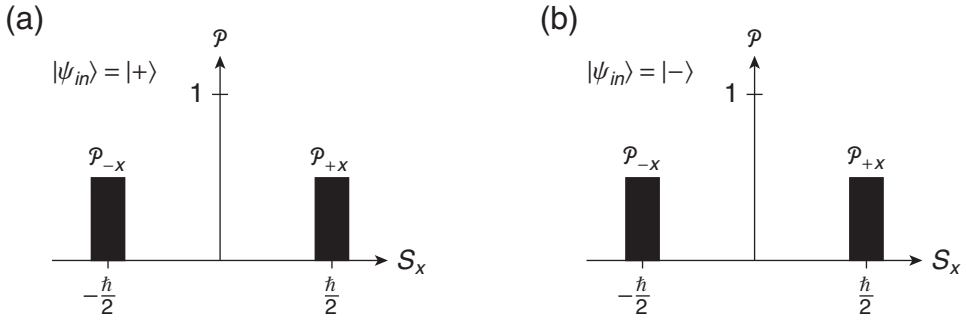


FIGURE 1.11 Histograms of S_x spin component measurements for Experiment 2 for different input states (a) $|\psi_{in}\rangle = |+\rangle$ and (b) $|\psi_{in}\rangle = |-\rangle$.

enough information about how the states $|+\rangle_x$ and $|-\rangle_x$ behave mathematically. Rather, we will use the results of the experiment to determine these states. Recalling that the experimental results would be the same if the first analyzer prepared the system to be in the $|-\rangle$ state [see Fig. 1.11(b)], we have four results for the two experiments:

$$\begin{aligned}
\mathcal{P}_{1,+x} &= |{}_x\langle +|+\rangle|^2 = \tfrac{1}{2} \\
\mathcal{P}_{1,-x} &= |{}_x\langle -|+\rangle|^2 = \tfrac{1}{2} \\
\mathcal{P}_{2,+x} &= |{}_x\langle +|-\rangle|^2 = \tfrac{1}{2} \\
\mathcal{P}_{2,-x} &= |{}_x\langle -|-\rangle|^2 = \tfrac{1}{2}.
\end{aligned} \tag{1.28}$$

Because the kets $|+\rangle$ and $|-\rangle$ form a complete basis, the kets describing the S_x measurement, $|+\rangle_x$ and $|-\rangle_x$, can be written in terms of them. We do not yet know the specific coefficients of the $|\pm\rangle_x$ states, so we use general expressions

$$\begin{aligned}
|+\rangle_x &= a|+\rangle + b|-\rangle \\
|-\rangle_x &= c|+\rangle + d|-\rangle,
\end{aligned} \tag{1.29}$$

and now our task is to use the results of Experiment 2 to determine the coefficients a, b, c, and d. The first measured probability in Eq. (1.28) is

$$\mathcal{P}_{1,+x} = |{}_x\langle +|+\rangle|^2 = \tfrac{1}{2}. \tag{1.30}$$

Using the general expression for $|+\rangle_x$ in Eq. (1.29), we calculate the probability that the $|+\rangle$ input state is measured to be in the $|+\rangle_x$ output state, that is, to have $S_x = +\hbar/2$:

$$\begin{aligned}
\mathcal{P}_{1,+x} &= |{}_x\langle +|+\rangle|^2 \\
&= |\{a^*\langle +| + b^*\langle -|\}|+\rangle|^2 \\
&= |a^*|^2 = |a|^2,
\end{aligned} \tag{1.31}$$

where we convert the $|+\rangle_x$ ket to a bra ${}_x\langle +|$ in order to calculate the inner product. Equating the experimental result in Eq. (1.30) and the prediction in Eq. (1.31), we find

$$|a|^2 = \tfrac{1}{2}. \tag{1.32}$$

Similarly, one can calculate the other three probabilities to arrive at $|b|^2 = |c|^2 = |d|^2 = \frac{1}{2}$. (Problem 1.4) Because each coefficient is complex, each has an amplitude and phase. However, the overall phase of a quantum state vector is not physically meaningful (see Problem 1.3). Only the relative phase between different components of the state vector is physically measurable. Hence, we are free to choose *one* coefficient of each vector to be real and positive without any loss of generality. This allows us to write the desired states as

$$\begin{aligned} |+\rangle_x &= \tfrac{1}{\sqrt{2}}\big[\,|+\rangle + e^{i\alpha}|-\rangle\,\big] \\ |-\rangle_x &= \tfrac{1}{\sqrt{2}}\big[\,|+\rangle + e^{i\beta}|-\rangle\,\big], \end{aligned} \tag{1.33}$$

where α and β are relative phases that we have yet to determine. Note that these states are already normalized because we used all of the experimental results, which reflect the fact that the probability for all possible results of an experiment must sum to unity.

We have used all the experimental results from Experiment 2, but the $|\pm\rangle_x$ kets are still not determined. We need some more information. If we perform Experiment 1 with both analyzers aligned along the x-axis, the results will be as you expect—all $|+\rangle_x$ states from the first analyzer will be measured to have $S_x = +\hbar/2$ at the second analyzer, that is, all atoms exit in the $|+\rangle_x$ state and none in the $|-\rangle_x$. The probability calculations for this experiment are

$$\begin{aligned} \mathcal{P}_{+x} &= |{}_x\langle +|+\rangle_x|^2 = 1 \\ \mathcal{P}_{-x} &= |{}_x\langle -|+\rangle_x|^2 = 0, \end{aligned} \tag{1.34}$$

which tell us mathematically that the $|\pm\rangle_x$ states are orthonormal to each other, just like the $|\pm\rangle$ states. This also implies that the $|\pm\rangle_x$ kets form a basis, the S_x basis, which you might expect because they correspond to the distinct results of a different spin component measurement. The general expressions we used for the $|\pm\rangle_x$ kets are already normalized but are not yet orthogonal. That is the new piece of information we need. The orthogonality condition leads to

$$\begin{aligned} &{}_x\langle -|+\rangle_x = 0 \\ &\tfrac{1}{\sqrt{2}}\big[\langle +| + e^{-i\beta}\langle -|\big]\tfrac{1}{\sqrt{2}}\big[\,|+\rangle + e^{i\alpha}|-\rangle\,\big] = 0 \\ &\tfrac{1}{2}\big[\,1 + e^{i(\alpha-\beta)}\,\big] = 0 \\ &e^{i(\alpha-\beta)} = -1 \\ &e^{i\alpha} = -e^{i\beta}, \end{aligned} \tag{1.35}$$

where the complex conjugation of the second coefficient of the ${}_x\langle -|$ bra should be noted.

We now have an equation relating the remaining coefficients α and β, but we need some more information to determine their values. Unfortunately, there is no more information to be obtained, so we are free to choose the value of the phase α. This freedom comes from the fact that we have required only that the x-axis be perpendicular to the z-axis, which limits the x-axis only to a plane rather than to a unique direction. We follow convention here and choose the phase $\alpha = 0$. Thus we can express the S_x basis kets in terms of the S_z basis kets as

$$\begin{aligned} |+\rangle_x &= \tfrac{1}{\sqrt{2}}\big[\,|+\rangle + |-\rangle\,\big] \\ |-\rangle_x &= \tfrac{1}{\sqrt{2}}\big[\,|+\rangle - |-\rangle\,\big]. \end{aligned} \tag{1.36}$$

We generally use the S_z basis as the preferred basis for writing general states, but we could use any basis we choose. If we were to use the S_x basis, then we could write the $|\pm\rangle$ kets as general states in terms of the $|\pm\rangle_x$ kets. This can be done by solving Eq. (1.36) for the $|\pm\rangle$ kets, yielding

$$\begin{aligned} |+\rangle &= \tfrac{1}{\sqrt{2}}\big[|+\rangle_x + |-\rangle_x\big] \\ |-\rangle &= \tfrac{1}{\sqrt{2}}\big[|+\rangle_x - |-\rangle_x\big]. \end{aligned} \tag{1.37}$$

With respect to the measurements performed in Experiment 2, Eq. (1.37) tells us that the $|+\rangle$ state is a combination of the states $|+\rangle_x$ and $|-\rangle_x$. The coefficients tell us that there is a 50% probability for measuring the spin component to be up along the x-axis, and likewise for the down possibility, which is in agreement with the histogram of measurements shown in Fig. 1.11(a). We must now take a moment to describe carefully what a combination of states, such as in Eqs. (1.36) and (1.37), is and what it is not.

1.2.3 ■ Superposition States

A general spin-1/2 state vector $|\psi\rangle$ can be expressed as a combination of the basis kets $|+\rangle$ and $|-\rangle$

$$|\psi\rangle = a|+\rangle + b|-\rangle. \tag{1.38}$$

We refer to such a combination of states as a **superposition state**. To understand the importance of a quantum mechanical superposition state, consider the particular state

$$|\psi\rangle = \tfrac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \tag{1.39}$$

and measurements on this state, as shown in Fig. 1.12(a). Note that the state $|\psi\rangle$ is none other than the state $|+\rangle_x$ that we found in Eq. (1.36), so we already know what the measurement results are. If we measure the spin component along the x-axis for this state, then we record the result $S_x = +\hbar/2$ with 100% probability (Experiment 1 with both analyzers along the x-axis). If we measure the spin component along the orthogonal z-axis, then we record the two results $S_z = \pm\hbar/2$ with 50% probability each (Experiment 2 with the first and second analyzers along the x- and z-axes, respectively). Based upon this second set of results, one might be tempted to consider the state $|\psi\rangle$ as describing a beam that contains a mixture of atoms with 50% of the atoms in the $|+\rangle$ state and 50% in the $|-\rangle$ state. Such a state is called a **mixed state** and is very different from a superposition state.

To clarify the difference between a mixed state and a superposition state, let's carefully examine the results of experiments on the proposed mixed-state beam, as shown in Fig. 1.12(b). If we measure the spin component along the z-axis, then each atom in the $|+\rangle$ state yields the result $S_z = +\hbar/2$ with 100% certainty and each atom in the $|-\rangle$ state yields the result $S_z = -\hbar/2$ with 100% certainty. The net result is that 50% of the atoms yield $S_z = +\hbar/2$ and 50% yield $S_z = -\hbar/2$. This is exactly the same result as that obtained with all atoms in the $|+\rangle_x$ state, as seen in Fig. 1.12(a). If we instead measure the spin component along the x-axis, then each atom in the $|+\rangle$ state yields the two results $S_x = \pm\hbar/2$ with 50% probability each (Experiment 2 with the first and second analyzers along the z- and x-axes, respectively). The atoms in the $|-\rangle$ state yield the same results. The net result is that 50% of the atoms yield $S_x = +\hbar/2$ and 50% yield $S_x = -\hbar/2$. This is in stark contrast to the results of Experiment 1, which tell us that once we have prepared the state to be $|+\rangle_x$, then subsequent measurements yield $S_x = +\hbar/2$ with certainty, as seen in Fig. 1.12(a).

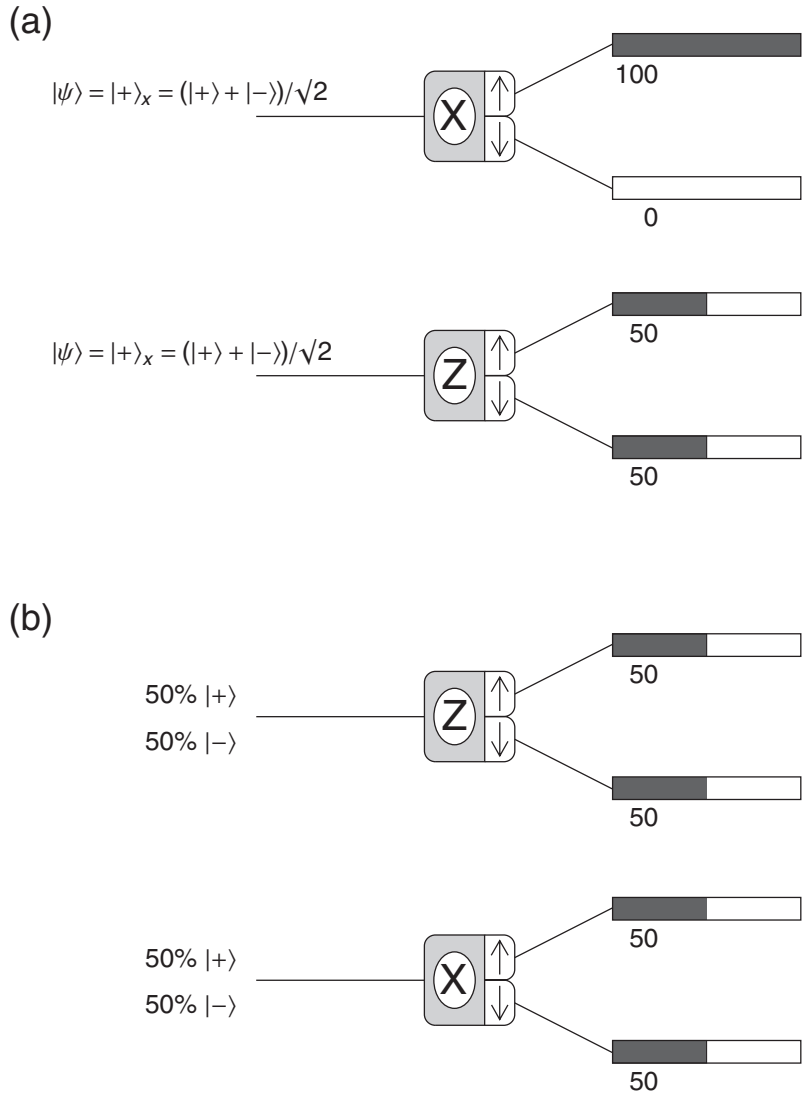


FIGURE 1.12 (a) Superposition state measurements and (b) mixed state measurements.

Hence we must conclude that the system described by the $|\psi\rangle = |+\rangle_x$ state is not a mixed state with some atoms in the $|+\rangle$ state and some in the $|-\rangle$ state. Rather, each atom in the $|+\rangle_x$ beam is in a state that itself is a superposition of the $|+\rangle$ and $|-\rangle$ states. A superposition state is often called a **coherent superposition** because the relative phase of the two terms is important. For example, if the input beam were in the $|-\rangle_x$ state, then there would be a relative minus sign between the two coefficients, which would result in an $S_x = -\hbar/2$ measurement but would not affect the S_z measurement.

We will not have any further need to speak of mixed states, so any combination of states we use is a superposition state. Note that we cannot even write down a ket describing a mixed state. So if someone gives you a quantum state written as a ket, then it must be a superposition state and not a mixed state. The random option in the SPINS program produces a mixed state, while the unknown states are all superposition states.

Example 1.2 Consider the input state

$$|\psi_{in}\rangle = 3|+\rangle + 4|-\rangle. \tag{1.40}$$

Normalize this state vector and find the probabilities of measuring the spin component along the z-axis to be $S_z = \pm\hbar/2$.

To normalize this state, introduce an overall complex multiplicative factor and solve for this factor by imposing the normalization condition:

$$\begin{aligned}
&|\psi_{in}\rangle = C[3|+\rangle + 4|-\rangle] \\
&\langle\psi_{in}|\psi_{in}\rangle = 1 \\
&\{C^*[3\langle+| + 4\langle-|]\}\{C[3|+\rangle + 4|-\rangle]\} = 1 \\
&C^*C[9\langle+|+\rangle + 12\langle+|-\rangle + 12\langle-|+\rangle + 16\langle-|-\rangle] = 1 \\
&C^*C[25] = 1 \\
&|C|^2 = \frac{1}{25}.
\end{aligned} \tag{1.41}$$

Because an overall phase is physically meaningless, we choose C to be real and positive: $C = 1/5$. Hence the normalized input state is

$$|\psi_{in}\rangle = \tfrac{3}{5}|+\rangle + \tfrac{4}{5}|-\rangle. \tag{1.42}$$

The probability of measuring $S_z = +\hbar/2$ is

$$\begin{aligned}
\mathcal{P}_+ &= \left|\langle+|\psi_{in}\rangle\right|^2 \\
&= \left|\langle+|[\tfrac{3}{5}|+\rangle + \tfrac{4}{5}|-\rangle]\right|^2 \\
&= \left|\tfrac{3}{5}\langle+|+\rangle + \tfrac{4}{5}\langle+|-\rangle\right|^2 \\
&= \left|\tfrac{3}{5}\right|^2 = \tfrac{9}{25}.
\end{aligned} \tag{1.43}$$

The probability of measuring $S_z = -\hbar/2$ is

$$\begin{aligned}
\mathcal{P}_- &= \left|\langle-|\psi_{in}\rangle\right|^2 \\
&= \left|\langle-|[\tfrac{3}{5}|+\rangle + \tfrac{4}{5}|-\rangle]\right|^2 \\
&= \left|\tfrac{3}{5}\langle-|+\rangle + \tfrac{4}{5}\langle-|-\rangle\right|^2 \\
&= \left|\tfrac{4}{5}\right|^2 = \tfrac{16}{25}.
\end{aligned} \tag{1.44}$$

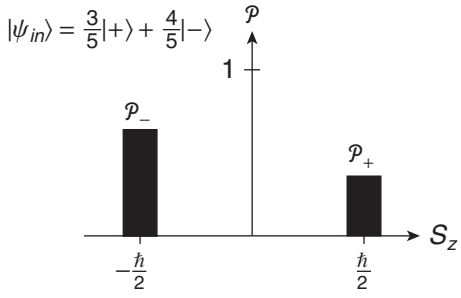


FIGURE 1.13 Histogram of S_z spin component measurements.

Note that the two probabilities add to unity, which indicates that we normalized the input state properly. A histogram of the predicted measurement results is shown in Fig. 1.13.

1.3 ■ MATRIX NOTATION

Up to this point, we have defined kets mathematically in terms of their inner products with other kets. Thus, in the general case we write a ket as

$$|\psi\rangle = \langle +|\psi\rangle|+\rangle + \langle -|\psi\rangle|-\rangle, \tag{1.45}$$

or in a specific case, we write

$$\begin{aligned} |+\rangle_x &= \langle +|+\rangle_x |+\rangle + \langle -|+\rangle_x |-\rangle \\ &= \tfrac{1}{\sqrt{2}}|+\rangle + \tfrac{1}{\sqrt{2}}|-\rangle. \end{aligned} \tag{1.46}$$

In both of these cases, we have chosen to write the kets in terms of the $|+\rangle$ and $|-\rangle$ basis kets. If we agree on that choice of basis as a convention, then the two coefficients $\langle +|+\rangle_x$ and $\langle -|+\rangle_x$ uniquely specify the quantum state, and we can simplify the notation by using just those numbers. Thus, we represent a ket as a **column vector** containing the two coefficients that multiply each basis ket. For example, we represent $|+\rangle_x$ as

$$|+\rangle_x \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{1.47}$$

where we have used the new symbol $\doteq$ to signify "is represented by," and it is understood that we are using the $|+\rangle$ and $|-\rangle$ basis or the S_z basis. We cannot say that the ket *equals* the column vector, because the ket is an abstract vector in the state space and the column vector is just two complex numbers. If we were to choose a different basis for representing the vector, then the complex coefficients would be different even though the vector is unchanged. We need to have a convention for the ordering of the amplitudes in the column vector. The standard convention is to put the spin up amplitude first (at the top). Thus, the representation of the $|-\rangle_x$ state in Eq. (1.36) is

$$|-\rangle_x \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{matrix} \leftarrow |+\rangle \\ \leftarrow |-\rangle \end{matrix}, \tag{1.48}$$

where we have explicitly labeled the rows according to their corresponding basis kets. Using this convention, it should be clear that the basis kets themselves are written as

$$\begin{aligned} |+\rangle &\doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |-\rangle &\doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \tag{1.49}$$

This demonstrates the important feature that *basis kets are unit vectors when written in their own basis*.

This new way of expressing a ket simply as the collection of coefficients that multiply the basis kets is referred to as a **representation**. Because we have assumed the S_z kets as the basis kets, this is called the S_z representation. It is always true that basis kets have the simple form shown in Eq. (1.49) when written in their own representation. A general ket $|\psi\rangle$ is written as

$$|\psi\rangle \doteq \begin{pmatrix} \langle +|\psi\rangle \\ \langle -|\psi\rangle \end{pmatrix}. \tag{1.50}$$

This use of matrix notation simplifies the mathematics of bras and kets. The advantage is not so evident for the simple two-dimensional state space of spin-1/2 systems, but it is very evident for larger dimensional problems. This notation is indispensable when using computers to calculate quantum mechanical results. For example, the SPINS program employs matrix calculations coded in the Java computer language to simulate the Stern-Gerlach experiments using the same probability rules you are learning here.

