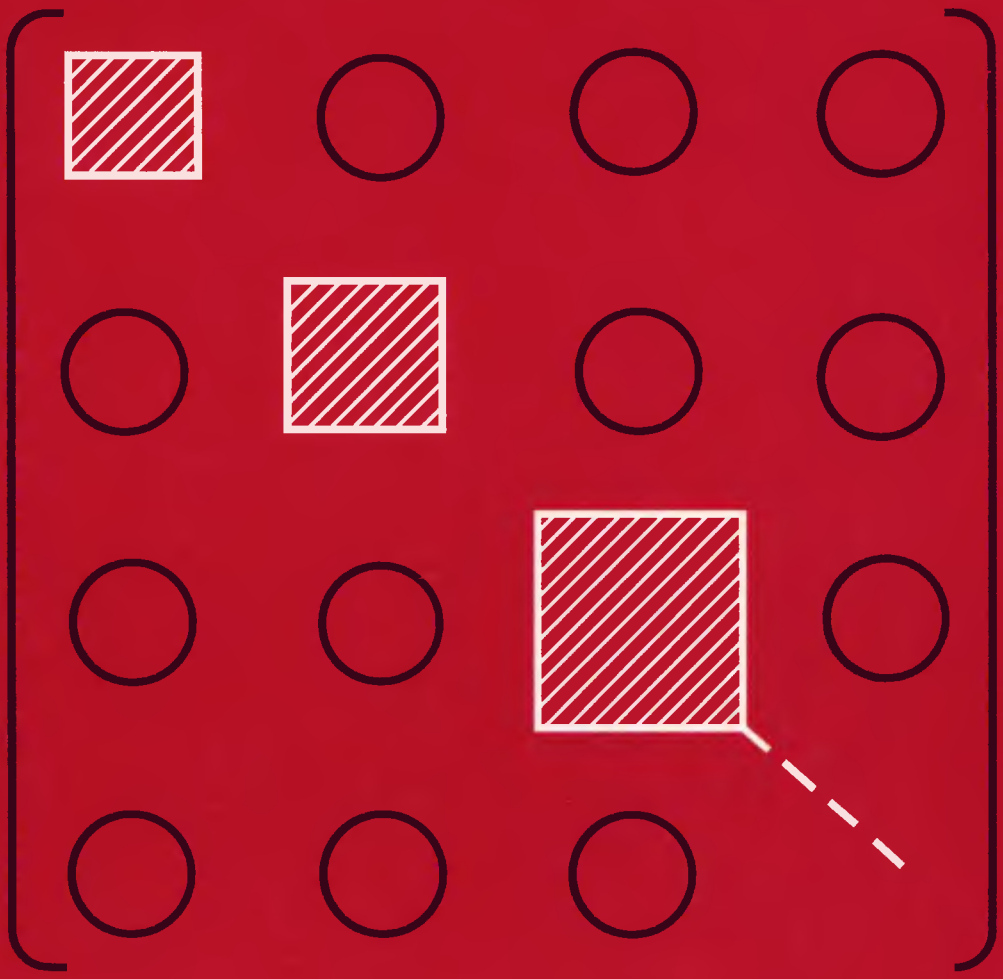
Modern Quantum Mechanics

J. J. Sakurai



Revised Edition

Modern Quantum Mechanics Revised Edition

J. J. Sakurai
Late, University of California, Los Angeles

San Fu Tuan, Editor
University of Hawaii, Manoa

Addison-Wesley Publishing Company
Reading, Massachusetts • Menlo Park, California • New York
Don Mills, Ontario • Wokingham, England • Amsterdam • Bonn
Sydney • Singapore • Tokyo • Madrid • San Juan • Milan • Paris

Sponsoring Editor: Stuart W. Johnson
Assistant Editor: Jennifer Duggan
Senior Production Coordinator: Amy Willcutt
Manufacturing Manager: Roy Logan

Library of Congress Cataloging-in-Publication Data

Sakurai, J. J. (Jun John), 1933–1982.
Modern quantum mechanics / J. J. Sakurai ; San Fu Tuan, editor.—
Rev. ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-201-53929-2
1. Quantum theory. I. Tuan, San Fu, 1932– . II. Title.
QC174.12.S25 1994
530.1'2—dc20 93-17803
CIP

5 6 7 8 9 10–MA–9695

Foreword

J. J. Sakurai was always a very welcome guest here at CERN, for he was one of those rare theorists to whom the experimental facts are even more interesting than the theoretical game itself. Nevertheless, he delighted in theoretical physics and in its teaching, a subject on which he held strong opinions. He thought that much theoretical physics teaching was both too narrow and too remote from application: "...we see a number of sophisticated, yet uneducated, theoreticians who are conversant in the LSZ formalism of the Heisenberg field operators, but do not know why an excited atom radiates, or are ignorant of the quantum theoretic derivation of Rayleigh's law that accounts for the blueness of the sky." And he insisted that the student must be able to use what has been taught: "The reader who has read the book but cannot do the exercises has learned nothing."

He put these principles to work in his fine book *Advanced Quantum Mechanics* (1967) and in *Invariance Principles and Elementary Particles* (1964), both of which have been very much used in the CERN library. This new book, *Modern Quantum Mechanics*, should be used even more, by a larger and less specialized group. The book combines breadth of interest with a thorough practicality. Its readers will find here what they need to know, with a sustained and successful effort to make it intelligible.

J. J. Sakurai's sudden death on November 1, 1982 left this book unfinished. Reinhold Bertlmann and I helped Mrs. Sakurai sort out her husband's papers at CERN. Among them we found a rough, handwritten version of most of the book and a large collection of exercises. Though only three chapters had been completely finished, it was clear that the bulk of the creative work had been done. It was also clear that much work remained to fill in gaps, polish the writing, and put the manuscript in order.

That the book is now finished is due to the determination of Noriko Sakurai and the dedication of San Fu Tuan. Upon her husband's death, Mrs. Sakurai resolved immediately that his last effort should not go to waste. With great courage and dignity she became the driving force behind the project, overcoming all obstacles and setting the high standards to be maintained. San Fu Tuan willingly gave his time and energy to the editing and completion of Sakurai's work. Perhaps only others close to the hectic field of high-energy theoretical physics can fully appreciate the sacrifice involved.

For me personally, J. J. had long been far more than just a particularly distinguished colleague. It saddens me that we will never again laugh together at physics and physicists and life in general, and that he will not see the success of his last work. But I am happy that it has been brought to fruition.

John S. Bell
CERN, Geneva

Preface to the Revised Edition

Since 1989 the Editor has enthusiastically pursued a revised edition of *Modern Quantum Mechanics* by his late great friend J. J. Sakurai, in order to extend this text's usefulness into the twenty-first century. Much consultation took place with the panel of Sakurai friends who helped with the original edition, but in particular with Professor Yasuo Hara of Tsukuba University and Professor Akio Sakurai of Kyoto Sangyo University in Japan.

The major motivation for this project is to revise the main text. There are three important additions and/or changes to the revised edition, which otherwise preserves the original version unchanged. These include a reworking of certain portions of Section 5.2 on time-independent perturbation theory for the degenerate case by Professor Kenneth Johnson of M.I.T., taking into account a subtle point that has not been properly treated by a number of texts on quantum mechanics in this country. Professor Roger Newton of Indiana University contributed refinements on lifetime broadening in Stark effect, additional explanations of phase shifts at resonances, the optical theorem, and on non-normalizable state. These appear as "remarks by the editor" or "editor's note" in the revised edition. Professor Thomas Fulton of the Johns Hopkins University reworked his Coulomb Scattering contribution (Section 7.13) so that it now appears as a shorter text portion emphasizing the physics, with the mathematical details relegated to Appendix C.

Though not a major part of the text, some additions were deemed necessary to take into account developments in quantum mechanics that have become prominent since November 1, 1982. To this end, two supplements are included at the end of the text. Supplement I is on adiabatic change and geometrical phase (popularized by M. V. Berry since 1983) and is actually an English translation of the supplement on this subject written by Professor Akio Sakurai for the Japanese version of *Modern Quantum Mechanics* (copyright © Yoshioka-Shoten Publishing of Kyoto). Supplement II is on non-exponential decays written by my colleague here, Professor Xerxes Tata, and read over by Professor E. C. G. Sudarshan of the University of Texas at Austin. Though non-exponential decays have a long history theoretically, experimental work on transition rates that tests indirectly such decays was done only in 1990. Introduction of additional material is of course a subjective matter on the part of the Editor; the readers will evaluate for themselves its appropriateness. Thanks to Professor Akio Sakurai, the revised edition has been "finely toothcombed" for misprint errors of the first ten printings of the original edition. My colleague, Professor Sandip Pakvasa, provided overall guidance and encouragement to me throughout this process of revision.

In addition to the acknowledgments above, my former students Li Ping, Shi Xiaohong, and Yasunaga Suzuki provided the sounding board for ideas on the revised edition when taking my graduate quantum mechanics course at the University of Hawaii during the spring of 1992. Suzuki provided the initial translation from Japanese of Supplement I as a course term paper. Dr. Andy Acker provided me with computer graphic assistance. The Department of Physics and Astronomy and particularly the High Energy Physics Group of the University of Hawaii at Manoa provided again both the facilities and a conducive atmosphere for me to carry out my editorial task. Finally I wish to express my gratitude to Physics (and sponsoring) Senior Editor, Stuart Johnson, and his Editorial Assistant, Jennifer Duggan, as well as Senior Production Coordinator Amy Willcutt, of Addison-Wesley for their encouragement and optimism that the revised edition will indeed materialize.

San Fu TUAN
Honolulu, Hawaii

J. J. Sakurai 1933–1982

In Memoriam

Jun John Sakurai was born in 1933 in Tokyo and came to the United States as a high school student in 1949. He studied at Harvard and at Cornell, where he received his Ph.D. in 1958. He was then appointed assistant professor of Physics at the University of Chicago, and became a full professor in 1964. He stayed at Chicago until 1970 when he moved to the University of California at Los Angeles, where he remained until his death. During his lifetime he wrote 119 articles in theoretical physics of elementary particles as well as several books and monographs on both quantum and particle theory.

The discipline of theoretical physics has as its principal aim the formulation of theoretical descriptions of the physical world that are at once concise and comprehensive. Because nature is subtle and complex, the pursuit of theoretical physics requires bold and enthusiastic ventures to the frontiers of newly discovered phenomena. This is an area in which Sakurai reigned supreme with his uncanny physical insight and intuition and also his ability to explain these phenomena in illuminating physical terms to the unsophisticated. One has but to read his very lucid textbooks on *Invariance Principles and Elementary Particles* and *Advanced Quantum Mechanics* as well as his reviews and summer school lectures to appreciate this. Without exaggeration I could say that much of what I did understand in particle physics came from these and from his articles and private tutoring.

When Sakurai was still a graduate student, he proposed what is now known as the V-A theory of weak interactions, independently of (and simultaneously with) Richard Feynman, Murray Gell-Mann, Robert Marshak, and George Sudarshan. In 1960 he published in *Annals of Physics* a prophetic paper, probably his single most important one. It was concerned with the first serious attempt to construct a theory of strong interactions based on Abelian and non-Abelian (Yang-Mills) gauge invariance. This seminal work induced theorists to attempt an understanding of the mechanisms of mass generation for gauge (vector) fields, now realized as the Higgs mechanism. Above all it stimulated the search for a realistic unification of forces under the gauge principle, now crowned with success in the celebrated Glashow-Weinberg-Salam unification of weak and electromagnetic forces. On the phenomenological side, Sakurai pursued and vigorously advocated the vector mesons dominance model of hadron dynamics. He was the first to discuss the mixing of ω and ϕ meson states. Indeed, he made numerous important contributions to particle physics phenomenology in a

much more general sense, as his heart was always close to experimental activities.

I knew Jun John for more than 25 years, and I had the greatest admiration not only for his immense powers as a theoretical physicist but also for the warmth and generosity of his spirit. Though a graduate student himself at Cornell during 1957–1958, he took time from his own pioneering research in K-nucleon dispersion relations to help me (via extensive correspondence) with my Ph.D. thesis on the same subject at Berkeley. Both Sandip Pakvasa and I were privileged to be associated with one of his last papers on weak couplings of heavy quarks, which displayed once more his infectious and intuitive style of doing physics. It is of course gratifying to us in retrospect that Jun John counted this paper among the score of his published works that he particularly enjoyed.