We saw earlier [Eq. (1.11)] that the coefficients of a bra are the complex conjugates of the coefficients of the corresponding ket. We also know that an inner product of a bra and a ket yields a single complex number. In order for the matrix rules of multiplication to be used, a bra must be represented by a **row vector**, with the entries being the coefficients ordered in the same sense as for the ket. For example, if we use the general ket

$$|\psi\rangle = a|+\rangle + b|-\rangle, \tag{1.51}$$

which is represented as

$$|\psi\rangle \doteq \begin{pmatrix} a \\ b \end{pmatrix}, \tag{1.52}$$

then the corresponding bra

$$\langle\psi| = a^*\langle +| + b^*\langle -| \tag{1.53}$$

is represented by a row vector as

$$\langle\psi| \doteq \begin{pmatrix} a^* & b^* \end{pmatrix}. \tag{1.54}$$

The rules of matrix algebra can then be applied to find an inner product. For example,

$$\begin{aligned} \langle\psi|\psi\rangle &= \begin{pmatrix} a^* & b^* \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \\ &= |a|^2 + |b|^2. \end{aligned} \tag{1.55}$$

So a bra is represented by a row vector that is the complex conjugate and transpose of the column vector representing the corresponding ket.

Example 1.3 To get some practice using this new matrix notation, and to learn some more about the spin-1/2 system, use the results of Experiment 2 to determine the S_y basis kets using the matrix approach instead of the Dirac bra-ket approach.

Consider Experiment 2 in the case where the second Stern-Gerlach analyzer is aligned along the y-axis. We said before that the results are the same as in the case shown in Fig. 1.4. Thus, we have

$$\begin{aligned}
\mathcal{P}_{1,+y} &= \left|{}_y\langle +|+\rangle\right|^2 = \tfrac{1}{2} \\
\mathcal{P}_{1,-y} &= \left|{}_y\langle -|+\rangle\right|^2 = \tfrac{1}{2} \\
\mathcal{P}_{2,+y} &= \left|{}_y\langle +|-\rangle\right|^2 = \tfrac{1}{2} \\
\mathcal{P}_{2,-y} &= \left|{}_y\langle -|-\rangle\right|^2 = \tfrac{1}{2},
\end{aligned} \tag{1.56}$$

as depicted in the histograms of Fig. 1.14.

These results allow us to determine the kets $|\pm\rangle_y$ corresponding to the spin component up and down along the y-axis. The argument and calculation proceeds exactly as it did earlier for the $|\pm\rangle_x$ states up until the point [Eq. (1.35)] where we arbitrarily chose the phase α to be zero. Having done that for the $|\pm\rangle_x$ states, we are no longer free to make that same choice for the $|\pm\rangle_y$ states. Thus we use Eq. (1.35) to write the $|\pm\rangle_y$ states as

$$\begin{aligned}
|+\rangle_y &= \frac{1}{\sqrt{2}}\left[|+\rangle + e^{i\alpha}|-\rangle\right] \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ e^{i\alpha} \end{pmatrix} \\
|-\rangle_y &= \frac{1}{\sqrt{2}}\left[|+\rangle - e^{i\alpha}|-\rangle\right] \doteq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -e^{i\alpha} \end{pmatrix}.
\end{aligned} \tag{1.57}$$

To determine the phase α, we use some more information at our disposal. Experiment 2 could be performed with the first Stern-Gerlach analyzer aligned along the x-axis and the second analyzer along the y-axis. Again the results would be identical (50% at each output port), yielding

$$\mathcal{P}_{+y} = \left|{}_y\langle +|+\rangle_x\right|^2 = \tfrac{1}{2} \tag{1.58}$$

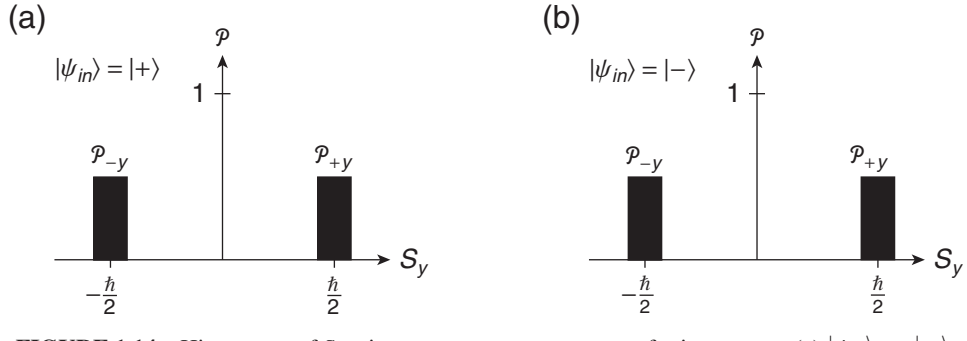


FIGURE 1.14 Histograms of S_y spin component measurements for input states (a) $|\psi_{in}\rangle = |+\rangle$ and (b) $|\psi_{in}\rangle = |-\rangle$.

as one of the measured quantities. Now use matrix algebra to calculate this:

$$\begin{aligned} {}_y\langle + | + \rangle_x &= \tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & e^{-i\alpha}\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \\ &= \tfrac{1}{2}(1 + e^{-i\alpha}) \\ \left|{}_y\langle + | + \rangle_x\right|^2 &= \tfrac{1}{2}(1 + e^{-i\alpha})\tfrac{1}{2}(1 + e^{i\alpha}) \\ &= \tfrac{1}{4}(1 + e^{i\alpha} + e^{-i\alpha} + 1) \\ &= \tfrac{1}{2}(1 + \cos\alpha) = \tfrac{1}{2}. \end{aligned} \tag{1.59}$$

This result requires that $\cos\alpha = 0$, or that $\alpha = \pm\pi/2$. The two choices for the phase correspond to the two possibilities for the direction of the y-axis relative to the already determined x- and z-axes. The choice $\alpha = +\pi/2$ can be shown to correspond to a right-handed coordinate system, which is the standard convention, so we choose that phase. We thus represent the $|\pm\rangle_y$ kets as

$$\begin{aligned} |+\rangle_y &\doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix} \\ |-\rangle_y &\doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}. \end{aligned} \tag{1.60}$$

Note that the imaginary components of these kets are required. They are not merely a mathematical convenience as one sees in classical mechanics. In general, quantum mechanical state vectors have complex coefficients. But this does not mean that the results of physical measurements are complex. On the contrary, we always calculate a measurement probability using a complex square, so all quantum mechanics predictions of probabilities are real.

1.4 ■ GENERAL QUANTUM SYSTEMS

The machinery we have developed for spin-1/2 systems can be generalized to other quantum systems. For example, if an observable A yields quantized measurement results a_n for some finite range of n, then we generalize the schematic depiction of a Stern-Gerlach measurement to a measurement of the

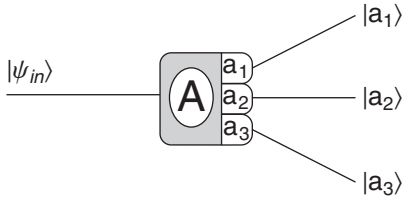


FIGURE 1.15 Generic depiction of the quantum mechanical measurement of observable A.

observable A, as shown in Fig. 1.15. The observable A labels the measurement device and the possible results a_1, a_2, a_3, etc. label the output ports. The basis kets corresponding to the results a_n are then $|a_n\rangle$. The mathematical rules about kets in this general case are

$$\begin{aligned} \langle a_i|a_j\rangle &= \delta_{ij} && \textit{orthonormality} \\ |\psi\rangle &= \sum_i \langle a_i|\psi\rangle |a_i\rangle && \textit{completeness}, \end{aligned} \tag{1.61}$$

where we use the **Kronecker delta**

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{1.62}$$

to express the orthonormality condition compactly. In this case, the generalization of postulate 4 says that the probability of a measurement of one of the possible results a_n is

$$\mathcal{P}_{a_n} = |\langle a_n|\psi_{in}\rangle|^2. \tag{1.63}$$

Example 1.4 Imagine a quantum system with an observable A that has three possible measurement results: a_1, a_2, and a_3. The three kets $|a_1\rangle$, $|a_2\rangle$, and $|a_3\rangle$ corresponding to these possible results form a complete orthonormal basis. The system is prepared in the state

$$|\psi\rangle = 2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle. \tag{1.64}$$

Calculate the probabilities of all possible measurement results of the observable A.

The state vector in Eq. (1.64) is not normalized, so we must normalize it before calculating probabilities. Introducing a complex normalization constant C, we find

$$\begin{aligned} 1 &= \langle\psi|\psi\rangle \\ &= C^*(2\langle a_1| - 3\langle a_2| - 4i\langle a_3|)C(2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle) \\ &= |C|^2\{4\langle a_1|a_1\rangle - 6\langle a_1|a_2\rangle + 8i\langle a_1|a_3\rangle \\ &\qquad - 6\langle a_2|a_1\rangle + 9\langle a_2|a_2\rangle - 12i\langle a_2|a_3\rangle \\ &\qquad - 8i\langle a_3|a_1\rangle + 12i\langle a_3|a_2\rangle + 16\langle a_3|a_3\rangle\} \\ &= |C|^2\{4 + 9 + 16\} = |C|^2 29 \\ &\Rightarrow C = \tfrac{1}{\sqrt{29}}. \end{aligned} \tag{1.65}$$

The normalized state is

$$|\psi\rangle = \tfrac{1}{\sqrt{29}}(2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle). \tag{1.66}$$

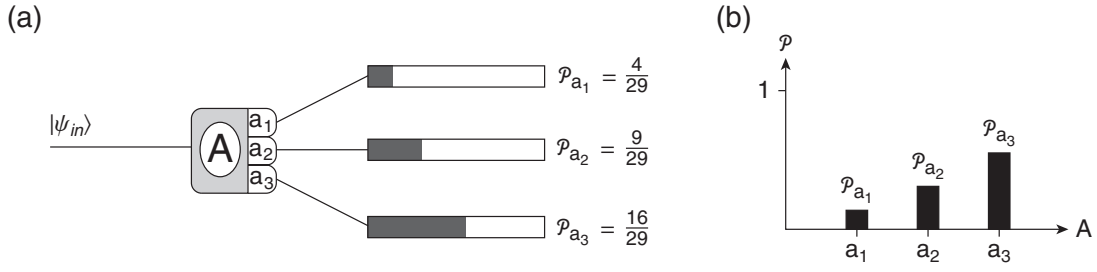


FIGURE 1.16 (a) Schematic diagram of the measurement of observable A and (b) histogram of the predicted measurement probabilities.

The probabilities of measuring the results a_1, a_2, and a_3 are

$$\begin{aligned}
\mathcal{P}_{a_1} &= |\langle a_1|\psi\rangle|^2 \\
&= \left|\langle a_1|\tfrac{1}{\sqrt{29}}\{2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle\}\right|^2 \\
&= \tfrac{1}{29}|2\langle a_1|a_1\rangle - 3\langle a_1|a_2\rangle + 4i\langle a_1|a_3\rangle|^2 = \tfrac{4}{29} \qquad (1.67)\\
\mathcal{P}_{a_2} &= |a_2|\psi\rangle|^2 = \left|\langle a_2|\tfrac{1}{\sqrt{29}}\{2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle\}\right|^2 = \tfrac{9}{29} \\
\mathcal{P}_{a_3} &= |\langle a_3|\psi\rangle|^2 = \left|\langle a_3|\tfrac{1}{\sqrt{29}}\{2|a_1\rangle - 3|a_2\rangle + 4i|a_3\rangle\}\right|^2 = \tfrac{16}{29}.
\end{aligned}$$

A schematic of this experiment is shown in Fig. 1.16(a) and a histogram of the predicted probabilities is shown in Fig. 1.16(b).

1.5 ■ POSTULATES

We have introduced two of the postulates of quantum mechanics in this chapter. The postulates of quantum mechanics dictate how to treat a quantum mechanical system mathematically and how to interpret the mathematics to learn about the physical system in question. These postulates cannot be proven, but they have been successfully tested by many experiments, and so we accept them as an accurate way to describe quantum mechanical systems. New results could force us to reevaluate these postulates at some later time. All six postulates are listed below to give you an idea where we are headed and a framework into which you can place the new concepts as we confront them.

Postulates of Quantum Mechanics

1. The state of a quantum mechanical system, including all the information you can know about it, is represented mathematically by a normalized ket $|\psi\rangle$.
2. A physical observable is represented mathematically by an operator A that acts on kets.
3. The only possible result of a measurement of an observable is one of the eigenvalues a_n of the corresponding operator A.

4. The probability of obtaining the eigenvalue a_n in a measurement of the observable A on the system in the state $|\psi\rangle$ is

$$\mathcal{P}_{a_n} = |\langle a_n|\psi\rangle|^2,$$

where $|a_n\rangle$ is the normalized eigenvector of A corresponding to the eigenvalue a_n.

5. After a measurement of A that yields the result a_n, the quantum system is in a new state that is the normalized projection of the original system ket onto the ket (or kets) corresponding to the result of the measurement:

$$|\psi'\rangle = \frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}}.$$

6. The time evolution of a quantum system is determined by the Hamiltonian or total energy operator $H(t)$ through the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle.$$

As you read these postulates for the first time, you will undoubtedly encounter new terms and concepts. Rather than explain them all here, the plan of this text is to continue to explain them through their manifestation in the Stern-Gerlach spin-1/2 experiment. We have chosen this example because it is inherently quantum mechanical and forces us to break away from reliance on classical intuition or concepts. Moreover, this simple example is a paradigm for many other quantum mechanical systems. By studying it in detail, we can appreciate much of the richness of quantum mechanics.

SUMMARY

Through the Stern-Gerlach experiment we have learned several key concepts about quantum mechanics in this chapter.

- Quantum mechanics is probabilistic.
  We cannot predict the results of experiments precisely. We can predict only the probability that a certain result is obtained in a measurement.
- Spin measurements are quantized.
  The possible results of a spin component measurement are quantized. Only these discrete values are measured.
- Quantum measurements disturb the system.
  Measuring one physical observable can "destroy" information about other observables.

We have learned how to describe the state of a quantum mechanical system mathematically using a ket, which represents all the information we can know about that state. The kets $|+\rangle$ and $|-\rangle$ result when the spin component S_z along the z-axis is measured to be up or down, respectively. These kets form an orthonormal basis, which we denote by the inner products

$$\begin{aligned}\langle+|+\rangle &= 1\\ \langle-|-\rangle &= 1\\ \langle+|-\rangle &= 0.\end{aligned} \tag{1.68}$$

The basis is also complete, which means that it can be used to express all possible kets as superposition states

$$|\psi\rangle = a|+\rangle + b|-\rangle. \tag{1.69}$$

For spin component measurements, the kets corresponding to spin up or down along the three Cartesian axes are

$$\begin{array}{lll} |+\rangle & |+\rangle_x = \frac{1}{\sqrt{2}}\big[|+\rangle + |-\rangle\big] & |+\rangle_y = \frac{1}{\sqrt{2}}\big[|+\rangle + i|-\rangle\big] \\ |-\rangle & |-\rangle_x = \frac{1}{\sqrt{2}}\big[|+\rangle - |-\rangle\big] & |-\rangle_y = \frac{1}{\sqrt{2}}\big[|+\rangle - i|-\rangle\big]. \end{array} \tag{1.70}$$

We also found it useful to introduce a matrix notation for calculations. In this matrix language the kets in Eq. (1.70) are represented by

$$\begin{array}{lll} |+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} & |+\rangle_x \doteq \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} & |+\rangle_y \doteq \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} \\ |-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} & |-\rangle_x \doteq \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} & |-\rangle_y \doteq \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}. \end{array} \tag{1.71}$$

The most important tool we have learned so far is the probability postulate (postulate 4). To calculate the probability that a measurement on an input state $|\psi_{in}\rangle$ will yield a particular result, for example $S_z = \hbar/2$, we complex square the inner product of the input state with the ket corresponding to the measured result, $|+\rangle$ in this case:

$$\mathcal{P}_+ = |\langle + |\psi_{in}\rangle|^2. \tag{1.72}$$

This is generalized to other systems where a measurement yields a particular result a_n corresponding to the ket $|a_n\rangle$ as:

$$\mathcal{P}_{a_n} = |\langle a_n|\psi_{in}\rangle|^2. \tag{1.73}$$

PROBLEMS

1.1 Consider the following state vectors:

$$\begin{aligned} |\psi_1\rangle &= 3|+\rangle + 4|-\rangle \\ |\psi_2\rangle &= |+\rangle + 2i|-\rangle \\ |\psi_3\rangle &= 3|+\rangle - e^{i\pi/3}|-\rangle. \end{aligned}$$

a) Normalize each state vector.

b) For each state vector, calculate the probability that the spin component is up or down along each of the three Cartesian axes. Use bra-ket notation for the entire calculation.

c) Write each normalized state in matrix notation.

d) Repeat part (b) using matrix notation for the entire calculation.

1.2 Consider the three quantum states:

$$|\psi_1\rangle = \tfrac{1}{\sqrt{3}}|+\rangle + i\tfrac{\sqrt{2}}{\sqrt{3}}|-\rangle$$
$$|\psi_2\rangle = \tfrac{1}{\sqrt{5}}|+\rangle - \tfrac{2}{\sqrt{5}}|-\rangle$$
$$|\psi_3\rangle = \tfrac{1}{\sqrt{2}}|+\rangle + e^{i\pi/4}\tfrac{1}{\sqrt{2}}|-\rangle.$$

Use bra-ket notation (not matrix notation) to solve the following problems. Note that $\langle+|+\rangle = 1$, $\langle-|-\rangle = 1$, and $\langle+|-\rangle = 0$.

a) For each of the $|\psi_i\rangle$ above, find the normalized vector $|\phi_i\rangle$ that is orthogonal to it.

b) Calculate the inner products $\langle\psi_i|\psi_j\rangle$ for i and $j = 1, 2, 3$.

1.3 Show that a change in the overall phase of a quantum state vector does not change the probability of obtaining a particular result in a measurement. To do this, consider how the probability is affected by changing the state $|\psi\rangle$ to the state $e^{i\delta}|\psi\rangle$.

1.4 Show by explicit bra-ket calculations using the states in Eq. (1.29) that the four experimental results in Eq. (1.28) lead to the results $|b|^2 = |c|^2 = |d|^2 = \frac{1}{2}$.

1.5 A beam of spin-1/2 particles is prepared in the state

$$|\psi\rangle = \tfrac{2}{\sqrt{13}}|+\rangle + i\tfrac{3}{\sqrt{13}}|-\rangle.$$

a) What are the possible results of a measurement of the spin component S_z, and with what probabilities would they occur?

b) What are the possible results of a measurement of the spin component S_x, and with what probabilities would they occur?

c) Plot histograms of the predicted measurement results from parts (a) and (b).

1.6 A beam of spin-1/2 particles is prepared in the state

$$|\psi\rangle = \tfrac{2}{\sqrt{13}}|+\rangle_x + i\tfrac{3}{\sqrt{13}}|-\rangle_x.$$

a) What are the possible results of a measurement of the spin component S_z, and with what probabilities would they occur?

b) What are the possible results of a measurement of the spin component S_x, and with what probabilities would they occur?

c) Plot histograms of the predicted measurement results from parts (a) and (b).

1.7 A classical coin is thrown in the air and lands on the ground, where a measurement is made of its state.

a) What are the possible results of this measurement?

b) What are the predicted probabilities for these possible outcomes?

c) Plot a histogram of the predicted measurement results.

1.8 A classical cubical die is thrown onto a table and comes to rest, where a measurement is made of its state.

a) What are the possible results of this measurement?

b) What are the predicted probabilities for these possible outcomes?

c) Plot a histogram of the predicted measurement results.

1.9 A pair of dice (classical cubes) are thrown onto a table and come to rest, where a measurement is made of the state of the system (i.e., the sum of the two dice).

a) What are the possible results of this measurement?

b) What are the predicted probabilities for these possible outcomes?

c) Plot a histogram of the predicted measurement results.

1.10 Consider the three quantum states:

$$|\psi_1\rangle = \tfrac{4}{5}|+\rangle + i\tfrac{3}{5}|-\rangle$$
$$|\psi_2\rangle = \tfrac{4}{5}|+\rangle - i\tfrac{3}{5}|-\rangle$$
$$|\psi_3\rangle = -\tfrac{4}{5}|+\rangle + i\tfrac{3}{5}|-\rangle.$$

a) For each of the $|\psi_i\rangle$ above, calculate the probabilities of spin component measurements along the x-, y-, and z-axes.

b) Use your results from (a) to comment on the importance of the overall phase and of the relative phases of the quantum state vector.

1.11 A beam of spin-1/2 particles is prepared in the state

$$|\psi\rangle = \tfrac{3}{\sqrt{34}}|+\rangle + i\tfrac{5}{\sqrt{34}}|-\rangle.$$

a) What are the possible results of a measurement of the spin component S_z, and with what probabilities would they occur?

b) Suppose that the S_z measurement yields the result $S_z = -\hbar/2$. Subsequent to that result a second measurement is performed to measure the spin component S_x. What are the possible results of that measurement, and with what probabilities would they occur?

c) Draw a schematic diagram depicting the successive measurements in parts (a) and (b).