The physics community suffered a great loss at Jun John Sakurai's death. The personal sense of loss is a severe one for me. Hence I am profoundly thankful for the opportunity to edit and complete his manuscript on *Modern Quantum Mechanics* for publication. In my faith no greater gift can be given me than an opportunity to show my respect and love for Jun John through meaningful service.

San Fu Tuan

Contents

Modern Quantum Mechanics

CHAPTER 1

Fundamental Concepts

The revolutionary change in our understanding of microscopic phenomena that took place during the first 27 years of the twentieth century is unprecedented in the history of natural sciences. Not only did we witness severe limitations in the validity of classical physics, but we found the alternative theory that replaced the classical physical theories to be far richer in scope and far richer in its range of applicability.

The most traditional way to begin a study of quantum mechanics is to follow the historical developments—Planck's radiation law, the Einstein-Debye theory of specific heats, the Bohr atom, de Broglie's matter waves, and so forth—together with careful analyses of some key experiments such as the Compton effect, the Franck-Hertz experiment, and the Davisson-Germer-Thompson experiment. In that way we may come to appreciate how the physicists in the first quarter of the twentieth century were forced to abandon, little by little, the cherished concepts of classical physics and how, despite earlier false starts and wrong turns, the great masters—Heisenberg, Schrödinger, and Dirac, among others—finally succeeded in formulating quantum mechanics as we know it today.

However, we do not follow the historical approach in this book. Instead, we start with an example that illustrates, perhaps more than any other example, the inadequacy of classical concepts in a fundamental way. We hope that by exposing the reader to a "shock treatment" at the onset, he

or she may be attuned to what we might call the "quantum-mechanical way of thinking" at a very early stage.

1.1. THE STERN-GERLACH EXPERIMENT

The example we concentrate on in this section is the Stern-Gerlach experiment, originally conceived by O. Stern in 1921 and carried out in Frankfurt by him in collaboration with W. Gerlach in 1922. This experiment illustrates in a dramatic manner the necessity for a radical departure from the concepts of classical mechanics. In the subsequent sections the basic formalism of quantum mechanics is presented in a somewhat axiomatic manner but always with the example of the Stern-Gerlach experiment in the back of our minds. In a certain sense, a two-state system of the Stern-Gerlach type is the least classical, most quantum-mechanical system. A solid understanding of problems involving two-state systems will turn out to be rewarding to any serious student of quantum mechanics. It is for this reason that we refer repeatedly to two-state problems throughout this book.

Description of the Experiment

We now present a brief discussion of the Stern-Gerlach experiment, which is discussed in almost any book on modern physics.* First, silver (Ag) atoms are heated in an oven. The oven has a small hole through which some of the silver atoms escape. As shown in Figure 1.1, the beam goes through a collimator and is then subjected to an inhomogeneous magnetic field produced by a pair of pole pieces, one of which has a very sharp edge.

We must now work out the effect of the magnetic field on the silver atoms. For our purpose the following oversimplified model of the silver atom suffices. The silver atom is made up of a nucleus and 47 electrons, where 46 out of the 47 electrons can be visualized as forming a spherically symmetrical electron cloud with no net angular momentum. If we ignore the nuclear spin, which is irrelevant to our discussion, we see that the atom as a whole does have an angular momentum, which is due solely to the spin—intrinsic as opposed to orbital—angular momentum of the single 47th ($5s$) electron. The 47 electrons are attached to the nucleus, which is $\sim 2\times 10^5$ times heavier than the electron; as a result, the heavy atom as a whole possesses a magnetic moment equal to the spin magnetic moment of the 47th electron. In other words, the magnetic moment $\boldsymbol{\mu}$ of the atom is

*For an elementary but enlightening discussion of the Stern-Gerlach experiment, see French and Taylor (1978, 432–38).

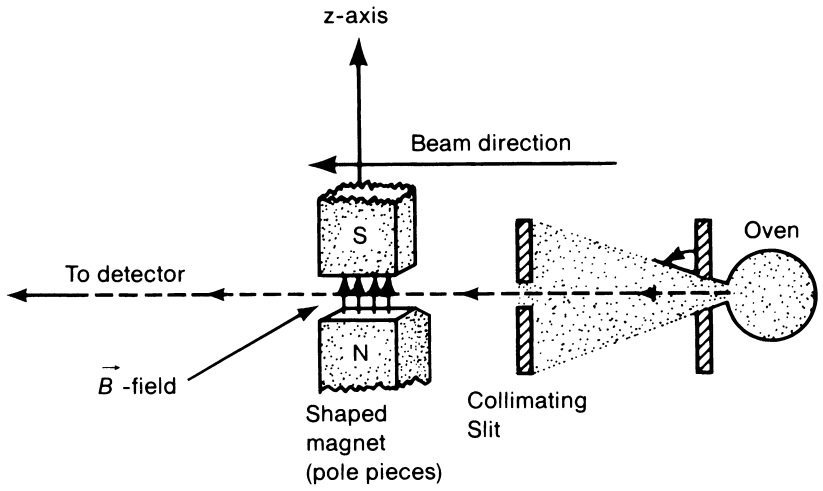


FIGURE 1.1. The Stern-Gerlach experiment.

proportional to the electron spin **S**,

$$\boldsymbol{\mu} \propto \mathbf{S}, \tag{1.1.1}$$

where the precise proportionality factor turns out to be $e/m_e c$ ($e < 0$ in this book) to an accuracy of about 0.2%.

Because the interaction energy of the magnetic moment with the magnetic field is just $-\boldsymbol{\mu}\cdot\mathbf{B}$, the z-component of the force experienced by the atom is given by

$$F_z = \frac{\partial}{\partial z}(\boldsymbol{\mu}\cdot\mathbf{B}) \simeq \mu_z \frac{\partial B_z}{\partial z}, \tag{1.1.2}$$

where we have ignored the components of **B** in directions other than the z-direction. Because the atom as a whole is very heavy, we expect that the classical concept of trajectory can be legitimately applied, a point which can be justified using the Heisenberg uncertainty principle to be derived later. With the arrangement of Figure 1.1, the $\mu_z > 0$ ($S_z < 0$) atom experiences a downward force, while the $\mu_z < 0$ ($S_z > 0$) atom experiences an upward force. The beam is then expected to get split according to the values of μ_z. In other words, the SG (Stern-Gerlach) apparatus "measures" the z-component of $\boldsymbol{\mu}$ or, equivalently, the z-component of **S** up to a proportionality factor.

The atoms in the oven are randomly oriented; there is no preferred direction for the orientation of $\boldsymbol{\mu}$. If the electron were like a classical spinning object, we would expect all values of μ_z to be realized between $|\boldsymbol{\mu}|$ and $-|\boldsymbol{\mu}|$. This would lead us to expect a continuous bundle of beams coming out of the SG apparatus, as shown in Figure 1.2a. Instead, what we

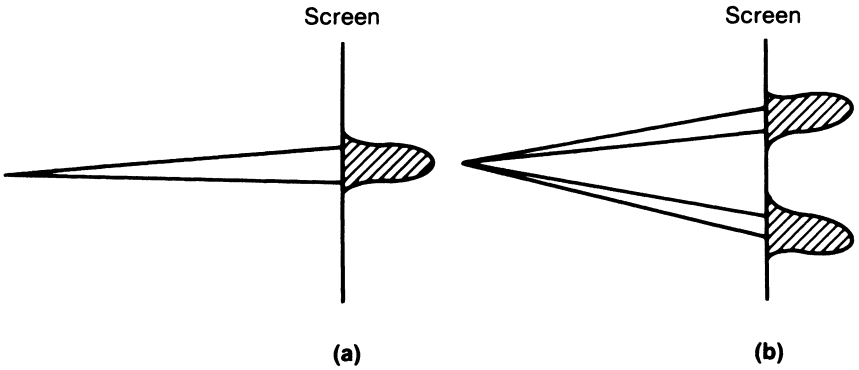


FIGURE 1.2. Beams from the SG apparatus; (a) is expected from classical physics, while (b) is actually observed.

experimentally observe is more like the situation in Figure 1.2b. In other words, the SG apparatus splits the original silver beam from the oven into *two distinct* components, a phenomenon referred to in the early days of quantum theory as "space quantization." To the extent that μ can be identified within a proportionality factor with the electron spin **S**, only two possible values of the z-component of **S** are observed to be possible, S_z up and S_z down, which we call S_z+ and S_z-. The two possible values of S_z are multiples of some fundamental unit of angular momentum; numerically it turns out that $S_z = \hbar/2$ and $-\hbar/2$, where

$$\begin{aligned}\hbar &= 1.0546\times10^{-27}\,\text{erg-s}\\ &= 6.5822\times10^{-16}\,\text{eV-s}\end{aligned} \tag{1.1.3}$$

This "quantization" of the electron spin angular momentum is the first important feature we deduce from the Stern-Gerlach experiment.

Of course, there is nothing sacred about the up-down direction or the z-axis. We could just as well have applied an inhomogeneous field in a horizontal direction, say in the x-direction, with the beam proceeding in the y-direction. In this manner we could have separated the beam from the oven into an S_x+ component and an S_x- component.

Sequential Stern-Gerlach Experiments

Let us now consider a sequential Stern-Gerlach experiment. By this we mean that the atomic beam goes through two or more SG apparatuses in sequence. The first arrangement we consider is relatively straightforward. We subject the beam coming out of the oven to the arrangement shown in Figure 1.3a, where SG$\hat{\mathbf{z}}$ stands for an apparatus with the inhomogeneous magnetic field in the z-direction, as usual. We then block the S_z- compo-

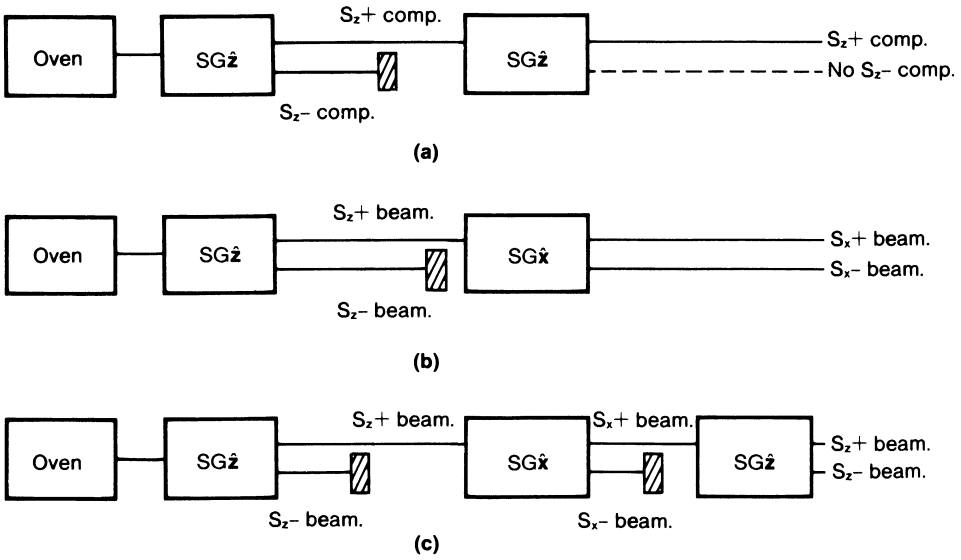


FIGURE 1.3. Sequential Stern-Gerlach experiments.

nent coming out of the first SG$\hat{\mathbf{z}}$ apparatus and let the remaining S_z+ component be subjected to another SG$\hat{\mathbf{z}}$ apparatus. This time there is only one beam component coming out of the second apparatus—just the S_z+ component. This is perhaps not so surprising; after all if the atom spins are up, they are expected to remain so, short of any external field that rotates the spins between the first and the second SG$\hat{\mathbf{z}}$ apparatuses.

A little more interesting is the arrangement shown in Figure 1.3b. Here the first SG apparatus is the same as before but the second one (SG$\hat{\mathbf{x}}$) has an inhomogeneous magnetic field in the x-direction. The S_z+ beam that enters the second apparatus (SG$\hat{\mathbf{x}}$) is now split into two components, an S_x+ component and an S_x- component, with equal intensities. How can we explain this? Does it mean that 50% of the atoms in the S_z+ beam coming out of the first apparatus (SG$\hat{\mathbf{z}}$) are made up of atoms characterized by both S_z+ and S_x+, while the remaining 50% have both S_z+ and S_x-? It turns out that such a picture runs into difficulty, as will be shown below.