1.12 Consider a quantum system with an observable A that has three possible measurement results: a_1, a_2, and a_3. Write down the orthogonality, normalization, and completeness relations for the three kets comprising the basis corresponding to the possible results of the A measurement.

1.13 Consider a quantum system with an observable A that has three possible measurement results: a_1, a_2, and a_3.

a) Write down the three kets $|a_1\rangle$, $|a_2\rangle$, and $|a_3\rangle$ corresponding to these possible results using matrix notation.

b) The system is prepared in the state

$$|\psi\rangle = 1|a_1\rangle - 2|a_2\rangle + 5|a_3\rangle.$$

Write this state in matrix notation and calculate the probabilities of all possible measurement results of the observable A. Plot a histogram of the predicted measurement results.

c) In a different experiment, the system is prepared in the state

$$|\psi\rangle = 2|a_1\rangle + 3i|a_2\rangle.$$

Write this state in matrix notation and calculate the probabilities of all possible measurement results of the observable A. Plot a histogram of the predicted measurement results.

1.14 Consider a quantum system in which the energy E is measured and there are four possible measurement results: 2 eV, 4 eV, 7 eV, and 9 eV. The system is prepared in the state

$$|\psi\rangle = \tfrac{1}{\sqrt{39}}\{3|2\text{ eV}\rangle - i|4\text{ eV}\rangle + 2e^{i\pi/7}|7\text{ eV}\rangle + 5|9\text{ eV}\rangle\}.$$

Calculate the probabilities of all possible measurement results of the energy E. Plot a histogram of the predicted measurement results.

1.15 Consider a quantum system described by a basis $|a_1\rangle$, $|a_2\rangle$, and $|a_3\rangle$. The system is initially in a state

$$|\psi_i\rangle = \tfrac{i}{\sqrt{3}}|a_1\rangle + \sqrt{\tfrac{2}{3}}\,|a_2\rangle.$$

Find the probability that the system is measured to be in the final state

$$|\psi_f\rangle = \tfrac{1+i}{\sqrt{3}}|a_1\rangle + \tfrac{1}{\sqrt{6}}|a_2\rangle + \tfrac{1}{\sqrt{6}}|a_3\rangle.$$

1.16 The spin components of a beam of atoms prepared in the state $|\psi_{in}\rangle$ are measured and the following experimental probabilities are obtained:

$$\mathcal{P}_+ = \tfrac{1}{2} \qquad \mathcal{P}_{+x} = \tfrac{3}{4} \qquad \mathcal{P}_{+y} = 0.067$$
$$\mathcal{P}_- = \tfrac{1}{2} \qquad \mathcal{P}_{-x} = \tfrac{1}{4} \qquad \mathcal{P}_{-y} = 0.933.$$

From the experimental data, determine the input state.

1.17 In part (1) of SPINS Lab #2, you measured the probabilities of all the possible spin components for each of the unknown initial states $|\psi_i\rangle$ ($i = 1, 2, 3, 4$). Using your data from that lab, find the unknown states $|\psi_1\rangle$, $|\psi_2\rangle$, $|\psi_3\rangle$, and $|\psi_4\rangle$. Express each of the unknown states as a linear superposition of the S_z basis states $|+\rangle$ and $|-\rangle$. For each state, use your result to calculate the theoretical values of the probabilities for each component measurement and compare these theoretical predictions with your experimental results.

RESOURCES

Activities

SPINS: A software program to simulate Stern-Gerlach spin experiments. The Java software runs on all platforms and can be downloaded in two forms:

Open Source Physics framework

www.physics.oregonstate.edu/~mcintyre/ph425/spins/index_SPINS_OSP.html

or

Standalone Java

www.physics.oregonstate.edu/~mcintyre/ph425/spins

The bulleted activities are available at

www.physics.oregonstate.edu/qmactivities

- **SPINS Lab 1:** An introduction to successive Stern-Gerlach spin-1/2 measurements. The randomness of measurements is demonstrated and students use statistical analysis to deduce probabilities from measurements.

- **SPINS Lab 2:** Students deduce unknown quantum state vectors from measurements of spin projections (part 3 requires material from Chapter 2 to do the calculations).

Stern-Gerlach simulation: A different simulation of the Stern-Gerlach experiment from the PHET group at the University of Colorado (somewhat Flashier version):

http://phet.colorado.edu/en/simulation/stern-gerlach

Further Reading

The history of the Stern-Gerlach experiment and how a bad cigar helped are chronicled in a *Physics Today* article:

B. Friedrich and D. Herschbach, "Stern and Gerlach: How a Bad Cigar Helped Reorient Atomic Physics," *Phys. Today* **56**(12), 53–59 (2003).

http://dx.doi.org/10.1063/1.1650229

A different spin on the quantum mechanics of socks is discussed by John S. Bell in this article:

J. S. Bell, "Bertlmann's socks and the nature of reality, " *J. Phys. Colloq.* **42**, C22 C2.41-C2.62 (1981).

http://cdsweb.cern.ch/record/142461

Nature has published a supplement on the milestones in spin physics. An extensive timeline of historical events, review articles, and links to original articles are included.

Nature Phys. 4, S1–S43 (2008).

www.nature.com/milestones/spin

The SPINS lab software is described in this pedagogical article:

D. V. Schroeder and T. A. Moore, "A computer-simulated Stern-Gerlach laboratory," *Am. J. Phys.* **61**, 798–805 (1993).

http://dx.doi.org/10.1119/1.17172

Some other textbooks that take a spins-first approach or have an extensive treatment of Stern-Gerlach experiments:

R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics,* Volume 3, *Quantum Mechanics*, Reading, MA: Addison-Wesley Publishing Company, Inc., 1965.

J. J. Sakurai, *Modern Quantum Mechanics*, Redwood City, CA: Addison-Wesley Publishing Company, Inc., 1985.

J. S. Townsend, *A Modern Approach to Quantum Mechanics*, New York: McGraw Hill, Inc., 1992.

C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*, New York: John Wiley & Sons, 1977.

D. F. Styer, *The Strange World of Quantum Mechanics*, Cambridge: Cambridge University Press, 2000.

CHAPTER

2 Operators and Measurement

In Chapter 1 we used the results of experiments to deduce a mathematical description of the spin-1/2 system. The Stern-Gerlach experiments demonstrated that spin component measurements along the x-, y-, or z-axes yield only $\pm\hbar/2$ as possible results. We learned how to predict the probabilities of these measurements using the basis kets of the spin component observables S_x, S_y, and S_z, and these predictions agreed with the experiments. However, the real power of a theory is its ability to predict results of experiments that you haven't yet done. For example, what are the possible results of a measurement of the spin component S_n along an arbitrary direction $\hat{\mathbf{n}}$ and what are the predicted probabilities? To make these predictions, we need to learn about the operators of quantum mechanics.

2.1 ■ OPERATORS, EIGENVALUES, AND EIGENVECTORS

The mathematical theory we developed in Chapter 1 used only quantum state vectors. We said that the state vector represents all the information we can know about the system and we used the state vectors to calculate probabilities. With each observable S_x, S_y, and S_z we associated a pair of kets corresponding to the possible measurement results of that observable. The observables themselves are not yet included in our mathematical theory, but the distinct association between an observable and its measurable kets provides the means to do so.

The role of physical observables in the mathematics of quantum theory is described by the two postulates listed below. Postulate 2 states that physical observables are represented by mathematical operators, in the same sense that physical states are represented by mathematical vectors or kets (postulate 1). An **operator** is a mathematical object that acts or operates on a ket and transforms it into a new ket, for example $A|\psi\rangle = |\phi\rangle$. However, there are special kets that are not changed by the operation of a particular operator, except for a possible multiplicative constant, which we know does not change anything measurable about the state. An example of a ket that is not changed by an operator would be $A|\psi\rangle = a|\psi\rangle$. Such kets are known as **eigenvectors** of the operator A and the multiplicative constants are known as the **eigenvalues** of the operator. These are important because postulate 3 states that the only possible result of a measurement of a physical observable is one of the eigenvalues of the corresponding operator.

Postulate 2

A physical observable is represented mathematically by an operator A that acts on kets.

Postulate 3

The only possible result of a measurement of an observable is one of the eigenvalues a_n of the corresponding operator A.

We now have a mathematical description of that special relationship we saw in Chapter 1 between a physical observable, S_z say, the possible results $\pm\hbar/2$, and the kets $|\pm\rangle$ corresponding to those results. This relationship is known as the **eigenvalue equation** and is depicted in Fig. 2.1 for the case of the spin up state in the z-direction. In the eigenvalue equation, the observable is represented by an operator, the eigenvalue is one of the possible measurement results of the observable, and the eigenvector is the ket corresponding to the chosen eigenvalue of the operator. The eigenvector appears on both sides of the equation because it is unchanged by the operator.

The eigenvalue equations for the S_z operator in a spin-1/2 system are:

$$\begin{aligned} S_z|+\rangle &= +\frac{\hbar}{2}|+\rangle \\ S_z|-\rangle &= -\frac{\hbar}{2}|-\rangle. \end{aligned} \tag{2.1}$$

These equations tell us that $+\hbar/2$ is the eigenvalue of S_z corresponding to the eigenvector $|+\rangle$ and $-\hbar/2$ is the eigenvalue of S_z corresponding to the eigenvector $|-\rangle$. Equations (2.1) are sufficient to define how the S_z operator acts mathematically on kets. However, it is useful to use matrix notation to represent operators in the same sense that we used column vectors and row vectors in Chapter 1 to represent bras and kets, respectively. For Eqs. (2.1) to be satisfied using matrix algebra with the kets represented as column vectors of size 1×2, the operator S_z must be represented by a 2×2 matrix. The eigenvalue equations (2.1) provide sufficient information to determine this matrix.

To determine the matrix representing the operator S_z, assume the most general form for a 2×2 matrix

$$S_z \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{2.2}$$

where we are again using the $\doteq$ symbol to mean "is represented by." Now write the eigenvalue equations in matrix form:

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} &= +\frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} &= -\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \tag{2.3}$$

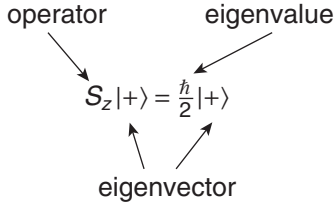


FIGURE 2.1 Eigenvalue equation for the spin up state.

Note that we are still using the convention that the $|\pm\rangle$ kets are used as the basis for the representation. It is crucial that the rows and columns of the operator matrix are ordered in the same manner as used for the ket column vectors; anything else would amount to nonsense. An explicit labeling of the rows and columns of the operator and the basis kets makes this clear:

$$\begin{array}{c|cc} S_z & |+\rangle & |-\rangle \\ \hline \langle +| & a & b \\ \langle -| & c & d \end{array} \qquad \begin{array}{c|c} |+\rangle & \\ \hline \langle +| & 1 \\ \langle -| & 0 \end{array} \qquad \begin{array}{c|c} |-\rangle & \\ \hline \langle +| & 0 \\ \langle -| & 1 \end{array} . \tag{2.4}$$

Carrying through the multiplication in Eqs. (2.3) yields

$$\begin{aligned} \begin{pmatrix} a \\ c \end{pmatrix} &= +\frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} b \\ d \end{pmatrix} &= -\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \end{aligned} \tag{2.5}$$

which results in

$$\begin{aligned} a &= +\frac{\hbar}{2} & b &= 0 \\ c &= 0 & d &= -\frac{\hbar}{2}. \end{aligned} \tag{2.6}$$

Thus the matrix representation of the operator S_z is

$$\begin{aligned} S_z &\doteq \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \\ &\doteq \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{2.7}$$

Note two important features of this matrix: (1) it is a **diagonal matrix**—it has only diagonal elements—and (2) the diagonal elements are the eigenvalues of the operator, ordered in the same manner as the corresponding eigenvectors. In this example, the basis used for the matrix representation is that formed by the eigenvectors $|\pm\rangle$ of the operator S_z. That the matrix representation of the operator in this case is a diagonal matrix is a necessary and general result of linear algebra that will prove valuable as we study quantum mechanics. In simple terms, we say that *an operator is always diagonal in its own basis*. This special form of the matrix representing the operator is similar to the special form that the eigenvectors $|\pm\rangle$ take in this same representation—the *eigenvectors are unit vectors in their own basis*. These ideas cannot be overemphasized, so we repeat them:

An operator is always diagonal in its own basis.
Eigenvectors are unit vectors in their own basis.

Let's also summarize the matrix representations of the S_z operator and its eigenvectors:

$$S_z \doteq \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad |+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.8}$$

2.1.1 ■ Matrix Representation of Operators

Now consider how matrix representation works in general. Consider a general operator A describing a physical observable (still in the two-dimensional spin-1/2 system), which we represent by the general matrix

$$A \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{2.9}$$

in the S_z basis. The operation of A on the basis ket $|+\rangle$ yields

$$A|+\rangle \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}. \tag{2.10}$$

The inner product of this new ket $A|+\rangle$ with the ket $|+\rangle$ (converted to a bra following the rules) results in

$$\langle +|A|+\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ c \end{pmatrix} = a, \tag{2.11}$$

which serves to isolate one of the elements of the matrix. Hence an individual element such as $\langle +|A|+\rangle$ or $\langle +|A|-\rangle$ is generally referred to as a **matrix element**. This "sandwich" of a bra, an operator, and a ket

$$\langle bra|OPERATOR|ket\rangle \tag{2.12}$$

plays an important role in many quantum mechanical calculations. Even in cases where the bra and ket are not basis kets, such as in $\langle \psi|A|\phi\rangle$, we still refer to this as a matrix element. A schematic diagram of a generic matrix element is depicted in Fig. 2.2(a).

All four elements of the matrix representation of A can be determined in the same manner as Eq. (2.11), with the final result

$$A \doteq \begin{pmatrix} \langle +|A|+\rangle & \langle +|A|-\rangle \\ \langle -|A|+\rangle & \langle -|A|-\rangle \end{pmatrix}. \tag{2.13}$$

To emphasize the structure of the matrix, let's write it with explicit labeling of the rows and columns:

A	$\lvert+\rangle$	$\lvert-\rangle$
$\langle +\rvert$	$\langle +\lvert A\rvert+\rangle$	$\langle +\lvert A\rvert-\rangle$
$\langle -\rvert$	$\langle -\lvert A\rvert+\rangle$	$\langle -\lvert A\rvert-\rangle$

. (2.14)

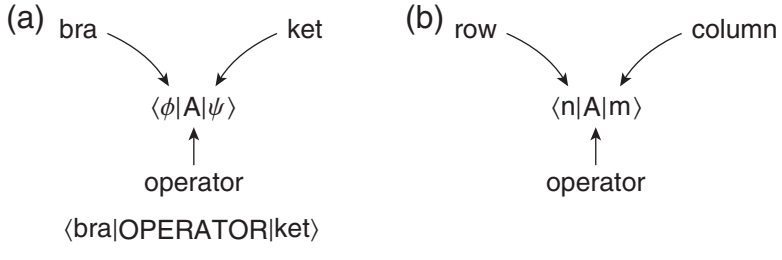


FIGURE 2.2 (a) Schematic diagram of a generic matrix element. (b) Schematic diagram of the row and column labeling convention for matrix elements.

In a more general problem with more than two dimensions in the complex vector space, the matrix representation of an operator is

$$A \doteq \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.15}$$

where the matrix elements are

$$A_{ij} = \langle i|A|j\rangle \tag{2.16}$$

and the basis is assumed to be the states labeled $|i\rangle$, with the subscripts i and j labeling the rows and columns respectively, as depicted in Fig. 2.2(b). Using this matrix representation, the action of this operator on a general ket $|\psi\rangle = \sum_i c_i|i\rangle$ is

$$A|\psi\rangle \doteq \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11}c_1 + A_{12}c_2 + A_{13}c_3 + \cdots \\ A_{21}c_1 + A_{22}c_2 + A_{23}c_3 + \cdots \\ A_{31}c_1 + A_{32}c_2 + A_{33}c_3 + \cdots \\ \vdots \end{pmatrix}. \tag{2.17}$$

If we write the new ket $|\phi\rangle = A|\psi\rangle$ as $|\phi\rangle = \sum_i b_i|i\rangle$, then from Eq. (2.17) the coefficients b_i are

$$b_i = \sum_j A_{ij}c_j \tag{2.18}$$

in summation notation.

2.1.2 ■ Diagonalization of Operators

In the case of the operator S_z above, we used the experimental results and the eigenvalue equations to find the matrix representation of the operator in Eq. (2.7). It is more common to work the other way. That is, one is given the matrix representation of an operator and is asked to find the possible results of a measurement of the corresponding observable. According to the third postulate, the possible results are the eigenvalues of the operator, and the eigenvectors are the quantum states representing them. In the case of a general operator A in a two-state system, the eigenvalue equation is

$$A|a_n\rangle = a_n|a_n\rangle, \tag{2.19}$$

where we have labeled the eigenvalues a_n and we have labeled the eigenvectors with the corresponding eigenvalues. In matrix notation, the eigenvalue equation is

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} c_{n1} \\ c_{n2} \end{pmatrix} = a_n\begin{pmatrix} c_{n1} \\ c_{n2} \end{pmatrix}, \tag{2.20}$$

where c_{n1} and c_{n2} are the unknown coefficients of the eigenvector $|a_n\rangle$ corresponding to the eigenvalue a_n. This matrix equation yields the set of homogeneous equations

$$\begin{aligned} (A_{11} - a_n)c_{n1} + A_{12}c_{n2} &= 0 \\ A_{21}c_{n1} + (A_{22} - a_n)c_{n2} &= 0. \end{aligned} \tag{2.21}$$

The rules of linear algebra dictate that a set of homogeneous equations has solutions for the unknowns c_{n1} and c_{n2} only if the determinant of the coefficients vanishes:

$$\begin{vmatrix} A_{11} - a_n & A_{12} \\ A_{21} & A_{22} - a_n \end{vmatrix} = 0. \tag{2.22}$$

It is common notation to use the symbol λ for the eigenvalues, in which case this equation is

$$\det(A - \lambda I) = 0, \tag{2.23}$$

where I is the identity matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.24}$$

Equation (2.23) is known as the **secular** or **characteristic** equation. It is a second order equation in the parameter λ and the two roots are identified as the two eigenvalues a_1 and a_2 that we are trying to find. Once these eigenvalues are found, they are then individually substituted back into Eqs. (2.21), which are solved to find the coefficients of the corresponding eigenvector.

Example 2.1 Assume that we know (e.g., from Problem 2.1) that the matrix representation for the operator S_y is

$$S_y \doteq \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{2.25}$$

Find the eigenvalues and eigenvectors of the operator S_y.

The general eigenvalue equation is

$$S_y|\lambda\rangle = \lambda|\lambda\rangle, \tag{2.26}$$

and the possible eigenvalues λ are found using the secular equation

$$\det|S_y - \lambda I| = 0. \tag{2.27}$$

The secular equation is

$$\begin{vmatrix} -\lambda & -i\dfrac{\hbar}{2} \\ i\dfrac{\hbar}{2} & -\lambda \end{vmatrix} = 0, \tag{2.28}$$

and solving yields the eigenvalues

$$\begin{aligned} \lambda^2 + i^2\left(\frac{\hbar}{2}\right)^2 &= 0 \\ \lambda^2 - \left(\frac{\hbar}{2}\right)^2 &= 0 \\ \lambda^2 &= \left(\frac{\hbar}{2}\right)^2 \\ \lambda &= \pm\frac{\hbar}{2}, \end{aligned} \tag{2.29}$$

which was to be expected, because we know that the only possible results of a measurement of any spin component are $\pm\hbar/2$.

As before, we label the eigenvectors $|\pm\rangle_y$. The eigenvalue equation for the positive eigenvalue is

$$S_y|+\rangle_y = +\frac{\hbar}{2}|+\rangle_y, \tag{2.30}$$

or in matrix notation

$$\frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = +\frac{\hbar}{2}\begin{pmatrix}a\\ b\end{pmatrix}, \tag{2.31}$$

where we must solve for a and b to determine the eigenvector. Multiplying through and canceling the common factor yields

$$\begin{pmatrix}-ib\\ ia\end{pmatrix} = \begin{pmatrix}a\\ b\end{pmatrix}. \tag{2.32}$$

This results in two equations, but they are not linearly independent, so we need some more information. The normalization condition provides what we need. Thus we have two equations that determine the eigenvector coefficients:

$$\begin{aligned} b &= ia\\ |a|^2 + |b|^2 &= 1. \end{aligned} \tag{2.33}$$

Solving these yields

$$\begin{aligned} |a|^2 + |ia|^2 &= 1\\ |a|^2 &= \tfrac{1}{2}. \end{aligned} \tag{2.34}$$

Again we follow the convention of choosing the first coefficient to be real and positive, resulting in

$$\begin{aligned} a &= \tfrac{1}{\sqrt{2}}\\ b &= i\tfrac{1}{\sqrt{2}}. \end{aligned} \tag{2.35}$$

Thus the eigenvector corresponding to the positive eigenvalue is

$$|+\rangle_y \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}. \tag{2.36}$$

Likewise, one can find the eigenvector for the negative eigenvalue to be

$$|-\rangle_y \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix}. \tag{2.37}$$

These are, of course, the same states we found in Chapter 1 (Eq. 1.60).