We now consider a third step, the arrangement shown in Figure 1.3(c), which most dramatically illustrates the peculiarities of quantum-mechanical systems. This time we add to the arrangement of Figure 1.3b yet a third apparatus, of the SG$\hat{\mathbf{z}}$ type. It is observed experimentally that *two* components emerge from the third apparatus, not one; the emerging beams are seen to have *both* an S_z+ component and an S_z- component. This is a complete surprise because after the atoms emerged from the first

apparatus, we made sure that the $S_z -$ component was completely blocked. How is it possible that the $S_z -$ component which, we thought, we eliminated earlier reappears? The model in which the atoms entering the third apparatus are visualized to have both $S_z +$ and $S_x +$ is clearly unsatisfactory.

This example is often used to illustrate that in quantum mechanics we cannot determine both S_z and S_x simultaneously. More precisely, we can say that the selection of the $S_x +$ beam by the second apparatus (SG$\hat{\mathbf{x}}$) completely destroys any *previous* information about S_z.

It is amusing to compare this situation with that of a spinning top in classical mechanics, where the angular momentum

$$\mathbf{L} = I\boldsymbol{\omega} \tag{1.1.4}$$

can be measured by determining the components of the angular-velocity vector $\boldsymbol{\omega}$. By observing how fast the object is spinning in which direction we can determine ω_x, ω_y, and ω_z simultaneously. The moment of inertia I is computable if we know the mass density and the geometric shape of the spinning top, so there is no difficulty in specifying both L_z and L_x in this classical situation.

It is to be clearly understood that the limitation we have encountered in determining S_z and S_x is not due to the incompetence of the experimentalist. By improving the experimental techniques we cannot make the $S_z -$ component out of the third apparatus in Figure 1.3c disappear. The peculiarities of quantum mechanics are imposed upon us by the experiment itself. The limitation is, in fact, inherent in microscopic phenomena.

Analogy with Polarization of Light

Because this situation looks so novel, some analogy with a familiar classical situation may be helpful here. To this end we now digress to consider the polarization of light waves.

Consider a monochromatic light wave propagating in the z-direction. A linearly polarized (or plane polarized) light with a polarization vector in the x-direction, which we call for short an *x-polarized light*, has a space-time dependent electric field oscillating in the x-direction

$$\mathbf{E} = E_0\hat{\mathbf{x}}\cos(kz - \omega t). \tag{1.1.5}$$

Likewise, we may consider a y-polarized light, also propagating in the z-direction,

$$\mathbf{E} = E_0\hat{\mathbf{y}}\cos(kz - \omega t). \tag{1.1.6}$$

Polarized light beams of type (1.1.5) or (1.1.6) can be obtained by letting an unpolarized light beam go through a Polaroid filter. We call a filter that selects only beams polarized in the x-direction an *x-filter*. An x-filter, of course, becomes a y-filter when rotated by 90° about the propagation (z)

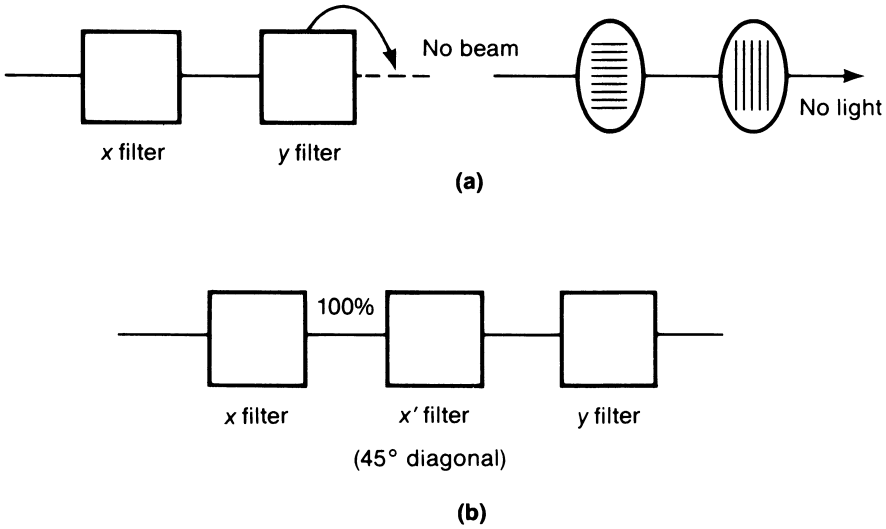


FIGURE 1.4. Light beams subjected to Polaroid filters.

direction. It is well known that when we let a light beam go through an x-filter and subsequently let it impinge on a y-filter, no light beam comes out provided, of course, we are dealing with 100% efficient Polaroids; see Figure 1.4a.

The situation is even more interesting if we insert between the x-filter and the y-filter yet another Polaroid that selects only a beam polarized in the direction—which we call the x'-direction—that makes an angle of 45° with the x-direction in the xy plane; see Figure 1.4b. This time, there is a light beam coming out of the y-filter despite the fact that right after the beam went through the x-filter it did not have any polarization component in the y-direction. In other words, once the x'-filter intervenes and selects the x'-polarized beam, it is immaterial whether the beam was previously x-polarized. The selection of the x'-polarized beam by the second Polaroid destroys any previous information on light polarization. Notice that this situation is quite analogous to the situation that we encountered earlier with the SG arrangement of Figure 1.3b, provided that the following correspondence is made:

$$\begin{aligned} S_z \pm \text{ atoms} &\leftrightarrow x\text{-, } y\text{-polarized light} \\ S_x \pm \text{ atoms} &\leftrightarrow x'\text{-, } y'\text{-polarized light,} \end{aligned} \tag{1.1.7}$$

where the x'- and the y'-axes are defined as in Figure 1.5.