This procedure of finding the eigenvalues and eigenvectors of a matrix is known as **diagonalization** of the matrix and is the key step in many quantum mechanics problems. Generally, if we find a new operator, the first thing we do is diagonalize it to find its eigenvalues and eigenvectors. However, we stop short of the mathematical exercise of finding the matrix that transforms the original matrix to its new diagonal form. This would amount to a change of basis from the original basis to a new basis of the eigenvectors we have just found, much like a rotation in three dimensions changes from one coordinate system to another. We don't want to make this change of basis. In the example above, the S_y matrix is not diagonal, whereas the S_z matrix is diagonal, because we are using the S_z basis. It is

common practice to use the S_z basis as the default basis, so you can assume that is the case unless you are told otherwise.

In summary, we now know three operators and their eigenvalues and eigenvectors. The spin component operators S_x, S_y, and S_z all have eigenvalues $\pm\hbar/2$. The matrix representations of the operators and eigenvectors are (see Problem 2.1)

$$\boxed{\begin{array}{lll} S_x \doteq \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} & |+\rangle_x \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} & |-\rangle_x \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix} \\ S_y \doteq \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} & |+\rangle_y \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix} & |-\rangle_y \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix} \\ S_z \doteq \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} & |+\rangle \doteq \begin{pmatrix}1\\ 0\end{pmatrix} & |-\rangle \doteq \begin{pmatrix}0\\ 1\end{pmatrix} \end{array}}. \tag{2.38}$$

2.2 ■ NEW OPERATORS

2.2.1 ■ Spin Component in a General Direction

Now that we know the three operators corresponding to the spin components along the three Cartesian axes, we can use them to find the operator S_n for the spin component along a general direction $\hat{\mathbf{n}}$. This new operator will allow us to predict results of experiments we have not yet performed. The direction $\hat{\mathbf{n}}$ is specified by the polar and azimuthal angles θ and ϕ as shown in Fig. 2.3. The unit vector $\hat{\mathbf{n}}$ is

$$\hat{\mathbf{n}} = \hat{\mathbf{i}}\sin\theta\cos\phi + \hat{\mathbf{j}}\sin\theta\sin\phi + \hat{\mathbf{k}}\cos\theta. \tag{2.39}$$

The spin component along this direction is obtained by projecting the spin vector $\mathbf{S}$ onto this new unit vector

$$\begin{aligned} S_n &= \mathbf{S}\cdot\hat{\mathbf{n}} \\ &= S_x\sin\theta\cos\phi + S_y\sin\theta\sin\phi + S_z\cos\theta. \end{aligned} \tag{2.40}$$

The matrix representations we found for S_x, S_y, and S_z lead to the matrix representation of the spin component operator S_n (Problem 2.6):

$$S_n \doteq \frac{\hbar}{2}\begin{pmatrix}\cos\theta & \sin\theta\, e^{-i\phi}\\ \sin\theta\, e^{i\phi} & -\cos\theta\end{pmatrix}. \tag{2.41}$$

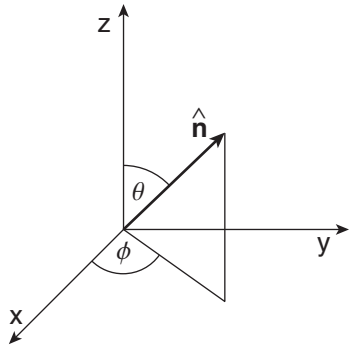


FIGURE 2.3 General direction along which to measure the spin component.

We have found a new operator, so to learn about its properties, we diagonalize it. Following the diagonalization procedure outlined in Section 2.1.2, we find that the eigenvalues of S_n are $\pm\hbar/2$ (Problem 2.7). So if we measure the spin component along any direction, we get only two possible results. This is to be expected from the experiments in Chapter 1. The eigenvectors for these two possible measurements are (Problem 2.7):

$$\begin{aligned} |+\rangle_n &= \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{i\phi}|-\rangle \\ |-\rangle_n &= \sin\frac{\theta}{2}|+\rangle - \cos\frac{\theta}{2}e^{i\phi}|-\rangle, \end{aligned} \tag{2.42}$$

where we again use the convention of choosing the first coefficient to be real and positive. It is important to point out that the $|+\rangle_n$ eigenstate (or equivalently the $|-\rangle_n$ eigenstate) can be used to represent any possible ket in a spin-1/2 system, if one allows for all possible angles $0 \leq \theta < \pi$ and $0 \leq \phi < 2\pi$. We generally write the most general state as $|\psi\rangle = a|+\rangle + b|-\rangle$, where a and b are complex. Requiring that the state be normalized and using the freedom to choose the first coefficient real and positive reduces this to

$$|\psi\rangle = |a||+\rangle + \sqrt{1 - |a|^2}e^{i\phi}|-\rangle. \tag{2.43}$$

If we change the parametrization of $|a|$ to $\cos(\theta/2)$, we see that $|+\rangle_n$ is equivalent to the most general state $|\psi\rangle$. This correspondence between the $|+\rangle_n$ eigenstate and the most general state is only valid in a two-state system such as spin 1/2. In systems with more dimensionality, it does not hold because more parameters are needed to specify the most general state than are afforded by the two angles θ and ϕ.

Example 2.2 Find the probabilities of the measurements shown in Fig. 2.4, assuming that the first Stern-Gerlach analyzer is aligned along the direction $\hat{\mathbf{n}}$ defined by the angles $\theta = 2\pi/3$ and $\phi = \pi/4$.

The measurement by the first Stern-Gerlach analyzer prepares the system in the spin up state $|+\rangle_n$ along the direction $\hat{\mathbf{n}}$. This state is then the input state to the second Stern-Gerlach analyzer. The input state is

$$\begin{aligned} |\psi_{in}\rangle = |+\rangle_n &= \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{i\phi}|-\rangle \\ &= \cos\frac{\pi}{3}|+\rangle + \sin\frac{\pi}{3}e^{i\pi/4}|-\rangle \\ &= \frac{1}{2}|+\rangle + \frac{\sqrt{3}}{2}e^{i\pi/4}|-\rangle. \end{aligned} \tag{2.44}$$

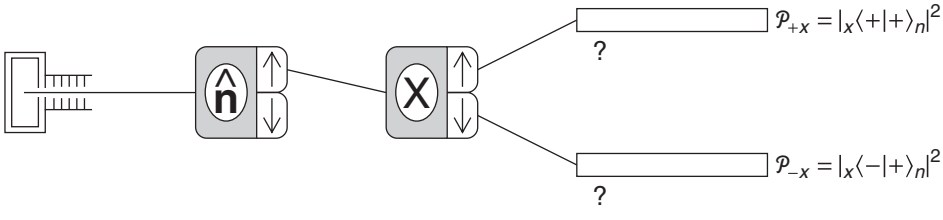


FIGURE 2.4 Measurement of the spin component after state preparation in a new direction.

The second analyzer is aligned along the x-axis, so the probabilities are

$$\begin{aligned} \mathcal{P}_{+x} &= \left|{}_x\langle +|\psi_{in}\rangle\right|^2 = \left|{}_x\langle +|+\rangle_n\right|^2 \\ \mathcal{P}_{-x} &= \left|{}_x\langle -|\psi_{in}\rangle\right|^2 = \left|{}_x\langle -|+\rangle_n\right|^2. \end{aligned} \tag{2.45}$$

Let's calculate the first probability using bra-ket notation, recalling that $|+\rangle_x = \frac{1}{\sqrt{2}}[|+\rangle + |-\rangle]$:

$$\begin{aligned} \mathcal{P}_{+x} &= \left|{}_x\langle +|+\rangle_n\right|^2 \\ &= \left|\tfrac{1}{\sqrt{2}}\big[\langle +| + \langle -|\big]\tfrac{1}{2}\big[|+\rangle + \sqrt{3}e^{i\pi/4}|-\rangle\big]\right|^2 \\ &= \left|\tfrac{1}{2\sqrt{2}}\big[1 + \sqrt{3}e^{i\pi/4}\big]\right|^2 \\ &= \tfrac{1}{8}\big[1 + \sqrt{3}e^{i\pi/4}\big]\big[1 + \sqrt{3}e^{-i\pi/4}\big] \\ &= \tfrac{1}{8}\big[1 + \sqrt{3}\big(e^{i\pi/4} + e^{-i\pi/4}\big) + 3\big] \\ &= \tfrac{1}{8}\big[4 + 2\sqrt{3}\cos(\pi/4)\big] \\ &= \tfrac{1}{8}\big[4 + 2\sqrt{3}/\sqrt{2}\big] \cong 0.806. \end{aligned} \tag{2.46}$$

Let's calculate the second probability using matrix notation, recalling that $|-\rangle_x = \frac{1}{\sqrt{2}}[|+\rangle - |-\rangle]$:

$$\begin{aligned} \mathcal{P}_{-x} &= \left|{}_x\langle -|+\rangle_n\right|^2 \\ &= \left|\tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\end{pmatrix}\tfrac{1}{2}\begin{pmatrix}1\\ \sqrt{3}e^{i\pi/4}\end{pmatrix}\right|^2 \\ &= \left|\tfrac{1}{2\sqrt{2}}\big[1 - \sqrt{3}e^{i\pi/4}\big]\right|^2 \\ &= \tfrac{1}{8}\big[4 - 2\sqrt{3}\cos(\pi/4)\big] \\ &= \tfrac{1}{8}\big[4 - 2\sqrt{3}/\sqrt{2}\big] \cong 0.194. \end{aligned} \tag{2.47}$$

The two results sum to unity as they must. A histogram of the measured results is shown in Fig. 2.5.

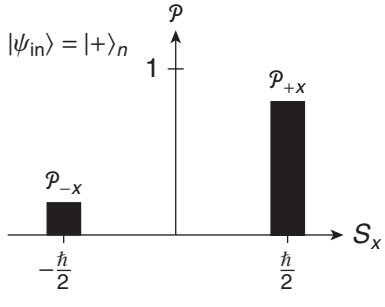


FIGURE 2.5 Histogram of spin component S_x measurement.

2.2.2 ■ Hermitian Operators

So far we have defined how operators act upon kets. For example, an operator A acts on a ket $|\psi\rangle$ to produce a new ket $|\phi\rangle = A|\psi\rangle$. The operator acts on the ket from the left; if the operator is on the right of the ket, the result is not defined, which is clear if you try to use matrix representation. Similarly, an operator acting on a bra must be on the right side of the bra

$$\langle\xi| = \langle\psi|A \tag{2.48}$$

and the result is another bra. However, the bra $\langle\xi| = \langle\psi|A$ is *not* the bra $\langle\phi|$ that corresponds to the ket $|\phi\rangle = A|\psi\rangle$. Rather the bra $\langle\phi|$ is found by defining a new operator $A^\dagger$ that obeys

$$\langle\phi| = \langle\psi|A^\dagger. \tag{2.49}$$

This new operator $A^\dagger$ is called the **Hermitian adjoint** of the operator A. We can learn something about the Hermitian adjoint by taking the inner product of the state $|\phi\rangle = A|\psi\rangle$ with another (unspecified) state $|\beta\rangle$

$$\begin{aligned} \langle\phi|\beta\rangle &= \langle\beta|\phi\rangle^* \\ [\langle\psi|A^+]|\beta\rangle &= \{\langle\beta|[A|\psi\rangle]\}^* \\ \langle\psi|A^+|\beta\rangle &= \langle\beta|A|\psi\rangle^*, \end{aligned} \tag{2.50}$$

which relates the matrix elements of A and $A^\dagger$. Equation (2.50) tells us that the matrix representing the Hermitian adjoint $A^\dagger$ is found by transposing and complex conjugating the matrix representing A. This is consistent with the definition of Hermitian adjoint used in matrix algebra.

An operator A is said to be **Hermitian** if it is equal to its Hermitian adjoint $A^\dagger$. If an operator is Hermitian, then the bra $\langle\psi|A$ is equal to the bra $\langle\phi|$ that corresponds to the ket $|\phi\rangle = A|\psi\rangle$. That is, a Hermitian operator can act to the right on a ket or to the left on a bra with the same result. In quantum mechanics, all operators that correspond to physical observables are Hermitian. This includes the spin operators we have already encountered as well as the energy, position, and momentum operators that we will introduce in later chapters. The Hermiticity of physical observables is important in light of two features of Hermitian matrices: (1) Hermitian matrices have real eigenvalues, which ensures that results of measurements are always real; and (2) the eigenvectors of a Hermitian matrix comprise a complete set of basis states, which ensures that we can use the eigenvectors of any observable as a valid basis.

2.2.3 ■ Projection Operators

For the spin-1/2 system, we now know four operators: S_x, S_y, S_z, and S_n. Let's look for some other operators. Consider the ket $|\psi\rangle$ written in terms of its coefficients in the S_z basis

$$\begin{aligned} |\psi\rangle &= a|+\rangle + b|-\rangle \\ &= (\langle+|\psi\rangle)|+\rangle + (\langle-|\psi\rangle)|-\rangle. \end{aligned} \tag{2.51}$$

Looking for the moment only at the first term, we can write it as a number times a ket, or as a ket times a number:

$$(\langle+|\psi\rangle)|+\rangle = |+\rangle(\langle+|\psi\rangle) \tag{2.52}$$

without changing its meaning. Using the second form, we can separate the bra and ket that form the inner product and obtain

$$|+\rangle(\langle+|\psi\rangle) = (|+\rangle\langle+|)|\psi\rangle. \tag{2.53}$$

The new term in parentheses is a product of a ket and a bra but in the opposite order compared to the inner product defined earlier. This new object must be an operator because it acts on the ket $|\psi\rangle$ and produces another ket: $(\langle+|\psi\rangle)|+\rangle$. This new type of operator is known as an **outer product**.

Returning now to Eq. (2.51), we write $|\psi\rangle$ using these new operators:

$$\begin{aligned} |\psi\rangle &= \langle+|\psi\rangle|+\rangle + \langle-|\psi\rangle|-\rangle \\ &= |+\rangle\langle+|\psi\rangle + |-\rangle\langle-|\psi\rangle \\ &= (|+\rangle\langle+| + |-\rangle\langle-|)|\psi\rangle. \end{aligned} \tag{2.54}$$

The term in parentheses is a sum of two outer products and is clearly an operator because it acts on a ket to produce another ket. In this special case, the result is the same as the original ket, so the operator must be the **identity operator 1**. This relationship is often written as

$$|+\rangle\langle+| + |-\rangle\langle-| = \mathbf{1} \tag{2.55}$$

and is known as the **completeness relation** or **closure**. It expresses the fact that the basis states $|\pm\rangle$ comprise a complete set of states, meaning any arbitrary ket can be written in terms of them. To make it obvious that outer products are operators, it is useful to express Eq. (2.55) in matrix notation using the standard rules of matrix multiplication:

$$\begin{aligned} |+\rangle\langle+| + |-\rangle\langle-| &\doteq \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} \\ &\doteq \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} \\ &\doteq \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}. \end{aligned} \tag{2.56}$$

Each outer product is represented by a matrix, as we expect for operators, and the sum of these two outer products is represented by the identity matrix, which we expected from Eq. (2.54).

Now consider the individual operators $|+\rangle\langle+|$ and $|-\rangle\langle-|$. These operators are called **projection** operators, and for spin 1/2 they are given by

$$\begin{aligned} P_+ &= |+\rangle\langle+| \doteq \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} \\ P_- &= |-\rangle\langle-| \doteq \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}. \end{aligned} \tag{2.57}$$

In terms of these new operators the completeness relation can also be written as

$$P_+ + P_- = \mathbf{1}. \tag{2.58}$$

When a projection operator for a particular eigenstate acts on a state $|\psi\rangle$, it produces a new ket that is aligned along the eigenstate and has a magnitude equal to the amplitude (including the phase) for the state $|\psi\rangle$ to be in that eigenstate. For example,

$$\begin{aligned} P_+|\psi\rangle &= |+\rangle\langle+|\psi\rangle = (\langle+|\psi\rangle)|+\rangle \\ P_-|\psi\rangle &= |-\rangle\langle-|\psi\rangle = (\langle-|\psi\rangle)|-\rangle. \end{aligned} \tag{2.59}$$

Note also that a projector acting on its corresponding eigenstate results in that eigenstate, and a projector acting on an orthogonal state results in zero:

$$\begin{aligned} P_+|+\rangle &= |+\rangle\langle +|+\rangle = |+\rangle \\ P_-|+\rangle &= |-\rangle\langle -|+\rangle = 0. \end{aligned} \tag{2.60}$$

Because the projection operator produces the probability amplitude, we expect that it must be intimately tied to measurement in quantum mechanics.

We found in Chapter 1 that the probability of a measurement is given by the square of the inner product of initial and final states (postulate 4). Using the new projection operators, we rewrite the probability as

$$\begin{aligned} \mathcal{P}_+ &= |\langle +|\psi\rangle|^2 \\ &= \langle +|\psi\rangle^*\langle +|\psi\rangle \\ &= \langle \psi|+\rangle\langle +|\psi\rangle \\ &= \langle \psi|P_+|\psi\rangle. \end{aligned} \tag{2.61}$$

Thus we say that the probability of the measurement $S_z = \hbar/2$ can be calculated as a matrix element of the projection operator, using the input state $|\psi\rangle$ and the projector P_+ corresponding to the result.

The other important aspect of quantum measurement that we learned in Chapter 1 is that a measurement disturbs the system. That is, if an input state $|\psi\rangle$ is measured to have $S_z = +\hbar/2$, then the output state is no longer $|\psi\rangle$ but is changed to $|+\rangle$. We saw above that the projection operator does this operation for us, with a multiplicative constant of the probability amplitude. Thus, if we divide by this amplitude, which is the square root of the probability, then we can describe the abrupt change of the input state as

$$|\psi'\rangle = \frac{P_+|\psi\rangle}{\sqrt{\langle \psi|P_+|\psi\rangle}} = |+\rangle, \tag{2.62}$$

where $|\psi'\rangle$ is the output state. This effect is described by the fifth postulate, which is presented below and is often referred to as the **projection postulate**.

Postulate 5

After a measurement of A that yields the result a_n, the quantum system is in a new state that is the normalized projection of the original system ket onto the ket (or kets) corresponding to the result of the measurement:

$$|\psi'\rangle = \frac{P_n|\psi\rangle}{\sqrt{\langle \psi|P_n|\psi\rangle}}.$$

The projection postulate is at the heart of quantum measurement. This effect is often referred to as the **collapse** (or **reduction** or **projection**) of the quantum state vector. The projection postulate clearly states that quantum measurements cannot be made without disturbing the system (except in the case where the input state is the same as the output state), in sharp contrast to classical measurements. The collapse of the quantum state makes quantum mechanics irreversible, again in contrast to classical mechanics.

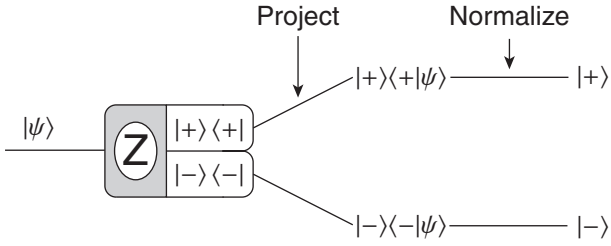


FIGURE 2.6 Schematic diagram of the role of the projection operator in a Stern-Gerlach spin measurement.

We can use the projection postulate to make a model of quantum measurement, as shown in the revised depiction of a Stern-Gerlach measurement system in Fig. 2.6. The projection operators act on the input state to produce output states with probabilities given by the squares of the amplitudes that the projection operations yield. For example, the input state $|\psi_{in}\rangle$ is acted on the projection operator $P_+ = |+\rangle\langle+|$, producing an output ket $|\psi_{out}\rangle = |+\rangle(\langle+|\psi_{in}\rangle)$ with probability $\mathcal{P}_+ = |\langle+|\psi_{in}\rangle|^2$. The output ket $|\psi_{out}\rangle = |+\rangle(\langle+|\psi_{in}\rangle)$ is really just a $|+\rangle$ ket that is not properly normalized, so we normalize it for use in any further calculations. We do not really know what is going on in the measurement process, so we cannot explain the mechanism of the collapse of the quantum state vector. This lack of understanding makes some people uncomfortable with this aspect of quantum mechanics and has been the source of much controversy surrounding quantum mechanics. Trying to better understand the measurement process in quantum mechanics is an ongoing research problem. However, despite our lack of understanding, the theory for predicting the results of experiments has been proven with very high accuracy.