Let us examine how we can quantitatively describe the behavior of 45°-polarized beams (x'- and y'-polarized beams) within the framework of

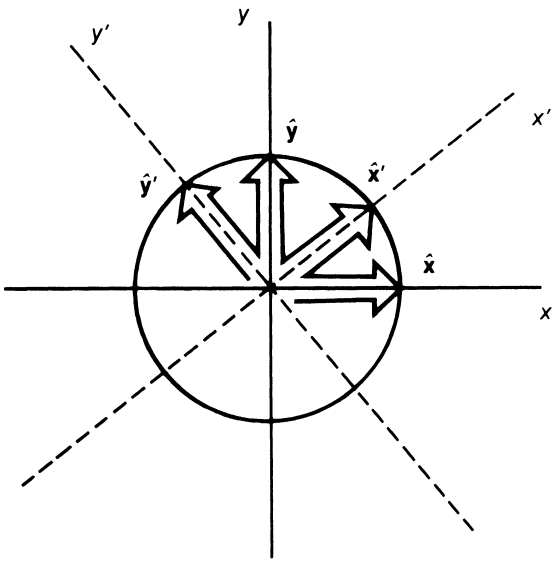


FIGURE 1.5. Orientations of the x'- and y'-axes.

classical electrodynamics. Using Figure 1.5 we obtain

$$E_0\hat{\mathbf{x}}'\cos(kz-\omega t)=E_0\left[\frac{1}{\sqrt{2}}\hat{\mathbf{x}}\cos(kz-\omega t)+\frac{1}{\sqrt{2}}\hat{\mathbf{y}}\cos(kz-\omega t)\right],$$
$$E_0\hat{\mathbf{y}}'\cos(kz-\omega t)=E_0\left[-\frac{1}{\sqrt{2}}\hat{\mathbf{x}}\cos(kz-\omega t)+\frac{1}{\sqrt{2}}\hat{\mathbf{y}}\cos(kz-\omega t)\right]. \tag{1.1.8}$$

In the triple-filter arrangement of Figure 1.4b the beam coming out of the first Polaroid is an $\hat{\mathbf{x}}$-polarized beam, which can be regarded as a linear combination of an x'-polarized beam and a y'-polarized beam. The second Polaroid selects the x'-polarized beam, which can in turn be regarded as a linear combination of an x-polarized and a y-polarized beam. And finally, the third Polaroid selects the y-polarized component.

Applying correspondence (1.1.7) from the sequential Stern-Gerlach experiment of Figure 1.3c, to the triple-filter experiment of Figure 1.4b suggests that we might be able to represent the spin state of a silver atom by some kind of vector in a new kind of two-dimensional vector space, an abstract vector space not to be confused with the usual two-dimensional (xy) space. Just as $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ in (1.1.8) are the base vectors used to decompose the polarization vector $\hat{\mathbf{x}}'$ of the $\hat{\mathbf{x}}'$-polarized light, it is reasonable to represent the S_x+ state by a vector, which we call a *ket* in the Dirac notation to be developed fully in the next section. We denote this vector by

$|S_x; +\rangle$ and write it as a linear combination of two base vectors, $|S_z; +\rangle$ and $|S_z; -\rangle$, which correspond to the S_z+ and the S_z- states, respectively. So we may conjecture

$$|S_x; +\rangle \stackrel{?}{=} \frac{1}{\sqrt{2}}|S_z; +\rangle + \frac{1}{\sqrt{2}}|S_z; -\rangle \tag{1.1.9a}$$

$$|S_x; -\rangle \stackrel{?}{=} -\frac{1}{\sqrt{2}}|S_z; +\rangle + \frac{1}{\sqrt{2}}|S_z; -\rangle \tag{1.1.9b}$$

in analogy with (1.1.8). Later we will show how to obtain these expressions using the general formalism of quantum mechanics.

Thus the unblocked component coming out of the second (SG$\hat{x}$) apparatus of Figure 1.3c is to be regarded as a superposition of S_z+ and S_z- in the sense of (1.1.9a). It is for this reason that two components emerge from the third (SG$\hat{z}$) apparatus.

The next question of immediate concern is, How are we going to represent the $S_y \pm$ states? Symmetry arguments suggest that if we observe an $S_z \pm$ beam going in the x-direction and subject it to an SG$\hat{\mathbf{y}}$ apparatus, the resulting situation will be very similar to the case where an $S_z \pm$ beam going in the y-direction is subjected to an SG$\hat{\mathbf{x}}$ apparatus. The kets for $S_y \pm$ should then be regarded as a linear combination of $|S_z; \pm\rangle$, but it appears from (1.1.9) that we have already used up the available possibilities in writing $|S_x; \pm\rangle$. How can our vector space formalism distinguish $S_y \pm$ states from $S_x \pm$ states?

An analogy with polarized light again rescues us here. This time we consider a circularly polarized beam of light, which can be obtained by letting a linearly polarized light pass through a quarter-wave plate. When we pass such a circularly polarized light through an x-filter or a y-filter, we again obtain either an x-polarized beam or a y-polarized beam of equal intensity. Yet everybody knows that the circularly polarized light is totally different from the 45°-linearly polarized (x'-polarized or y'-polarized) light.

Mathematically, how do we represent a circularly polarized light? A right circularly polarized light is nothing more than a linear combination of an x-polarized light and a y-polarized light, where the oscillation of the electric field for the y-polarized component is 90° out of phase with that of the x-polarized component:*

$$\mathbf{E} = E_0\left[\frac{1}{\sqrt{2}}\hat{\mathbf{x}}\cos(kz - \omega t) + \frac{1}{\sqrt{2}}\hat{\mathbf{y}}\cos\left(kz - \omega t + \frac{\pi}{2}\right)\right]. \tag{1.1.10}$$

It is more elegant to use complex notation by introducing $\boldsymbol{\epsilon}$ as follows:

$$\mathrm{Re}(\boldsymbol{\epsilon}) = \mathbf{E}/\mathbf{E}_0. \tag{1.1.11}$$

*Unfortunately, there is no unanimity in the definition of right versus left circularly polarized light in the literature.

For a right circularly polarized light, we can then write

$$\boldsymbol{\varepsilon} = \left[\frac{1}{\sqrt{2}} \hat{\mathbf{x}} e^{i(kz-\omega t)} + \frac{i}{\sqrt{2}} \hat{\mathbf{y}} e^{i(kz-\omega t)} \right], \tag{1.1.12}$$

where we have used $i = e^{i\pi/2}$.

We can make the following analogy with the spin states of silver atoms:

$$\begin{aligned} S_y + \text{ atom} &\leftrightarrow \text{ right circularly polarized beam,} \\ S_y - \text{ atom} &\leftrightarrow \text{ left circularly polarized beam.} \end{aligned} \tag{1.1.13}$$

Applying this analogy to (1.1.12), we see that if we are allowed to make the coefficients preceding base kets complex, there is no difficulty in accommodating the $S_y \pm$ atoms in our vector space formalism:

$$|S_y; \pm\rangle \stackrel{?}{=} \frac{1}{\sqrt{2}} |S_z; +\rangle \pm \frac{i}{\sqrt{2}} |S_z; -\rangle, \tag{1.1.14}$$

which are obviously different from (1.1.9). We thus see that the two-dimensional vector space needed to describe the spin states of silver atoms must be a *complex* vector space; an arbitrary vector in the vector space is written as a linear combination of the base vectors $|S_z; \pm\rangle$ with, in general, complex coefficients. The fact that the necessity of complex numbers is already apparent in such an elementary example is rather remarkable.