2.2.4 ■ Analysis of Experiments 3 and 4

We can now return to Experiments 3 and 4 from Chapter 1 and analyze them with these new tools. Recall that Experiment 3 is the same as Experiment 4a, and Experiments 4a and 4b are similar in that they each use only one of the output ports of the second Stern-Gerlach analyzer as input to the third analyzer. Figure 2.7 depicts these experiments again, with Fig. 2.7(a) showing a hybrid experiment that is essentially Experiment 4a in its upper half and Experiment 4b in its lower half, and Fig. 2.7(b) showing Experiment 4c. In this problem, we discuss the probability that an atom leaving the first analyzer in the $|+\rangle$ state is detected in one of the counters connected to the output ports of the third analyzer. Such a probability involves two measurements at the second and third analyzers. The total probability is the product of the individual probabilities of each measurement.

For the hybrid experiment shown in Fig. 2.7(a), the probability of measuring an atom at the topmost counter is the probability of measuring $S_x = +\hbar/2$ at the second analyzer, $|{}_x\langle+|+\rangle|^2$, times the probability of measuring $S_z = +\hbar/2$ at the third analyzer, $|\langle+|+\rangle_x|^2$, giving

$$\mathcal{P}_{\text{upper},+} = |\langle+|+\rangle_x|^2 |{}_x\langle+|+\rangle|^2. \tag{2.63}$$

Likewise the probability of measuring the atom to have $S_x = +\hbar/2$ and then $S_z = -\hbar/2$ is

$$\mathcal{P}_{\text{upper},-} = |\langle-|+\rangle_x|^2 |{}_x\langle+|+\rangle|^2, \tag{2.64}$$

where we have written the product so as to be read from right to left as is the usual practice with quantum mechanical amplitudes and probabilities. For atoms that take the lower path from the second analyzer, the final probabilities are

$$\begin{aligned} \mathcal{P}_{\text{lower},+} &= |\langle +|-\rangle_x|^2 |_x\langle -|+\rangle|^2 \\ \mathcal{P}_{\text{lower},-} &= |\langle -|-\rangle_x|^2 |_x\langle -|+\rangle|^2. \end{aligned} \tag{2.65}$$

For Experiment 4c, shown in Fig. 2.7(b), we have a new situation at the second analyzer. Both output ports are connected to the third analyzer, which means that the probability of an atom from the first analyzer being input to the third analyzer is 100%. So we need only calculate the probability of passage through the third analyzer. The crucial step is determining the input state, for which we use the projection postulate. Because both states are used, the relevant projection operator is the sum of the two projection operators for each port, $P_{+x} + P_{-x}$, where $P_{+x} = |+\rangle_{x\,x}\langle +|$ and $P_{-x} = |-\rangle_{x\,x}\langle -|$. Thus the state after the second analyzer is

$$\begin{aligned} |\psi_2\rangle &= \frac{(P_{+x} + P_{-x})|\psi_1\rangle}{\sqrt{\langle \psi_1|(P_{+x} + P_{-x})|\psi_1\rangle}} \\ &= \frac{(P_{+x} + P_{-x})|+\rangle}{\sqrt{\langle +|(P_{+x} + P_{-x})|+\rangle}}. \end{aligned} \tag{2.66}$$

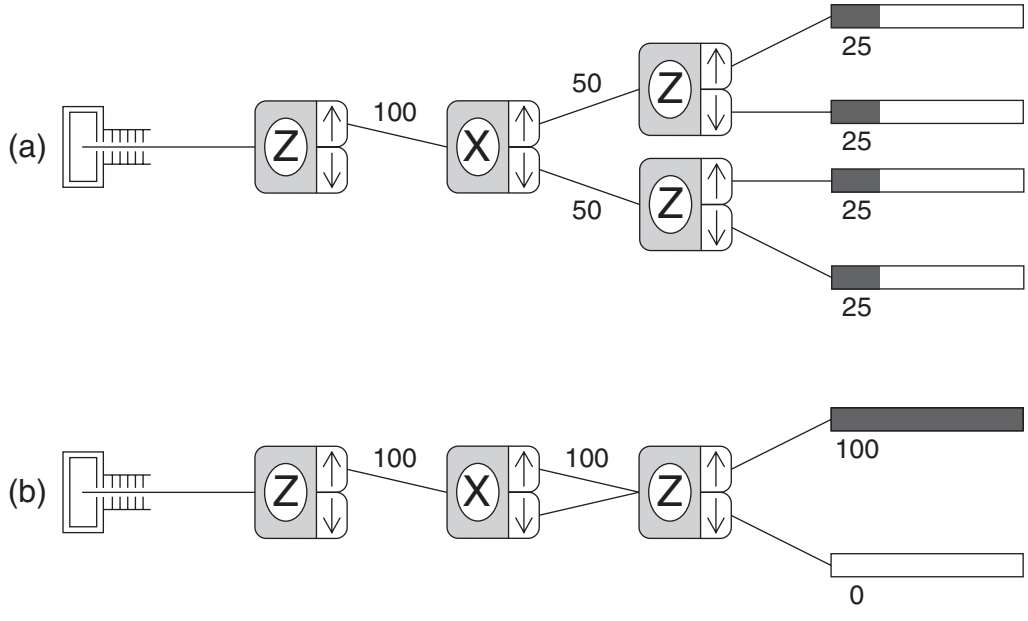


FIGURE 2.7 (a) Hybrid Experiment 4a and 4b, and (b) Experiment 4c.

In this simple example, the projector $P_{+x} + P_{-x}$ is equal to the identity operator because the two states form a complete basis. This clearly simplifies the calculation, giving $|\psi_2\rangle = |+\rangle$, but to illustrate our point, let's simplify only the denominator (which equals one), giving

$$\begin{aligned} |\psi_2\rangle &= (|+\rangle_{x\,x}\langle +| + |-\rangle_{x\,x}\langle -|)|+\rangle \\ &= |+\rangle_{x\,x}\langle +|+\rangle + |-\rangle_{x\,x}\langle -|+\rangle. \end{aligned} \tag{2.67}$$

Thus the beam entering the third analyzer can be viewed as a coherent superposition of the eigenstates of the second analyzer. Now calculate the probability of measuring spin up at the third analyzer:

$$\begin{aligned} \mathcal{P}_+ &= |\langle +|\psi_2\rangle|^2 \\ &= |\langle +|+\rangle_{x\,x}\langle +|+\rangle + \langle +|-\rangle_{x\,x}\langle -|+\rangle|^2. \end{aligned} \tag{2.68}$$

The probability of measuring spin down at the third analyzer is similarly

$$\begin{aligned} \mathcal{P}_- &= |\langle -|\psi_2\rangle|^2 \\ &= |\langle -|+\rangle_{x\,x}\langle +|+\rangle + \langle -|-\rangle_{x\,x}\langle -|+\rangle|^2. \end{aligned} \tag{2.69}$$

In each case, the probability is a square of a sum of amplitudes, each amplitude being the amplitude for a successive pair of measurements. For example, in $\mathcal{P}_-$ the amplitude $\langle -|+\rangle_{x\,x}\langle +|+\rangle$ refers to the upper path that the initial $|+\rangle$ state takes as it is first measured to be in the $|+\rangle_x$ state and then measured to be in the $|-\rangle$ state (read from right to left). This amplitude is added to the amplitude for the lower path because the beams of the second analyzer are combined, in the proper fashion, to create the input beam to the third analyzer. When the sum of amplitudes is squared, four terms are obtained, two squares and two cross terms, giving

$$\begin{aligned} \mathcal{P}_- = {} & |\langle -|+\rangle_{x\,x}\langle +|+\rangle|^2 + |\langle -|-\rangle_{x\,x}\langle -|+\rangle|^2 \\ & + \langle -|+\rangle^*_{x\,x}\langle +|+\rangle^*\langle -|-\rangle_{x\,x}\langle -|+\rangle \\ & + \langle -|+\rangle_{x\,x}\langle +|+\rangle\langle -|-\rangle^*_{x\,x}\langle -|+\rangle^* \\ = {} & \mathcal{P}_{\text{upper},-} + \mathcal{P}_{\text{lower},-} + \text{interference terms}. \end{aligned} \tag{2.70}$$

This tells us that the probability of detecting an atom to have spin down when both paths are used is the sum of the probabilities for detecting a spin down atom when either the upper path or the lower path is used alone *plus* additional cross terms involving both amplitudes, which are commonly called interference terms. It is these additional terms, which are not complex squares and so could be positive or negative, that allow the total probability to become zero in this case, illustrating the phenomenon of interference.

This interference arises from the nature of the superposition of states that enters the third analyzer. To illustrate, consider what happens if we change the superposition state to a mixed state, as we discussed previously in Section 1.2.3. Recall that a superposition state implies a beam with each atom in the same state, which is a combination of states, while a mixed state implies that the beam consists of atoms in separate states. As we have described it so far, Experiment 4c involves a superposition state as the input to the third analyzer. We can change this to a mixed state by "watching" to see which of the two output ports of the second analyzer each atom travels through. There are a variety of ways to imagine doing this experimentally. The usual idea proposed is to illuminate the paths with light and watch for the scattered light from the atoms. With proper design of the optics, the light can be localized

sufficiently to determine which path the atom takes. Hence, such experiments are generally referred to as "**Which Path**" or "**Welcher Weg**" experiments. Such experiments can be performed in the SPINS program by selecting the "Watch" feature. Once we know which path the atom takes, the state is not the superposition $|\psi_2\rangle$ described above, but is either $|+\rangle_x$ or $|-\rangle_x$, depending on which path produces the light signal. To find the probability that atoms are detected at the spin down counter of the third analyzer, we add the probabilities for atoms to follow the path $|+\rangle \rightarrow |+\rangle_x \rightarrow |-\rangle$ to the probability for other atoms to follow the path $|+\rangle \rightarrow |-\rangle_x \rightarrow |-\rangle$ because these are independent events, giving

$$\begin{aligned} \mathcal{P}_{\text{watch},-} &= |\langle -|+\rangle_x \, {}_x\langle +|+\rangle|^2 + |\langle -|-\rangle_x \, {}_x\langle -|+\rangle|^2 \\ &= \mathcal{P}_{\text{upper},-} + \mathcal{P}_{\text{lower},-}, \end{aligned} \tag{2.71}$$

in which no interference terms are present.

This interference example illustrates again the important distinction between a coherent superposition state and a statistical mixed state. In a coherent superposition, there is a definite relative phase between the different states, which gives rise to interference effects that are dependent on that phase. In a statistical mixed state, the phase relationship between the states has been destroyed and the interference is washed out. Now we can understand what it takes to have the beams "properly" combined after the second analyzer of Experiment 4c. The relative phases of the two paths must be preserved. Anything that randomizes the phase is equivalent to destroying the superposition and leaving only a statistical mixture. If the beams are properly combined to leave the superposition intact, the results of Experiment 4c are the same as if no measurement were made at the second analyzer. So even though we have used a measuring device in the middle of Experiment 4c, we generally say that no measurement was made there. We can summarize our conclusions by saying that if no measurement is made on the intermediate state, then we add amplitudes and then square to find the probability, while if an intermediate measurement is performed (i.e., watching), then we square the amplitudes first and then add to find the probability. One is the square of a sum and the other is the sum of squares, and only the former exhibits interference.

2.3 ■ MEASUREMENT

Let's discuss how the probabilistic nature of quantum mechanics affects the way experiments are performed and compared with theory. In classical physics, a theoretical prediction can be reliably compared to a single experimental result. For example, a prediction of the range of a projectile can be tested by doing an experiment. The experiment may be repeated several times in order to understand and possibly reduce any systematic errors (e.g., wind) and measurement errors (e.g., misreading the tape measure). In quantum mechanics, a single measurement is meaningless. If we measure an atom to have spin up in a Stern-Gerlach analyzer, we cannot discern whether the original state was $|+\rangle$ or $|-\rangle_x$ or any arbitrary state $|\psi\rangle$ (except $|-\rangle$). Moreover, we cannot repeat the measurement on the same atom, because the original measurement changed the state, per the projection postulate.

Thus, one must, by necessity, perform identical measurements on identically prepared systems. In the spin-1/2 example, an initial Stern-Gerlach analyzer is used to prepare atoms in a particular state $|\psi\rangle$. Then a second Stern-Gerlach analyzer is used to perform the same experiment on each identically prepared atom. Consider performing a measurement of S_z on N identically prepared atoms. Let N_+ be the number of times the result $+\hbar/2$ is recorded and N_- be the number of times the result $-\hbar/2$ is recorded. Because there are only two possible results for each measurement, we must have $N = N_+ + N_-$. The probability postulate (postulate 4) predicts that the probability of measuring $+\hbar/2$ is

$$\mathcal{P}_+ = |\langle +|\psi\rangle|^2. \tag{2.72}$$

For a finite number N of atoms, we expect that N_+ is only approximately equal to $\mathcal{P}_+ N$ due to the statistical fluctuations inherent in a random process. Only in the limit of an infinite number N do we expect exact agreement:

$$\lim_{N\to\infty} \frac{N_+}{N} = \mathcal{P}_+ = |\langle +|\psi\rangle|^2. \tag{2.73}$$

It is useful to characterize a data set in terms of the **mean** and **standard deviation** (see Appendix A for further information on probability). The mean value of a data set is the average of all the measurements. The expected or predicted mean value of a measurement is the sum of the products of each possible result and its probability, which for this spin-1/2 measurement is

$$\boxed{\langle S_z\rangle = \left(+\frac{\hbar}{2}\right)\mathcal{P}_+ + \left(-\frac{\hbar}{2}\right)\mathcal{P}_-}, \tag{2.74}$$

where the angle brackets signify average or mean value. Using the rules of quantum mechanics we rewrite this mean value as

$$\begin{aligned}
\langle S_z\rangle &= +\frac{\hbar}{2}|\langle +|\psi\rangle|^2 + \left(-\frac{\hbar}{2}\right)|\langle -|\psi\rangle|^2 \\
&= +\frac{\hbar}{2}\langle\psi|+\rangle\langle +|\psi\rangle + \left(-\frac{\hbar}{2}\right)\langle\psi|-\rangle\langle -|\psi\rangle \\
&= \langle\psi|\left[+\frac{\hbar}{2}|+\rangle\langle +|\psi\rangle + \left(-\frac{\hbar}{2}\right)|-\rangle\langle -|\psi\rangle\right] \\
&= \langle\psi|[S_z|+\rangle\langle +|\psi\rangle + S_z|-\rangle\langle -|\psi\rangle] \\
&= \langle\psi|S_z[|+\rangle\langle +| + |-\rangle\langle -|]|\psi\rangle.
\end{aligned} \tag{2.75}$$

According to the completeness relation, the term in square brackets in the last line is unity, so we obtain

$$\boxed{\langle S_z\rangle = \langle\psi|S_z|\psi\rangle}. \tag{2.76}$$

We now have two ways to calculate the predicted mean value, Eq. (2.74) and Eq. (2.76). Which you use generally depends on what quantities you have readily available. The matrix element version in Eq. (2.76) is more common and is especially useful in systems that are more complicated than the 2-level spin-1/2 system. This predicted mean value is commonly called the **expectation value**, but it is *not* the expected value of any single experiment. Rather it is the expected mean value of a large number of experiments. It is *not* a time average, but an average over many identical experiments. For a general quantum mechanical observable, the expectation value is

$$\boxed{\langle A\rangle = \langle\psi|A|\psi\rangle = \sum_n a_n \mathcal{P}_{a_n}}, \tag{2.77}$$

where a_n are the eigenvalues of the operator A.

To see how the concept of expectation values applies to our study of spin-1/2 systems, consider two examples. First consider a system prepared in the state $|+\rangle$. The expectation value of S_z is

$$\langle S_z\rangle = \langle +|S_z|+\rangle, \tag{2.78}$$

which we calculate with bra-ket notation

$$\begin{aligned}\langle S_z\rangle &= \langle +|S_z|+\rangle \\ &= \langle +|\frac{\hbar}{2}|+\rangle \\ &= \frac{\hbar}{2}\langle +|+\rangle \\ &= \frac{\hbar}{2}.\end{aligned} \tag{2.79}$$

This result should seem obvious because $+\hbar/2$ is the only possible result of a measurement of S_z for the $|+\rangle$ state, so it must be the expectation value.

Next consider a system prepared in the state $|+\rangle_x$. In this case, the expectation value of S_z is

$$\langle S_z\rangle = {}_x\langle +|S_z|+\rangle_x. \tag{2.80}$$

Using matrix notation, we obtain

$$\begin{aligned}\langle S_z\rangle &= \tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} \\ &= \frac{\hbar}{4}\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1\\ -1\end{pmatrix} = 0\hbar.\end{aligned} \tag{2.81}$$

Again this is what you expect, because the two possible measurement results $\pm\hbar/2$ each have 50% probability, so the average value is zero. Note that the value of zero is never measured, so it is not the value "expected" for any given measurement, but rather the expected mean value of an ensemble of measurements.

In addition to the mean value, it is common to characterize a measurement by the standard deviation, which quantifies the spread of measurements about the mean or expectation value. The standard deviation is defined as the *square root* of the *mean* of the *square* of the deviations from the mean, and for an observable A is given by

$$\Delta A = \sqrt{\langle (A - \langle A\rangle)^2\rangle}, \tag{2.82}$$

where the angle brackets signify average value as used in the definition of an expectation value. This result is also often called the **root-mean-square deviation**, or r.m.s. deviation. We need to square the deviations, because the deviations from the mean are equally distributed above and below the mean in such a way that the average of the deviations themselves is zero. This expression can be simplified by expanding the square and performing the averages, resulting in

$$\begin{aligned}\Delta A &= \sqrt{\langle \left(A^2 - 2A\langle A\rangle + \langle A\rangle^2\right)\rangle} \\ &= \sqrt{\langle A^2\rangle - 2\langle A\rangle\langle A\rangle + \langle A\rangle^2} \\ &= \sqrt{\langle A^2\rangle - \langle A\rangle^2},\end{aligned} \tag{2.83}$$

where one must be clear to distinguish between the square of the mean $\langle A \rangle^2$ and the mean of the square $\langle A^2 \rangle$. While the mean of the square of an observable may not be a common experimental quantity, it can be calculated using the definition of the expectation value

$$\langle A^2 \rangle = \langle \psi | A^2 | \psi \rangle. \tag{2.84}$$

The square of an operator means that the operator acts twice in succession:

$$A^2 |\psi\rangle = AA|\psi\rangle = A(A|\psi\rangle). \tag{2.85}$$

To gain experience with the standard deviation, return to the two examples used above. To calculate the standard deviation, we need to find the mean of the square of the operator S_z. In the first case ($|+\rangle$ initial state), we get

$$\begin{aligned} \langle S_z^2 \rangle &= \langle + | S_z^2 | + \rangle = \langle + | S_z S_z | + \rangle = \langle + | S_z \frac{\hbar}{2} | + \rangle \\ &= \langle + | \left(\frac{\hbar}{2} \right)^2 | + \rangle \\ &= \left(\frac{\hbar}{2} \right)^2. \end{aligned} \tag{2.86}$$

We already have the mean of the operator S_z in Eq. (2.79) so the standard deviation is

$$\begin{aligned} \Delta S_z &= \sqrt{\langle S_z^2 \rangle - \langle S_z \rangle^2} \\ &= \sqrt{\left(\frac{\hbar}{2} \right)^2 - \left(\frac{\hbar}{2} \right)^2} \\ &= 0\hbar, \end{aligned} \tag{2.87}$$

which is to be expected because there is only one possible result, and hence no spread in the results of the measurement, as shown in the histogram in Fig. 2.8(a).

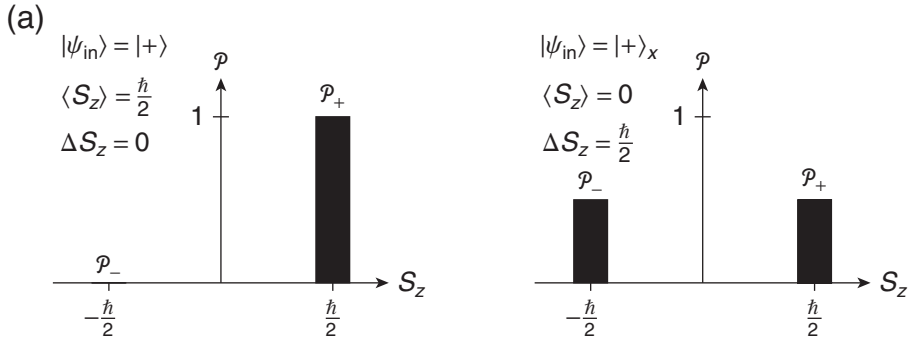


FIGURE 2.8 Idealized measurements of S_z with (a) a $|+\rangle$ input state and (b) with a $|+\rangle_x$ input state.

In the second case ($|+\rangle_x$ initial state), the mean of the square of the operator S_z is

$$\begin{aligned}\langle S_z^2\rangle &= {}_x\langle +|S_z^2|+\rangle_x \\ &= \tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} \\ &= \tfrac{1}{2}\left(\frac{\hbar}{2}\right)^2\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1\\ -1\end{pmatrix} \\ &= \tfrac{1}{2}\left(\frac{\hbar}{2}\right)^2\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix} \\ &= \left(\frac{\hbar}{2}\right)^2.\end{aligned} \tag{2.88}$$

The mean of the operator S_z is in Eq. (2.81), giving a standard deviation of

$$\begin{aligned}\Delta S_z &= \sqrt{\langle S_z^2\rangle - \langle S_z\rangle^2} \\ &= \sqrt{\left(\frac{\hbar}{2}\right)^2 - 0\hbar^2} \\ &= \frac{\hbar}{2}.\end{aligned} \tag{2.89}$$

Again this makes sense because each measurement deviates from the mean ($0\hbar$) by the same value of $\hbar/2$, as shown in the histogram in Fig. 2.8(b).

The standard deviation ΔA represents the uncertainty in the results of an experiment. In quantum mechanics, this uncertainty is inherent and fundamental, meaning that you cannot design the experiment any better to improve the result. What we have calculated then is the minimum uncertainty allowed by quantum mechanics. Any actual uncertainty may be larger due to experimental error. This is another ramification of the probabilistic nature of quantum mechanics and will lead us to the Heisenberg uncertainty relation in Section 2.5.

2.4 ■ COMMUTING OBSERVABLES

We found in Experiment 3 that two incompatible observables could not be known or measured simultaneously, because measurement of one somehow erased knowledge of the other. Let us now explore further what it means for two observables to be incompatible and how incompatibility affects the results of measurements. First we need to define a new object called a **commutator**. The commutator of two operators is defined as the difference between the products of the two operators taken in alternate orders:

$$[A,B] = AB - BA. \tag{2.90}$$

If the commutator is equal to zero, we say that the operators or observables **commute**; if it is not zero, we say they don't commute. Whether or not two operators commute has important ramifications in analyzing a quantum system and in making measurements of the two observables represented by those operators.

Consider what happens when two operators A and B do commute:

$$\begin{aligned} [A,B] &= 0 \\ AB - BA &= 0 \\ AB &= BA. \end{aligned} \tag{2.91}$$

Thus, for commuting operators the order of operation does not matter, whereas it does for noncommuting operators. Now let $|a\rangle$ be an eigenstate of the operator A with eigenvalue a:

$$A|a\rangle = a|a\rangle. \tag{2.92}$$

Operate on both sides of this equation with the operator B and use the fact that A and B commute:

$$\begin{aligned} BA|a\rangle &= Ba|a\rangle \\ AB|a\rangle &= aB|a\rangle \\ A(B|a\rangle) &= a(B|a\rangle). \end{aligned} \tag{2.93}$$

The last equation says that the state $B|a\rangle$ is also an eigenstate of the operator A with the same eigenvalue a. Assuming that each eigenvalue has a unique eigenstate (which is true if there is no degeneracy, but we haven't discussed degeneracy yet), the state $B|a\rangle$ must be some scalar multiple of the state $|a\rangle$. If we call this multiple b, then we can write

$$B|a\rangle = b|a\rangle, \tag{2.94}$$

which is just an eigenvalue equation for the operator B. Thus, we must conclude that the state $|a\rangle$ is also an eigenstate of the operator B, with the eigenvalue b. The assumption that the operators A and B commute has led us to the result that A and B have common or **simultaneous sets of eigenstates**. This result bears repeating:

Commuting operators share common eigenstates.

The ramifications of this result for experiments are very important. Recall that a measurement of the observable A projects the initial state $|\psi\rangle$ onto an eigenstate of A: $|a\rangle$. A subsequent measurement of the observable B then projects the input state $|a\rangle$ onto an eigenstate of B. But the eigenstates of the commuting operators A and B are the same, so the second measurement does not change the state $|a\rangle$. Thus, another measurement of A following the measurement of B yields the same result as the initial measurement of A, as illustrated in Fig. 2.9. Thus we say that we can know the eigenvalues of these two observables **simultaneously**. It is common to extend this language and say that these two observables can be measured simultaneously, although, as illustrated in Fig. 2.9, we do not really measure them simultaneously. What we mean is that we can measure one observable without erasing our knowledge of the previous results of the other observable. Observables A and B are said to be **compatible**.

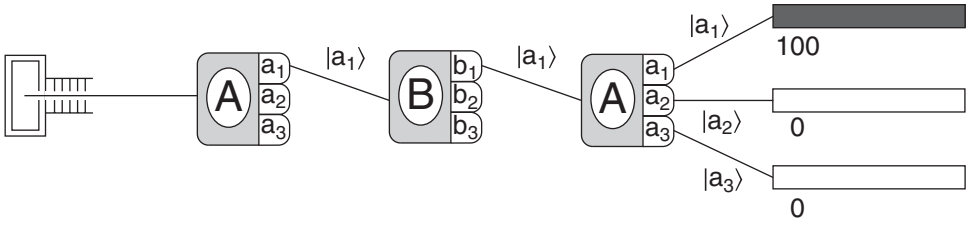


FIGURE 2.9 Successive measurements of commuting observables.

Conversely, if two operators do not commute, then they are incompatible observables and cannot be measured or known simultaneously. This is what we saw in Experiment 3 in Chapter 1. In that case, the two observables were S_x and S_z. Let's take a look at their commutator to show that they are not compatible:

$$\begin{aligned}
[S_z,S_x] &\doteq \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} - \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\\
&\doteq \left(\frac{\hbar}{2}\right)^2\left[\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} - \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\right]\\
&\doteq \left(\frac{\hbar}{2}\right)^2\begin{pmatrix}0 & 2\\ -2 & 0\end{pmatrix}\\
&= i\hbar S_y.
\end{aligned} \tag{2.95}$$

As expected, these two operators do not commute. In fact, none of the spin component operators commute with each other. The complete commutation relations are

$$\boxed{\begin{aligned}[S_x,S_y] &= i\hbar S_z\\ [S_y,S_z] &= i\hbar S_x\\ [S_z,S_x] &= i\hbar S_y\end{aligned}}, \tag{2.96}$$

so written to make the cyclic relations clear.

When we represent operators as matrices, we can often decide whether two operators commute by inspection of the matrices. Recall the important statement: *An operator is always diagonal in its own basis.* If you are presented with two matrices that are both diagonal, they must share a common basis, and so they commute with each other. To be explicit, the product of two diagonal matrices

$$\begin{aligned}
AB &\doteq \begin{pmatrix}a_1 & 0 & 0 & \cdots\\ 0 & a_2 & 0 & \cdots\\ 0 & 0 & a_3 & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{pmatrix}\begin{pmatrix}b_1 & 0 & 0 & \cdots\\ 0 & b_2 & 0 & \cdots\\ 0 & 0 & b_3 & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{pmatrix}\\
&\doteq \begin{pmatrix}a_1b_1 & 0 & 0 & \cdots\\ 0 & a_2b_2 & 0 & \cdots\\ 0 & 0 & a_3b_3 & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{pmatrix},
\end{aligned} \tag{2.97}$$

is clearly independent of the order of the product. Note, however, that you may *not* conclude that two operators do not commute if one is diagonal and one is not, nor if both are not diagonal.

2.5 ■ UNCERTAINTY PRINCIPLE

The intimate connection between the commutator of two observables and the possible precision of measurements of the two corresponding observables is reflected in an important relation that we simply state here (see more advanced texts for a derivation). The product of the uncertainties or standard deviations of two observables is related to the commutator of the two observables:

$$\boxed{\Delta A\Delta B \geq \tfrac{1}{2}\left|\langle[A,B]\rangle\right|}. \tag{2.98}$$

This is the **uncertainty principle** of quantum mechanics. Consider what it says about a simple Stern-Gerlach experiment. The uncertainty principle for the S_x and S_y spin components is

$$\begin{aligned}\Delta S_x \Delta S_y &\geq \tfrac{1}{2}|\langle [S_x, S_y]\rangle| \\ &\geq \tfrac{1}{2}|\langle i\hbar S_z\rangle| \\ &\geq \frac{\hbar}{2}|\langle S_z\rangle|.\end{aligned} \tag{2.99}$$

These uncertainties are the minimal quantum mechanical uncertainties that would arise in any experiment. Any experimental uncertainties due to experimenter error, apparatus errors, and statistical limitations would be additional.

Let's now apply the uncertainty principle to Experiment 3 where we first learned of the impact of measurements in quantum mechanics. If the initial state is $|+\rangle$, then a measurement of S_z results in an expectation value $\langle S_z\rangle = \hbar/2$ with an uncertainty $\Delta S_z = 0$, as illustrated in Fig. 2.8(a). Thus the uncertainty principle dictates that the product of the other uncertainties for measurements of the $|+\rangle$ state is

$$\Delta S_x \Delta S_y \geq \left(\frac{\hbar}{2}\right)^2, \tag{2.100}$$

or simply

$$\Delta S_x \Delta S_y \neq 0. \tag{2.101}$$

This implies that

$$\begin{aligned}\Delta S_x &\neq 0 \\ \Delta S_y &\neq 0.\end{aligned} \tag{2.102}$$

The conclusion to draw from this is that while we can know one spin component absolutely ($\Delta S_z = 0$), we can never know all three, nor even two, simultaneously. This is in agreement with our results from Experiment 3. This lack of ability to measure all spin components simultaneously implies that the spin does not really point in a given direction, as a classical spin or angular momentum does. So when we say that we have measured "spin up," we really mean only that the spin component along that axis is up, as opposed to down, and not that the complete spin angular momentum vector points up along that axis.

2.6 ■ S^2 OPERATOR

Another indication that the spin does not point along the axis along which you measure the spin component is obtained by considering a new operator that represents the magnitude of the spin vector but has no information about the direction. It is common to use the square of the spin vector for this task. This new operator is

$$\mathbf{S}^2 = S_x^2 + S_y^2 + S_z^2, \tag{2.103}$$

and it is calculated in the S_z representation as

$$\begin{aligned}\mathbf{S}^2 &\doteq \left(\frac{\hbar}{2}\right)^2\left[\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} + \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right] \\ &\doteq \left(\frac{\hbar}{2}\right)^2\left[\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right] \\ &\doteq \tfrac{3}{4}\hbar^2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.\end{aligned} \tag{2.104}$$

Thus the $\mathbf{S}^2$ operator is proportional to the identity operator, which means it must commute with all the other operators S_x, S_y, and S_z. It also means that all states are eigenstates of the $\mathbf{S}^2$ operator. Thus, we can write

$$\mathbf{S}^2|\psi\rangle = \tfrac{3}{4}\hbar^2|\psi\rangle \tag{2.105}$$

for any state $|\psi\rangle$ in the spin-1/2 system.

For the case of spin 1/2, note that the expectation value of the operator $\mathbf{S}^2$ is

$$\langle \mathbf{S}^2 \rangle = \tfrac{3}{4}\hbar^2, \tag{2.106}$$

which would imply that the "length" of the spin vector is

$$|\mathbf{S}| = \sqrt{\langle \mathbf{S}^2 \rangle} = \sqrt{3}\,\frac{\hbar}{2}. \tag{2.107}$$

This is appreciably longer than the measured component of $\hbar/2$, implying that the spin vector can never be fully aligned along any axis. A useful mental model of the spin vector and its component is shown in Fig. 2.10. In this **vector model**, one can imagine the total spin vector $\mathbf{S}$ precessing around the z-axis at a constant angle to form a cone, with a constant spin component S_z. For a spin-1/2 system in the "spin up" state $|+\rangle$, this classical model yields the same expectation values and uncertainties as the quantum model (Problem 2.9)

$$\begin{aligned} \langle S_z \rangle &= \frac{\hbar}{2} \qquad \Delta S_z = 0 \\ \langle S_x \rangle &= 0 \qquad \Delta S_x \neq 0 \\ \langle S_y \rangle &= 0 \qquad \Delta S_y \neq 0. \end{aligned} \tag{2.108}$$

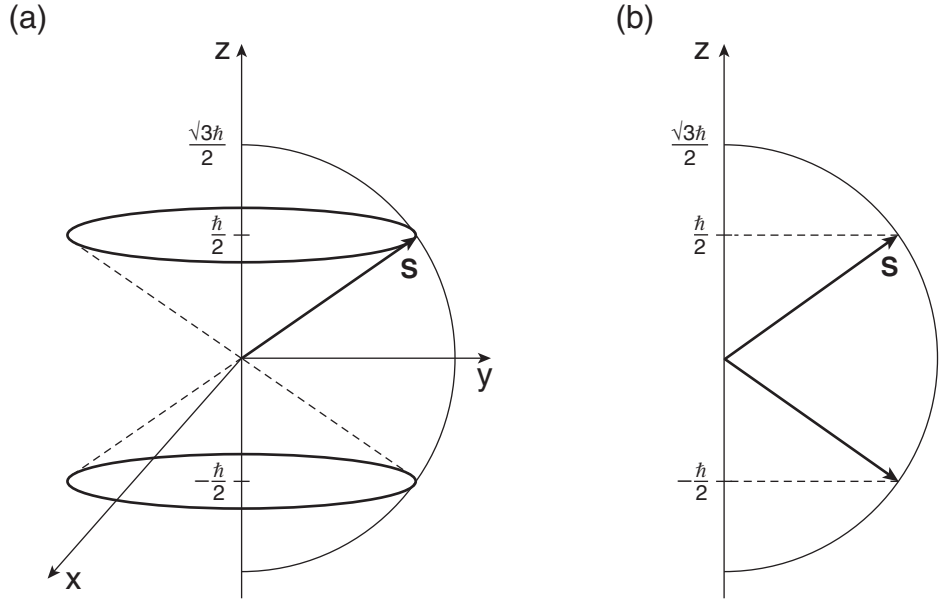


FIGURE 2.10 (a) Vector model illustrating the classical precision of a spin vector and the allowed quantum mechanical components. (b) Two-dimensional version of the vector model with constant spin vector length and two possible components.

However, a quantum mechanical experiment on a spin component eigenstate does not yield the time dependence of the precession implied by the picture in Fig. 2.10(a). Rather, the quantum mechanical spin vector is more accurately thought of as smeared out over the whole cone in a uniform random sense. This randomness is often termed **quantum fuzziness** and will be evident in other systems we will study later. To avoid the inaccurate precession part of the vector model, it is often illustrated as in Fig. 2.10(b).

2.7 ■ SPIN-1 SYSTEM

The Stern-Gerlach experiment depicted in Fig. 1.1 can be performed on a variety of atoms or particles. Such experiments always result in a finite number of discrete beams exiting the analyzer. For spin-1/2 particles, there are two output beams. For the case of three output beams, the deflections are consistent with magnetic moments arising from spin angular momentum components of $1\hbar$, $0\hbar$, and $-1\hbar$. For an analyzer aligned along the z-axis, the three output states are labeled $|1\rangle$, $|0\rangle$, and $|-1\rangle$, as shown in Fig. 2.11. This is what we call a **spin-1** system. (Note that the SPINS software and our Stern-Gerlach schematics use arrows for the $|1\rangle$ and $|-1\rangle$ output beams, but these outputs are not the same as the spin-1/2 states that are also denoted with arrows.)

The three eigenvalue equations for the spin component operator S_z of a spin-1 system are

$$\begin{aligned} S_z|1\rangle &= \hbar|1\rangle \\ S_z|0\rangle &= 0\hbar|0\rangle \\ S_z|-1\rangle &= -\hbar|-1\rangle. \end{aligned} \tag{2.109}$$

As we did in the spin-1/2 case, we choose the S_z basis as the standard basis in which to express kets and operators using matrix representation. In Section 2.1, we found that *eigenvectors are unit vectors in their own basis* and *an operator is always diagonal in its own basis*. Using the first rule, we can immediately write down the eigenvectors of the S_z operator:

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad |0\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad |-1\rangle \doteq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{2.110}$$

where we again use the convention that the ordering of the rows follows the eigenvalues in descending order. Using the second rule, we write down the S_z operator

$$S_z \doteq \begin{pmatrix} 1\hbar & 0 & 0 \\ 0 & 0\hbar & 0 \\ 0 & 0 & -1\hbar \end{pmatrix} = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{2.111}$$

with the eigenvalues $1\hbar$, $0\hbar$, and $-1\hbar$ ordered along the diagonal. The value zero is a perfectly valid eigenvalue in some systems.

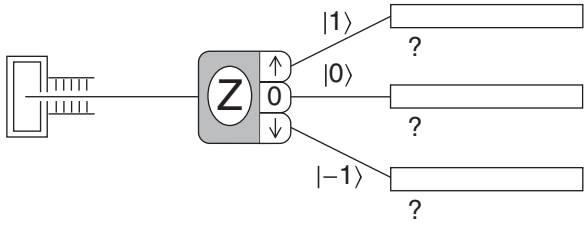


FIGURE 2.11 Spin-1 Stern-Gerlach experiment.

The same four experiments performed on the spin-1/2 system can be performed on a spin-1 system. Conceptually the results are the same. One important difference occurs in Experiment 2, where a measurement of S_z is first performed to prepare a particular state, and then a subsequent measurement of S_x (or S_y) is performed. Based upon the results of the spin-1/2 experiment, one might expect each of the possible components to have one-third probability. Such is not the case. Rather, one set of results is

$$\begin{aligned} \mathcal{P}_{1x} &= |{}_x\langle 1|1\rangle|^2 = \tfrac{1}{4} \\ \mathcal{P}_{0x} &= |{}_x\langle 0|1\rangle|^2 = \tfrac{1}{2} \\ \mathcal{P}_{-1x} &= |{}_x\langle -1|1\rangle|^2 = \tfrac{1}{4}, \end{aligned} \tag{2.112}$$

as illustrated in Fig. 2.12. These experimental results can be used to determine the S_x eigenstates in terms of the S_z basis

$$\begin{aligned} |1\rangle_x &= \tfrac{1}{2}|1\rangle + \tfrac{1}{\sqrt{2}}|0\rangle + \tfrac{1}{2}|-1\rangle \\ |0\rangle_x &= \tfrac{1}{\sqrt{2}}|1\rangle - \tfrac{1}{\sqrt{2}}|-1\rangle \\ |-1\rangle_x &= \tfrac{1}{2}|1\rangle - \tfrac{1}{\sqrt{2}}|0\rangle + \tfrac{1}{2}|-1\rangle. \end{aligned} \tag{2.113}$$

Likewise, we can find the S_y eigenstates:

$$\begin{aligned} |1\rangle_y &= \tfrac{1}{2}|1\rangle + i\tfrac{1}{\sqrt{2}}|0\rangle - \tfrac{1}{2}|-1\rangle \\ |0\rangle_y &= \tfrac{1}{\sqrt{2}}|1\rangle + \tfrac{1}{\sqrt{2}}|-1\rangle \\ |-1\rangle_y &= \tfrac{1}{2}|1\rangle - i\tfrac{1}{\sqrt{2}}|0\rangle - \tfrac{1}{2}|-1\rangle. \end{aligned} \tag{2.114}$$

The matrix representations of the S_x and S_y operators are

$$S_x \doteq \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad S_y \doteq \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{2.115}$$

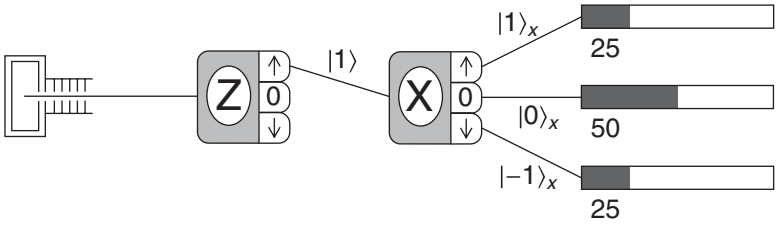


FIGURE 2.12 Experiment 2 in the spin-1 case.

Example 2.3 A spin-1 system is prepared in the state

$$|\psi_{in}\rangle = \tfrac{2}{\sqrt{6}}|1\rangle - \tfrac{i}{\sqrt{6}}|0\rangle + \tfrac{i}{\sqrt{6}}|-1\rangle. \tag{2.116}$$

Find the probabilities of measuring each of the possible spin components along the z-axis.