The reader must have noted by this time that we have deliberately avoided talking about photons. In other words, we have completely ignored the quantum aspect of light; nowhere did we mention the polarization states of individual photons. The analogy we worked out is between kets in an abstract vector space that describes the spin states of individual atoms with the polarization vectors of the *classical electromagnetic field*. Actually we could have made the analogy even more vivid by introducing the photon concept and talking about the probability of finding a circularly polarized photon in a linearly polarized state, and so forth; however, that is not needed here. Without doing so, we have already accomplished the main goal of this section: to introduce the idea that quantum-mechanical states are to be represented by vectors in an abstract complex vector space.*

1.2. KETS, BRAS, AND OPERATORS

In the preceding section we showed how analyses of the Stern-Gerlach experiment lead us to consider a complex vector space. In this and the

* The reader who is interested in grasping the basic concepts of quantum mechanics through a careful study of photon polarization may find Chapter 1 of Baym (1969) extremely illuminating.

following section we formulate the basic mathematics of vector spaces as used in quantum mechanics. Our notation throughout this book is the bra and ket notation developed by P. A. M. Dirac. The theory of linear vector spaces had, of course, been known to mathematicians prior to the birth of quantum mechanics, but Dirac's way of introducing vector spaces has many advantages, especially from the physicist's point of view.

Ket Space

We consider a complex vector space whose dimensionality is specified according to the nature of a physical system under consideration. In Stern-Gerlach–type experiments where the only quantum-mechanical degree of freedom is the spin of an atom, the dimensionality is determined by the number of alternative paths the atoms can follow when subjected to a SG apparatus; in the case of the silver atoms of the previous section, the dimensionality is just two, corresponding to the two possible values S_z can assume.* Later, in Section 1.6, we consider the case of continuous spectra—for example, the position (coordinate) or momentum of a particle —where the number of alternatives is nondenumerably infinite, in which case the vector space in question is known as a **Hilbert space** after D. Hilbert, who studied vector spaces in infinite dimensions.

In quantum mechanics a physical state, for example, a silver atom with a definite spin orientation, is represented by a **state vector** in a complex vector space. Following Dirac, we call such a vector a **ket** and denote it by $|\alpha\rangle$. This state ket is postulated to contain complete information about the physical state; everything we are allowed to ask about the state is contained in the ket. Two kets can be added:

$$|\alpha\rangle + |\beta\rangle = |\gamma\rangle. \tag{1.2.1}$$

The sum $|\gamma\rangle$ is just another ket. If we multiply $|\alpha\rangle$ by a complex number c, the resulting product $c|\alpha\rangle$ is another ket. The number c can stand on the left or on the right of a ket; it makes no difference:

$$c|\alpha\rangle = |\alpha\rangle c. \tag{1.2.2}$$

In the particular case where c is zero, the resulting ket is said to be a **null ket**.

One of the physics postulates is that $|\alpha\rangle$ and $c|\alpha\rangle$, with $c \neq 0$, represent the same physical state. In other words, only the "direction" in vector space is of significance. Mathematicians may prefer to say that we are here dealing with rays rather than vectors.

---

*For many physical systems the dimension of the state space is denumerably infinite. While we will usually indicate a finite number of dimensions, N, of the ket space, the results also hold for denumerably infinite dimensions.

An **observable**, such as momentum and spin components, can be represented by an **operator**, such as A, in the vector space in question. Quite generally, an operator acts on a ket *from the left*,

$$A \cdot (|\alpha\rangle) = A|\alpha\rangle, \tag{1.2.3}$$

which is yet another ket. There will be more on multiplication operations later.

In general, $A|\alpha\rangle$ is *not* a constant times $|\alpha\rangle$. However, there are particular kets of importance, known as **eigenkets** of operator A, denoted by

$$|a'\rangle, |a''\rangle, |a'''\rangle, \ldots \tag{1.2.4}$$

with the property

$$A|a'\rangle = a'|a'\rangle, A|a''\rangle = a''|a''\rangle, \ldots \tag{1.2.5}$$

where a', a'', ... are just numbers. Notice that applying A to an eigenket just reproduces the same ket apart from a multiplicative number. The set of numbers $\{a', a'', a''', \ldots\}$, more compactly denoted by $\{a'\}$, is called the set of **eigenvalues** of operator A. When it becomes necessary to order eigenvalues in a specific manner, $\{a^{(1)}, a^{(2)}, a^{(3)}, \ldots\}$ may be used in place of $\{a', a'', a''', \ldots\}$.

The physical state corresponding to an eigenket is called an **eigenstate**. In the simplest case of spin $\frac{1}{2}$ systems, the eigenvalue-eigenket relation (1.2.5) is expressed as

$$S_z|S_z; +\rangle = \frac{\hbar}{2}|S_z; +\rangle, \qquad S_z|S_z; -\rangle = -\frac{\hbar}{2}|S_z; -\rangle, \tag{1.2.6}$$

where $|S_z; \pm\rangle$ are eigenkets of operator S_z with eigenvalues $\pm \hbar/2$. Here we could have used just $|\hbar/2\rangle$ for $|S_z; +\rangle$ in conformity with the notation $|a'\rangle$, where an eigenket is labeled by its eigenvalue, but the notation $|S_z; \pm\rangle$, already used in the previous section, is more convenient here because we also consider eigenkets of S_x:

$$S_x|S_x; \pm\rangle = \pm \frac{\hbar}{2}|S_x; \pm\rangle. \tag{1.2.7}$$

We remarked earlier that the dimensionality of the vector space is determined by the number of alternatives in Stern-Gerlach–type experiments. More formally, we are concerned with an N-dimensional vector space spanned by the N eigenkets of observable A. Any arbitrary ket $|\alpha\rangle$ can be written as

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle, \tag{1.2.8}$$

with $a', a'', \ldots$ up to $a^{(N)}$, where $c_{a'}$ is a complex coefficient. The question of the uniqueness of such an expansion will be postponed until we prove the orthogonality of eigenkets.