The probability of measuring $S_z = +1\hbar$ is

$$\begin{aligned}\mathcal{P}_1 &= |\langle 1|\psi_{in}\rangle|^2 \\ &= \left|\langle 1|\left[\tfrac{2}{\sqrt{6}}|1\rangle - \tfrac{i}{\sqrt{6}}|0\rangle + \tfrac{i}{\sqrt{6}}|-1\rangle\right]\right|^2 \\ &= \left|\tfrac{2}{\sqrt{6}}\langle 1|1\rangle - \tfrac{i}{\sqrt{6}}\langle 1|0\rangle + \tfrac{i}{\sqrt{6}}\langle 1|-1\rangle\right|^2 \\ &= \left|\tfrac{2}{\sqrt{6}}\right|^2 = \tfrac{2}{3}.\end{aligned} \tag{2.117}$$

The probability of measuring $S_z = 0\hbar$ is

$$\begin{aligned}\mathcal{P}_0 &= |\langle 0|\psi_{in}\rangle|^2 \\ &= \left|\langle 0|\left[\tfrac{2}{\sqrt{6}}|1\rangle - \tfrac{i}{\sqrt{6}}|0\rangle + \tfrac{i}{\sqrt{6}}|-1\rangle\right]\right|^2 \\ &= \left|\tfrac{-i}{\sqrt{6}}\right|^2 = \tfrac{1}{6}.\end{aligned} \tag{2.118}$$

The probability of measuring $S_z = -1\hbar$ is

$$\begin{aligned}\mathcal{P}_{-1} &= |\langle -1|\psi_{in}\rangle|^2 \\ &= \left|\langle -1|\left[\tfrac{2}{\sqrt{6}}|1\rangle - \tfrac{i}{\sqrt{6}}|0\rangle + \tfrac{i}{\sqrt{6}}|-1\rangle\right]\right|^2 \\ &= \left|\tfrac{i}{\sqrt{6}}\right|^2 = \tfrac{1}{6}.\end{aligned} \tag{2.119}$$

The three probabilities add to unity, as they must. A histogram of the predicted measurement results is shown in Fig. 2.13.

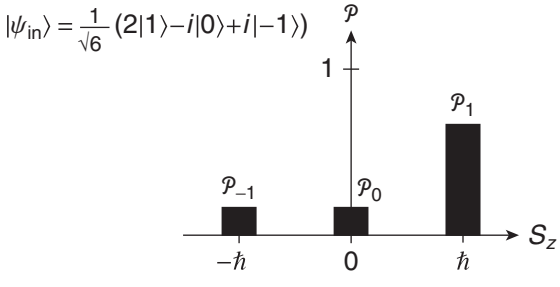


FIGURE 2.13 Histogram of measurements of z-component of spin for spin-1 particle.

To generalize to other possible spin systems, we need to introduce new labels. We use the label s to denote the spin of the system, such as spin 1/2, spin 1, spin 3/2. The number of beams exiting a Stern-Gerlach analyzer is $2s + 1$. In each of these cases, a measurement of a spin component along any axis yields results ranging from a maximum value of $s\hbar$ to a minimum value of $-s\hbar$, in unit steps of the value $\hbar$. We denote the possible values of the spin component along the z-axis by the label m, the integer or half-integer multiplying $\hbar$. A quantum state with specific values of s and m is denoted as $|sm\rangle$, yielding the eigenvalue equations

$$\begin{aligned} \mathbf{S}^2|sm\rangle &= s(s+1)\hbar^2|sm\rangle \\ S_z|sm\rangle &= m\hbar|sm\rangle. \end{aligned} \tag{2.120}$$

The label s is referred to as the **spin angular momentum quantum number** or the **spin quantum number** for short. The label m is referred to as the **spin component quantum number** or the **magnetic quantum number** because of its role in magnetic field experiments like the Stern-Gerlach experiment. The connection between this new $|sm\rangle$ notation and the spin-1/2 $|\pm\rangle$ notation is

$$\begin{aligned} \left|\tfrac{1}{2}\,\tfrac{1}{2}\right\rangle &= |+\rangle \\ \left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle &= |-\rangle. \end{aligned} \tag{2.121}$$

For the spin-1 case, the connection to this new notation is

$$\begin{aligned} |11\rangle &= |1\rangle \\ |10\rangle &= |0\rangle \\ |1,-1\rangle &= |-1\rangle. \end{aligned} \tag{2.122}$$

We will continue to use the $|\pm\rangle$ notation, but we will find the new notation useful later (Chapter 7).

2.8 ■ GENERAL QUANTUM SYSTEMS

Let's extend the important results of this chapter to general quantum mechanical systems. For a general observable A with quantized measurement results a_n, the eigenvalue equation is

$$A|a_n\rangle = a_n|a_n\rangle. \tag{2.123}$$

In the basis formed by the eigenstates $|a_n\rangle$, the operator A is represented by a matrix with the eigenvalues along the diagonal

$$A \doteq \begin{pmatrix} a_1 & 0 & 0 & \cdots \\ 0 & a_2 & 0 & \cdots \\ 0 & 0 & a_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.124}$$

whose size depends on the dimensionality of the system. In this same basis, the eigenstates are represented by the column vectors

$$|a_1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \quad |a_2\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \quad |a_3\rangle \doteq \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, \ldots . \tag{2.125}$$

The projection operators corresponding to measurement of the eigenvalues a_n are

$$P_{a_n} = |a_n\rangle\langle a_n|. \tag{2.126}$$

The completeness of the basis states is expressed by saying that the sum of the projection operators is the identity operator

$$\sum_n P_{a_n} = \sum_n |a_n\rangle\langle a_n| = \mathbf{1}. \tag{2.127}$$

SUMMARY

In this chapter we have extended the mathematical description of quantum mechanics by using operators to represent physical observables. The only possible results of measurements are the eigenvalues of operators. The eigenvectors of the operator are the basis states corresponding to each possible eigenvalue. We find the eigenvalues and eigenvectors by diagonalizing the matrix representing the operator, which allows us to predict the results of measurements. The eigenvalue equations for the spin-1/2 component operator S_z are

$$\begin{aligned} S_z|+\rangle &= +\frac{\hbar}{2}|+\rangle \\ S_z|-\rangle &= -\frac{\hbar}{2}|-\rangle. \end{aligned} \tag{2.128}$$

The matrices representing the spin-1/2 operators are

$$\begin{aligned} S_x &\doteq \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad & S_y &\doteq \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ S_z &\doteq \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad & \mathbf{S}^2 &\doteq \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.129}$$

We characterized quantum mechanical measurements of an observable A by the expectation value

$$\langle A\rangle = \langle\psi|A|\psi\rangle = \sum_n a_n \mathcal{P}_{a_n} \tag{2.130}$$

and the uncertainty

$$\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2}. \tag{2.131}$$

We made a connection between the commutator $[A, B] = AB - BA$ of two operators and the ability to measure the two observables. If two operators commute, then we can measure both observables simultaneously, but if they do not commute, then we cannot measure them simultaneously. We quantified this disturbance that measurement inflicts on quantum systems through the quantum mechanical uncertainty principle

$$\Delta A \Delta B \geq \tfrac{1}{2}\left|\langle [A,B]\rangle\right|. \tag{2.132}$$

We also introduced the projection postulate, which states how the quantum state vector is changed after a measurement.

PROBLEMS

2.1 Given the following information:

$$S_x|\pm\rangle_x = \pm\frac{\hbar}{2}|\pm\rangle_x \qquad S_y|\pm\rangle_y = \pm\frac{\hbar}{2}|\pm\rangle_y$$

$$|\pm\rangle_x = \tfrac{1}{\sqrt{2}}[|+\rangle \pm |-\rangle] \qquad |\pm\rangle_y = \tfrac{1}{\sqrt{2}}[|+\rangle \pm i|-\rangle]$$

find the matrix representations of S_x and S_y in the S_z basis.

2.2 From the previous problem we know that the matrix representation of S_x in the S_z basis is

$$S_x \doteq \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Diagonalize this matrix to find the eigenvalues and the eigenvectors of S_x.

2.3 Find the matrix representation of S_z in the S_x basis for spin 1/2. Diagonalize this matrix to find the eigenvalues and the eigenvectors in this basis. Show that the eigenvalue equations for S_z are satisfied in this new representation.

2.4 Show by explicit matrix calculation that the matrix elements of a general operator A (within a spin-1/2 system) are as shown in Eq. (2.13).

2.5 Calculate the commutators of the spin-1/2 operators S_x, S_y, and S_z, thus verifying Eqs. (2.96).

2.6 Verify that the spin component operator S_n along the direction $\hat{\mathbf{n}}$ has the matrix representation shown in Eq. (2.41).

2.7 Diagonalize the spin component operator S_n along the direction $\hat{\mathbf{n}}$ to find its eigenvalues and the eigenvectors.

2.8 Find the probabilities of the measurements shown below in Fig. 2.14. The first Stern-Gerlach analyzer is aligned along the direction $\hat{\mathbf{n}}$ defined by the angles $\theta = \pi/4$ and $\phi = 5\pi/3$.

2.9 For the state $|+\rangle$, calculate the expectation values and uncertainties for measurements of S_x, S_y, and S_z in order to verify Eq. (2.108).

2.10 For the state $|+\rangle_y$, calculate the expectation values and uncertainties for measurements of S_x, S_y, and S_z. Draw a diagram of the vector model applied to this state and reconcile your quantum mechanical calculations with the classical results.

2.11 Show that the $\mathbf{S}^2$ operator commutes with each of the spin component operators of S_x, S_y, and S_z. Do this once with matrix notation for a spin-1/2 system and a second time using only the component commutation relations in Eqs. (2.96) and the definition of $\mathbf{S}^2$ in Eq. (2.103).

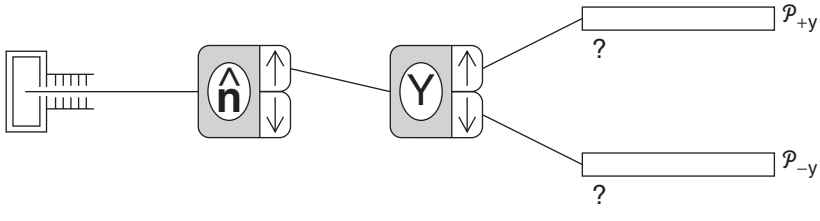


FIGURE 2.14 Measurement of spin components (Prob. 2.8).

2.12 Diagonalize the S_x and S_y operators in the spin-1 case to find the eigenvalues and the eigenvectors of both operators.

2.13 For a spin-1 system, show by explicit matrix calculation that the spin component operators obey the commutation relations in Eqs. (2.96).

2.14 Find the matrix representation of the $\mathbf{S}^2$ operator for a spin-1 system. Do this once by explicit matrix calculation and a second time by inspection of the $\mathbf{S}^2$ eigenvalue equation (2.120).

2.15 A beam of spin-1 particles is prepared in the state

$$|\psi\rangle = \tfrac{2}{\sqrt{29}}|1\rangle + i\tfrac{3}{\sqrt{29}}|0\rangle - \tfrac{4}{\sqrt{29}}|-1\rangle.$$

a) What are the possible results of a measurement of the spin component S_z, and with what probabilities would they occur?

b) What are the possible results of a measurement of the spin component S_x, and with what probabilities would they occur?

c) Plot histograms of the predicted measurement results from parts (a) and (b), and calculate the expectation values for both measurements.

2.16 A beam of spin-1 particles is prepared in the state

$$|\psi\rangle = \tfrac{2}{\sqrt{29}}|1\rangle_y + i\tfrac{3}{\sqrt{29}}|0\rangle_y - \tfrac{4}{\sqrt{29}}|-1\rangle_y.$$

a) What are the possible results of a measurement of the spin component S_z, and with what probabilities would they occur?

b) What are the possible results of a measurement of the spin component S_y, and with what probabilities would they occur?

c) Plot histograms of the predicted measurement results from parts (a) and (b), and calculate the expectation values for both measurements.

2.17 A spin-1 particle is in the state

$$|\psi\rangle \doteq \frac{1}{\sqrt{30}}\begin{pmatrix}1\\2\\5i\end{pmatrix}.$$

a) What are the possible results of a measurement of the spin component S_z, and with what probabilities would they occur? Calculate the expectation value of the spin component S_z.

b) Calculate the expectation value of the spin component S_x. *Suggestion*: Use matrix mechanics to evaluate the expectation value.

2.18 A spin-1 particle is prepared in the state

$$|\psi\rangle = \tfrac{1}{\sqrt{14}}|1\rangle - \tfrac{3}{\sqrt{14}}|0\rangle + i\tfrac{2}{\sqrt{14}}|-1\rangle.$$

a) What are the possible results of a measurement of the spin component S_z, and with what probabilities would they occur?

b) Suppose that the S_z measurement on the particle yields the result $S_z = -\hbar$. Subsequent to that result a second measurement is performed to measure the spin component S_x. What are the possible results of that measurement, and with what probabilities would they occur?

c) Draw a schematic diagram depicting the successive measurements in parts (a) and (b).

2.19 A spin-1 particle is prepared in the state

$$|\psi_i\rangle = \sqrt{\tfrac{1}{6}}|1\rangle - \sqrt{\tfrac{2}{6}}|0\rangle + i\sqrt{\tfrac{3}{6}}|-1\rangle.$$

Find the probability that the system is measured to be in the final state

$$|\psi_f\rangle = \tfrac{1+i}{\sqrt{7}}|1\rangle_y + \tfrac{2}{\sqrt{7}}|0\rangle_y - i\tfrac{1}{\sqrt{7}}|-1\rangle_y.$$

2.20 In part (2) of SPINS Lab #3, you measured the spin components of the unknown (spin 1) initial states $|\psi_i\rangle$ ($i = 1, 2, 3, 4$) along the three axes. Using your measured values, deduce the unknown initial states.

2.21 In part (3) of SPINS Lab #3, you built a spin-1 interferometer and measured the relative probabilities after the final Stern-Gerlach analyzer for the seven possible cases where one beam, a pair of beams, or all three beams from the second Stern-Gerlach analyzer were used. Show how you used the projection postulate to calculate the theoretical probabilities.

2.22 A beam of spin-1/2 particles is sent through a series of three Stern-Gerlach analyzers, as shown in Fig. 2.15. The second Stern-Gerlach analyzer is aligned along the $\hat{\mathbf{n}}$ direction, which makes an angle θ in the x-z plane with respect to the z-axis.

a) Find the probability that particles transmitted through the first Stern-Gerlach analyzer are measured to have spin down at the third Stern-Gerlach analyzer?

b) How must the angle θ of the second Stern-Gerlach analyzer be oriented so as to maximize the probability that particles are measured to have spin down at the third Stern-Gerlach analyzer? What is this maximum fraction?

c) What is the probability that particles have spin down at the third Stern-Gerlach analyzer if the second Stern-Gerlach analyzer is removed from the experiment?

2.23 Consider a three-dimensional ket space. In the basis defined by three orthogonal kets $|1\rangle$, $|2\rangle$, and $|3\rangle$, the operators A and B are represented by

$$A \doteq \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \qquad B \doteq \begin{pmatrix} b_1 & 0 & 0 \\ 0 & 0 & b_2 \\ 0 & b_2 & 0 \end{pmatrix},$$

where all the quantities are real.

a) Do the operators A and B commute?

b) Find the eigenvalues and normalized eigenvectors of both operators.

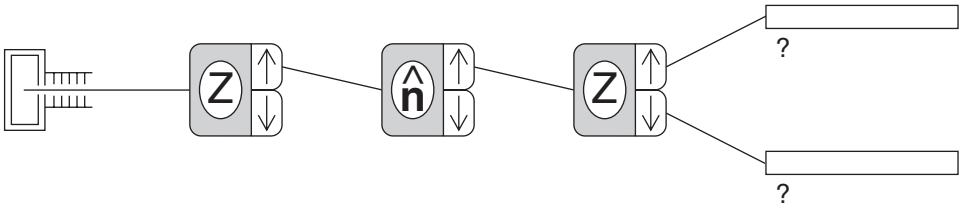


FIGURE 2.15 Measurement of spin components (Prob. 2.22).

c) Assume the system is initially in the state $|2\rangle$. Then the observable corresponding to the operator B is measured. What are the possible results of this measurement and the probabilities of each result? After this measurement, the observable corresponding to the operator A is measured. What are the possible results of this measurement and the probabilities of each result?

d) How are questions (a) and (c) above related?

2.24 If a beam of spin-3/2 particles is input to a Stern-Gerlach analyzer, there are four output beams whose deflections are consistent with magnetic moments arising from spin angular momentum components of $\frac{3}{2}\hbar$, $\frac{1}{2}\hbar$, $-\frac{1}{2}\hbar$, and $-\frac{3}{2}\hbar$. For a spin-3/2 system:

a) Write down the eigenvalue equations for the S_z operator.

b) Write down the matrix representation of the S_z eigenstates.

c) Write down the matrix representation of the S_z operator.

d) Write down the eigenvalue equations for the S^2 operator.

e) Write down the matrix representation of the S^2 operator.

2.25 Are the projection operators P_+ and P_- Hermitian? Explain.

RESOURCES

Activities

This activity is available at

www.physics.oregonstate.edu/qmactivities

Spins Lab 3: Stern-Gerlach measurements of a spin-1 system.

CHAPTER

3 Schrödinger Time Evolution

This chapter marks our final step in developing the mathematical basis of a quantum theory. In Chapter 1, we learned how to use kets to describe quantum states and how to predict the probabilities of results of measurements. In Chapter 2, we learned how to use operators to represent physical observables and how to determine the possible measurement results. The key missing aspect is the ability to predict the future. Physics theories are judged on their predictive power. Classical mechanics relies on Newton's second law $\mathbf{F} = m\mathbf{a}$ to predict the future of a particle's motion. The ability to predict the quantum future started with Erwin Schrödinger and bears his name.

3.1 ■ SCHRÖDINGER EQUATION

The sixth postulate of quantum mechanics says that the time evolution of a quantum system is governed by the differential equation

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle, \tag{3.1}$$

where the operator H corresponds to the total energy of the system and is called the **Hamiltonian** operator of the system because it is derived from the classical Hamiltonian. This equation is known as the **Schrödinger equation**.

> **Postulate 6**
>
> The time evolution of a quantum system is determined by the Hamiltonian or total energy operator $H(t)$ through the Schrödinger equation
>
> $$i\hbar \frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle.$$

The Hamiltonian is a new operator, but we can use the same ideas we developed in Chapter 2 to understand its basic properties. The Hamiltonian H is an observable, so it is a Hermitian operator. The eigenvalues of the Hamiltonian are the allowed energies of the quantum system, and the eigenstates of H are the energy eigenstates of the system. If we label the allowed energies as E_n, then the **energy eigenvalue equation** is

$$\boxed{H|E_n\rangle = E_n|E_n\rangle}. \tag{3.2}$$

If we have the Hamiltonian H in a matrix representation, then we diagonalize the matrix to find the eigenvalues E_n and the eigenvectors $|E_n\rangle$ just as we did with the spin operators in Chapter 2. For the moment, let's assume that we have already diagonalized the Hamiltonian [i.e., solved Eq. (3.2)] so that we know the eigenvalues E_n and the eigenvectors $|E_n\rangle$, and let's see what we can learn about quantum time evolution in general by solving the Schrödinger equation.

The eigenvectors of the Hamiltonian form a complete basis because the Hamiltonian is an observable, and therefore a Hermitian operator. Because H is the only operator appearing in the Schrödinger equation, it would seem reasonable (and will prove invaluable) to consider the energy eigenstates as the basis of choice for expanding general state vectors:

$$|\psi(t)\rangle = \sum_n c_n(t)|E_n\rangle. \tag{3.3}$$

The basis of eigenvectors of the Hamiltonian is also orthonormal, so

$$\langle E_k|E_n\rangle = \delta_{kn}. \tag{3.4}$$

We refer to this basis as the **energy basis**.

For now, we assume that the Hamiltonian is time independent (we will do the time-dependent case $H(t)$ in Section 3.4). The eigenvectors of a time-independent Hamiltonian come from the diagonalization procedure we used in Chapter 2, so there is no reason to expect the eigenvectors themselves to carry any time dependence. Thus if a general state $|\psi\rangle$ is to be time dependent, as the Schrödinger equation implies, then the time dependence must reside in the expansion coefficients $c_n(t)$, as expressed in Eq. (3.3). Substitute this general state into the Schrödinger equation (3.1)

$$i\hbar\frac{d}{dt}\sum_n c_n(t)|E_n\rangle = H\sum_n c_n(t)|E_n\rangle \tag{3.5}$$

and use the energy eigenvalue equation (3.2) to obtain

$$i\hbar\sum_n \frac{dc_n(t)}{dt}|E_n\rangle = \sum_n c_n(t)E_n|E_n\rangle. \tag{3.6}$$

Each side of this equation is a sum over all the energy states of the system. To simplify this equation, we isolate single terms in these two sums by taking the inner product of the ket on each side with one particular ket $|E_k\rangle$ (this ket can have any label k, but must not have the label n that is already used in the summation). The orthonormality condition $\langle E_k|E_n\rangle = \delta_{kn}$ then collapses the sums:

$$\begin{aligned}
\langle E_k|i\hbar\sum_n \frac{dc_n(t)}{dt}|E_n\rangle &= \langle E_k|\sum_n c_n(t)E_n|E_n\rangle \\
i\hbar\sum_n \frac{dc_n(t)}{dt}\langle E_k|E_n\rangle &= \sum_n c_n(t)E_n\langle E_k|E_n\rangle \\
i\hbar\sum_n \frac{dc_n(t)}{dt}\delta_{kn} &= \sum_n c_n(t)E_n\delta_{kn} \\
i\hbar\frac{dc_k(t)}{dt} &= c_k(t)E_k.
\end{aligned} \tag{3.7}$$

We are left with a single differential equation for each of the possible energy states of the systems $k = 1, 2, 3, \ldots$. This first-order differential equation can be rewritten as

$$\frac{dc_k(t)}{dt} = -i\frac{E_k}{\hbar}c_k(t). \tag{3.8}$$

The solution to Eq. (3.8) is a complex exponential

$$c_k(t) = c_k(0)e^{-iE_kt/\hbar}. \tag{3.9}$$

In Eq. (3.9), we have denoted the initial condition as $c_k(0)$, but we denote it simply as c_k hereafter. Each coefficient in the energy basis expansion of the state obeys the *same* form of the time dependence in Eq. (3.9), but with a *different* exponent due to the different energies. The time-dependent solution for the full state vector is summarized by saying that if the initial state of the system at time $t = 0$ is

$$|\psi(0)\rangle = \sum_n c_n|E_n\rangle, \tag{3.10}$$

then the time evolution of this state under the action of the time-independent Hamiltonian H is

$$\boxed{|\psi(t)\rangle = \sum_n c_n e^{-iE_nt/\hbar}|E_n\rangle}. \tag{3.11}$$

So the time dependence of the original state vector is found by multiplying *each* energy eigenstate coefficient by its own phase factor $e^{-iE_nt/\hbar}$ that depends on the energy of *that* eigenstate. Note that the factor $E/\hbar$ is an angular frequency, so that the time dependence is of the form $e^{-i\omega t}$, a form commonly found in many areas of physics. It is important to remember that one must use the *energy* eigenstates for the expansion in Eq. (3.10) in order to use the simple phase factor multiplication in Eq. (3.11) to account for the Schrödinger time evolution of the state. This key role of the energy basis accounts for the importance of the Hamiltonian operator and for the common practice of finding the energy eigenstates to use as the preferred basis.

A few examples help to illustrate some of the important consequences of this time evolution of the quantum mechanical state vector. First, consider the simplest possible situation where the system is initially in one particular energy eigenstate:

$$|\psi(0)\rangle = |E_1\rangle, \tag{3.12}$$

for example. The prescription for time evolution tells us that after some time t the system is in the state

$$|\psi(t)\rangle = e^{-iE_1t/\hbar}|E_1\rangle. \tag{3.13}$$

But this state differs from the original state only by an overall phase factor, which we have said before does not affect any measurements (Problem 1.3). For example, if we measure an observable A, then the probability of measuring an eigenvalue a_j is given by

$$\begin{aligned} \mathcal{P}_{a_j} &= |\langle a_j|\psi(t)\rangle|^2 \\ &= |\langle a_j|e^{-iE_1t/\hbar}|E_1\rangle|^2 \\ &= |\langle a_j|E_1\rangle|^2. \end{aligned} \tag{3.14}$$

This probability is time independent and is equal to the probability at the initial time. Thus, we conclude that there is no measureable time evolution for this state. Hence, the energy eigenstates are called **stationary states**. If a system begins in an energy eigenstate, then it remains in that state.

Now consider an initial state that is a superposition of two energy eigenstates:

$$|\psi(0)\rangle = c_1|E_1\rangle + c_2|E_2\rangle. \tag{3.15}$$

In this case, time evolution takes the initial state to the later state

$$|\psi(t)\rangle = c_1e^{-iE_1t/\hbar}|E_1\rangle + c_2e^{-iE_2t/\hbar}|E_2\rangle. \tag{3.16}$$

A measurement of the system energy at the time t would yield the value E_1 with a probability

$$\begin{aligned}\mathcal{P}_{E_1} &= \left|\langle E_1|\psi(t)\rangle\right|^2 \\ &= \left|\langle E_1|\left[c_1 e^{-iE_1 t/\hbar}|E_1\rangle + c_2 e^{-iE_2 t/\hbar}|E_2\rangle\right]\right|^2 \\ &= |c_1|^2,\end{aligned} \tag{3.17}$$

which is independent of time. The same is true for the probability of measuring the energy E_2. Thus, the probabilities of measuring the energies are stationary, as they were in the first example.

However, now consider what happens if another observable is measured on this system in this superposition state. There are two distinct situations: (1) If the other observable A commutes with the Hamiltonian H, then A and H have common eigenstates. In this case, measuring A is equivalent to measuring H because the inner products used to calculate the probabilities use the same eigenstates. Hence, the probability of measuring any particular eigenvalue of A is time independent, as in Eq. (3.17). (2) If A and H do not commute, then they do not share common eigenstates. In this case, the eigenstates of A in general consist of superpositions of energy eigenstates. For example, suppose that the eigenstate of A corresponding to the eigenvalue a_1 were

$$|a_1\rangle = \alpha_1|E_1\rangle + \alpha_2|E_2\rangle. \tag{3.18}$$

Then the probability of measuring the eigenvalue a_1 would be

$$\begin{aligned}\mathcal{P}_{a_1} &= \left|\langle a_1|\psi(t)\rangle\right|^2 \\ &= \left|\left[\alpha_1^*\langle E_1| + \alpha_2^*\langle E_2|\right]\left[c_1 e^{-iE_1 t/\hbar}|E_1\rangle + c_2 e^{-iE_2 t/\hbar}|E_2\rangle\right]\right|^2 \\ &= \left|\alpha_1^* c_1 e^{-iE_1 t/\hbar} + \alpha_2^* c_2 e^{-iE_2 t/\hbar}\right|^2.\end{aligned} \tag{3.19}$$

Factoring out the common phase gives

$$\begin{aligned}\mathcal{P}_{a_1} &= \left|e^{-iE_1 t/\hbar}\right|^2\left|\alpha_1^* c_1 + \alpha_2^* c_2 e^{-i(E_2-E_1)t/\hbar}\right|^2 \\ &= |\alpha_1|^2|c_1|^2 + |\alpha_2|^2|c_2|^2 + 2\mathrm{Re}\left(\alpha_1 c_1^* \alpha_2^* c_2 e^{-i(E_2-E_1)t/\hbar}\right).\end{aligned} \tag{3.20}$$

The different time-evolution phases of the two components of $|\psi(t)\rangle$ lead to a time dependence in the probability. The overall phase in Eq. (3.20) drops out, and only the relative phase remains in the probability calculation. Hence, the time dependence is determined by the *difference* of the energies of the two states involved in the superposition. The corresponding angular frequency of the time evolution

$$\omega_{21} = \frac{E_2 - E_1}{\hbar} \tag{3.21}$$

is called the **Bohr frequency**.

To summarize, we list below a recipe for solving a standard time-dependent quantum mechanics problem with a time-independent Hamiltonian.

> Given a Hamiltonian H and an initial state $|\psi(0)\rangle$, what is the probability that the eigenvalue a_j of the observable A is measured at time t?
>
> 1. Diagonalize H (find the eigenvalues E_n and eigenvectors $|E_n\rangle$).
> 2. Write $|\psi(0)\rangle$ in terms of the energy eigenstates $|E_n\rangle$.
> 3. Multiply each eigenstate coefficient by $e^{-iE_n t/\hbar}$ to get $|\psi(t)\rangle$.
> 4. Calculate the probability $\mathcal{P}_{a_j} = \left|\langle a_j|\psi(t)\rangle\right|^2$.

3.2 ■ SPIN PRECESSION

Now apply this new concept of Schrödinger time evolution to the case of a spin-1/2 system. The Hamiltonian operator represents the total energy of the system, but because only energy differences are important in time-dependent solutions (and because we can define the zero of potential energy as we wish), we need consider only energy terms that differentiate between the two possible spin states in the system. Our experience with the Stern-Gerlach apparatus tells us that the magnetic potential energy of the magnetic dipole differs for the two possible spin-component states. So to begin, we consider the potential energy of a single magnetic dipole (e.g., in a silver atom) in a uniform magnetic field as the sole term in the Hamiltonian. Recalling that the magnetic dipole is given by

$$\boldsymbol{\mu} = g\frac{q}{2m_e}\mathbf{S}, \tag{3.22}$$

the Hamiltonian is

$$\begin{aligned} H &= -\boldsymbol{\mu}\cdot\mathbf{B} \\ &= -g\frac{q}{2m_e}\mathbf{S}\cdot\mathbf{B} \\ &= \frac{e}{m_e}\mathbf{S}\cdot\mathbf{B}, \end{aligned} \tag{3.23}$$

where $q = -e$ and $g = 2$ have been used in the last line. The gyromagnetic ratio, g, is slightly different from 2, but we ignore that detail.

3.2.1 ■ Magnetic Field in the *z*-Direction

For our first example, we assume that the magnetic field is uniform and directed along the z-axis. Writing the magnetic field as

$$\mathbf{B} = B_0\hat{\mathbf{z}} \tag{3.24}$$

allows the Hamiltonian to be simplified to

$$\begin{aligned} H &= \frac{eB_0}{m_e}S_z \\ &= \omega_0 S_z, \end{aligned} \tag{3.25}$$

where we have introduced the definition

$$\omega_0 \equiv \frac{eB_0}{m_e}. \tag{3.26}$$

This definition of an angular frequency simplifies the notation now and will have an obvious interpretation at the end of the problem.

The Hamiltonian in Eq. (3.25) is proportional to the S_z operator, so H and S_z commute and therefore share common eigenstates. This is clear if we write the Hamiltonian as a matrix in the S_z representation:

$$H \doteq \frac{\hbar\omega_0}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.27}$$

Because H is diagonal, we have already completed step 1 of the Schrödinger time-evolution recipe. The eigenstates of H are the basis states of the representation, while the eigenvalues are the diagonal elements of the matrix in Eq. (3.27). The eigenvalue equations for the Hamiltonian are thus

$$\begin{aligned} H|+\rangle &= \omega_0 S_z|+\rangle = \frac{\hbar\omega_0}{2}|+\rangle = E_+|+\rangle \\ H|-\rangle &= \omega_0 S_z|-\rangle = -\frac{\hbar\omega_0}{2}|+\rangle = E_-|-\rangle, \end{aligned} \tag{3.28}$$

with eigenvalues and eigenvectors given by

$$\begin{aligned} E_+ &= \frac{\hbar\omega_0}{2} & E_- &= -\frac{\hbar\omega_0}{2} \\ |E_+\rangle &= |+\rangle & |E_-\rangle &= |-\rangle. \end{aligned} \tag{3.29}$$

The information regarding the energy eigenvalues and eigenvectors is commonly presented in a graphical diagram, which is shown in Fig. 3.1 for this case. The two energy states are separated by the energy $E_+ - E_- = \hbar\omega_0$, so the angular frequency ω_0 characterizes the energy scale of this system. The spin-up state $|+\rangle$ has a higher energy because the magnetic moment is aligned against the field in that state; the negative charge in Eq. (3.22) causes the spin and magnetic moment to be antiparallel.

Now we look at a few examples to illustrate the key features of the behavior of a spin-1/2 system in a uniform magnetic field. First, consider the case where the initial state is spin up along the z-axis:

$$|\psi(0)\rangle = |+\rangle. \tag{3.30}$$

This initial state is already expressed in the energy basis (step 2 of the Schrödinger recipe), so the Schrödinger equation time evolution takes this initial state to the state

$$\begin{aligned} |\psi(t)\rangle &= e^{-iE_+t/\hbar}|+\rangle \\ &= e^{-i\omega_0 t/2}|+\rangle \end{aligned} \tag{3.31}$$

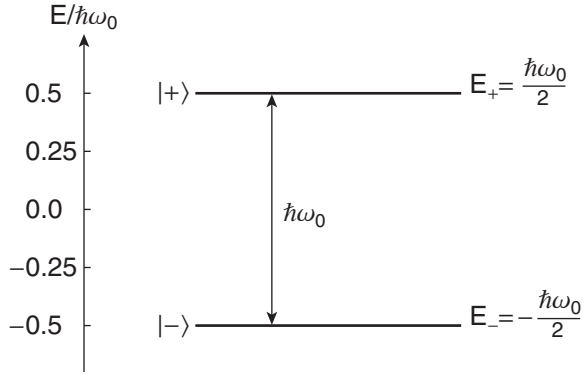


FIGURE 3.1 Energy level diagram of a spin-1/2 particle in a uniform magnetic field.

according to step 3 of the Schrödinger recipe. As we saw before [(Eq. (3.13)], because the initial state is an energy eigenstate, the time-evolved state acquires an overall phase factor, which does not represent a physical change of the state. The probability for measuring the spin to be up along the z-axis is (step 4 of the Schrödinger recipe)

$$\begin{aligned} \mathcal{P}_+ &= |\langle +|\psi(t)\rangle|^2 \\ &= \left|\langle +|e^{-i\omega_0 t/2}|+\rangle\right|^2 \\ &= 1. \end{aligned} \tag{3.32}$$

As expected, this probability is not time dependent, and we therefore refer to $|+\rangle$ as a stationary state for this system. A schematic diagram of this experiment is shown in Fig. 3.2, where we have introduced a new element to represent the applied field. This new depiction is the same as the depictions in the SPINS software, where the number in the applied magnetic field box (42 in Fig. 3.2) is a measure of the magnetic field strength. In this experiment, the results shown are independent of the applied field strength, as indicated by Eq. (3.32), and as you can verify with the software.

Next, consider the most general initial state, which we saw in Chapter 2 corresponds to spin up along an arbitrary direction defined by the polar angle θ and the azimuthal angle ϕ. The initial state is

$$|\psi(0)\rangle = |+\rangle_n = \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{i\phi}|-\rangle, \tag{3.33}$$

or using matrix notation:

$$|\psi(0)\rangle \doteq \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\sin(\theta/2) \end{pmatrix}. \tag{3.34}$$

Schrödinger time evolution introduces a time-dependent phase term for each component, giving

$$\begin{aligned} |\psi(t)\rangle &\doteq \begin{pmatrix} e^{-iE_+ t/\hbar}\cos(\theta/2) \\ e^{-iE_- t/\hbar}e^{i\phi}\sin(\theta/2) \end{pmatrix} \\ &\doteq \begin{pmatrix} e^{-i\omega_0 t/2}\cos(\theta/2) \\ e^{i\omega_0 t/2}e^{i\phi}\sin(\theta/2) \end{pmatrix} \\ &\doteq e^{-i\omega_0 t/2}\begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)}\sin(\theta/2) \end{pmatrix}. \end{aligned} \tag{3.35}$$

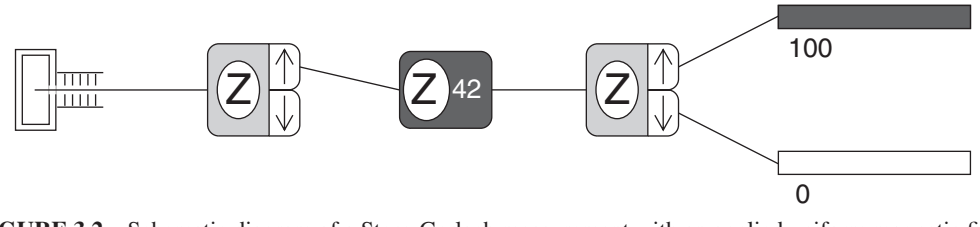


FIGURE 3.2 Schematic diagram of a Stern-Gerlach measurement with an applied uniform magnetic field represented by the box in the middle, with the number 42 representing the strength of the magnetic field.

Note again that an overall phase does not have a measurable effect, so the evolved state is a spin up eigenstate along a direction that has the same polar angle θ as the initial state and a new azimuthal angle $\phi + \omega_0 t$. The state appears to have simply rotated around the z-axis, the axis of the magnetic field, by the angle $\omega_0 t$. Of course, we have to limit our discussion to results of measurements, so let's first calculate the probability for measuring the spin component along the z-axis:

$$\begin{aligned}
\mathcal{P}_+ &= |\langle + | \psi(t) \rangle|^2 \\
&= \left| \begin{pmatrix} 1 & 0 \end{pmatrix} e^{-i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)} \sin(\theta/2) \end{pmatrix} \right|^2 \\
&= |e^{-i\omega_0 t/2} \cos(\theta/2)|^2 \\
&= \cos^2(\theta/2).
\end{aligned} \tag{3.36}$$

This probability is time independent because the S_z eigenstates are also energy eigenstates for this problem (i.e., H and S_z commute). The probability in Eq. (3.36) is consistent with the interpretation that the angle θ that the spin vector makes with the z-axis does not change.

The probability for measuring spin up along the x-axis is

$$\begin{aligned}
\mathcal{P}_{+x} &= |{}_x\langle + | \psi(t) \rangle|^2 \\
&= \left| \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} e^{-i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)} \sin(\theta/2) \end{pmatrix} \right|^2 \\
&= \tfrac{1}{2} \left| \cos(\theta/2) + e^{i(\phi+\omega_0 t)} \sin(\theta/2) \right|^2 \\
&= \tfrac{1}{2} \left[\cos^2(\theta/2) + \cos(\theta/2)\sin(\theta/2) \left(e^{i(\phi+\omega_0 t)} + e^{-i(\phi+\omega_0 t)} \right) + \sin^2(\theta/2) \right] \\
&= \tfrac{1}{2} \left[1 + \sin\theta \cos(\phi + \omega_0 t) \right].
\end{aligned} \tag{3.37}$$

This probability is time dependent because the S_x eigenstates are not stationary states (i.e., H and S_x do not commute). The time dependence in Eq. (3.37) is consistent with the spin precessing around the z-axis.

To illustrate this **spin precession** further, it is useful to calculate the expectation values for each of the spin components. For S_z, we have

$$\begin{aligned}
\langle S_z \rangle &= \langle \psi(t) | S_z | \psi(t) \rangle \\
&= e^{i\omega_0 t/2} \begin{pmatrix} \cos\left(\dfrac{\theta}{2}\right) & e^{-i(\phi+\omega_0 t)} \sin\left(\dfrac{\theta}{2}\right) \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} e^{-i\omega_0 t/2} \begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)} \sin(\theta/2) \end{pmatrix} \\
&= \frac{\hbar}{2} \left[\cos^2(\theta/2) - \sin^2(\theta/2) \right] \\
&= \frac{\hbar}{2} \cos\theta,
\end{aligned} \tag{3.38}$$

while the other components are

$$\begin{aligned}\langle S_y\rangle &= \langle\psi(t)|S_y|\psi(t)\rangle \\ &= e^{i\omega_0 t/2}\begin{pmatrix}\cos\left(\frac{\theta}{2}\right) & e^{-i(\phi+\omega_0 t)}\sin\left(\frac{\theta}{2}\right)\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}e^{-i\omega_0 t/2}\begin{pmatrix}\cos(\theta/2)\\ e^{i(\phi+\omega_0 t)}\sin(\theta/2)\end{pmatrix} \\ &= \frac{\hbar}{2}\sin\theta\sin(\phi+\omega_0 t)\end{aligned} \tag{3.39}$$

and

$$\begin{aligned}\langle S_x\rangle &= \langle\psi(t)|S_x|\psi(t)\rangle \\ &= \frac{\hbar}{2}\sin\theta\cos(\phi+\omega_0 t).\end{aligned} \tag{3.40}$$

The expectation value of the total spin vector $\langle\mathbf{S}\rangle$ is shown in Fig. 3.3, where it is seen to precess around the magnetic field direction with an angular frequency ω_0. The precession of the spin vector is known as **Larmor precession** and the frequency of precession is known as the **Larmor frequency**.

The quantum mechanical Larmor precession is analogous to the classical behavior of a magnetic moment in a uniform magnetic field. A classical magnetic moment $\boldsymbol{\mu}$ experiences a torque $\boldsymbol{\mu}\times\mathbf{B}$ when placed in a magnetic field. If the magnetic moment is associated with an angular momentum $\mathbf{L}$, then we can write

$$\boldsymbol{\mu} = \frac{q}{2m}\mathbf{L}, \tag{3.41}$$

where q and m are the charge and mass, respectively, of the system. The equation of motion for the angular momentum

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\mu}\times\mathbf{B} \tag{3.42}$$

then results in

$$\frac{d\boldsymbol{\mu}}{dt} = \frac{q}{2m}\boldsymbol{\mu}\times\mathbf{B}. \tag{3.43}$$

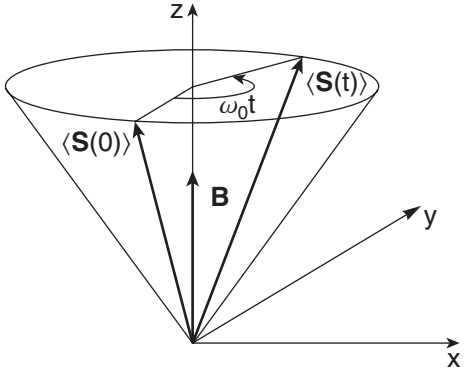


FIGURE 3.3 The expectation value of the spin vector precesses in a uniform magnetic field.