Bra Space and Inner Products

The vector space we have been dealing with is a ket space. We now introduce the notion of a **bra space**, a vector space "dual to" the ket space. We postulate that corresponding to every ket $|\alpha\rangle$ there exists a bra, denoted by $\langle\alpha|$, in this dual, or bra, space. The bra space is spanned by eigenbras $\{\langle a'|\}$ which correspond to the eigenkets $\{|a'\rangle\}$. There is a one-to-one correspondence between a ket space and a bra space:

$$\begin{gathered} |\alpha\rangle \overset{\text{DC}}{\leftrightarrow} \langle\alpha| \\ |a'\rangle, |a''\rangle, \ldots \overset{\text{DC}}{\leftrightarrow} \langle a'|, \langle a''|, \ldots \\ |\alpha\rangle + |\beta\rangle \overset{\text{DC}}{\leftrightarrow} \langle\alpha| + \langle\beta| \end{gathered} \tag{1.2.9}$$

where DC stands for **dual correspondence**. Roughly speaking, we can regard the bra space as some kind of mirror image of the ket space.

The bra dual to $c|\alpha\rangle$ is postulated to be $c^*\langle\alpha|$, *not* $c\langle\alpha|$, which is a very important point. More generally, we have

$$c_\alpha|\alpha\rangle + c_\beta|\beta\rangle \overset{\text{DC}}{\leftrightarrow} c_\alpha^*\langle\alpha| + c_\beta^*\langle\beta|. \tag{1.2.10}$$

We now define the **inner product** of a bra and a ket.* The product is written as a bra standing on the left and a ket standing on the right, for example,

$$\langle\beta|\alpha\rangle = (\langle\beta|)\cdot(|\alpha\rangle)\,. \tag{1.2.11}$$
$$\text{bra } (c) \text{ ket}$$

This product is, in general, a complex number. Notice that in forming an inner product we always take one vector from the bra space and one vector from the ket space.

We postulate two fundamental properties of inner products. First,

$$\langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^*. \tag{1.2.12}$$

In other words, $\langle\beta|\alpha\rangle$ and $\langle\alpha|\beta\rangle$ are complex conjugates of each other. Notice that even though the inner product is, in some sense, analogous to the familiar scalar product $\mathbf{a}\cdot\mathbf{b}$, $\langle\beta|\alpha\rangle$ must be clearly distinguished from $\langle\alpha|\beta\rangle$; the analogous distinction is not needed in real vector space because $\mathbf{a}\cdot\mathbf{b}$ is equal to $\mathbf{b}\cdot\mathbf{a}$. Using (1.2.12) we can immediately deduce that $\langle\alpha|\alpha\rangle$ must be a real number. To prove this just let $\langle\beta| \to \langle\alpha|$.

*In the literature an inner product is often referred to as a *scalar product* because it is analogous to $\mathbf{a}\cdot\mathbf{b}$ in Euclidean space; in this book, however, we reserve the term *scalar* for a quantity invariant under rotations in the usual three-dimensional space.

The second postulate on inner products is

$$\langle \alpha|\alpha\rangle \geq 0, \tag{1.2.13}$$

where the equality sign holds only if $|\alpha\rangle$ is a *null ket*. This is sometimes known as the postulate of **positive definite metric**. From a physicist's point of view, this postulate is essential for the probabilistic interpretation of quantum mechanics, as will become apparent later.*

Two kets $|\alpha\rangle$ and $|\beta\rangle$ are said to be **orthogonal** if

$$\langle \alpha|\beta\rangle = 0, \tag{1.2.14}$$

even though in the definition of the inner product the bra $\langle\alpha|$ appears. The orthogonality relation (1.2.14) also implies, via (1.2.12),

$$\langle \beta|\alpha\rangle = 0. \tag{1.2.15}$$

Given a ket which is not a null ket, we can form a **normalized ket** $|\tilde{\alpha}\rangle$, where

$$|\tilde{\alpha}\rangle = \left(\frac{1}{\sqrt{\langle \alpha|\alpha\rangle}}\right)|\alpha\rangle, \tag{1.2.16}$$

with the property

$$\langle \tilde{\alpha}|\tilde{\alpha}\rangle = 1. \tag{1.2.17}$$

Quite generally, $\sqrt{\langle \alpha|\alpha\rangle}$ is known as the **norm** of $|\alpha\rangle$, analogous to the magnitude of vector $\sqrt{\mathbf{a}\cdot\mathbf{a}} = |\vec{\mathbf{a}}|$ in Euclidean vector space. Because $|\alpha\rangle$ and $c|\alpha\rangle$ represent the same physical state, we might as well require that the kets we use for physical states be normalized in the sense of (1.2.17).†

Operators

As we remarked earlier, observables like momentum and spin components are to be represented by operators that can act on kets. We can consider a more general class of operators that act on kets; they will be denoted by X, Y, and so forth, while A, B, and so on will be used for a restrictive class of operators that correspond to observables.

An operator acts on a ket from the left side,

$$X\cdot(|\alpha\rangle) = X|\alpha\rangle, \tag{1.2.18}$$

and the resulting product is another ket. Operators X and Y are said to be **equal**,

$$X = Y, \tag{1.2.19}$$

*Attempts to abandon this postulate led to physical theories with "indefinite metric." We shall not be concerned with such theories in this book.

†For eigenkets of observables with continuous spectra, different normalization conventions will be used; see Section 1.6.

if

$$X|\alpha\rangle = Y|\alpha\rangle \tag{1.2.20}$$

for an *arbitrary* ket in the ket space in question. Operator X is said to be the **null operator** if, for any *arbitrary* ket $|\alpha\rangle$, we have

$$X|\alpha\rangle = 0. \tag{1.2.21}$$

Operators can be added; addition operations are commutative and associative:

$$X + Y = Y + X, \tag{1.2.21a}$$

$$X + (Y + Z) = (X + Y) + Z. \tag{1.2.21b}$$

With the single exception of the time-reversal operator to be considered in Chapter 4, the operators that appear in this book are all linear, that is,

$$X(c_\alpha|\alpha\rangle + c_\beta|\beta\rangle) = c_\alpha X|\alpha\rangle + c_\beta X|\beta\rangle. \tag{1.2.22}$$

An operator X always acts on a bra from the *right* side

$$(\langle\alpha|)\cdot X = \langle\alpha|X, \tag{1.2.23}$$

and the resulting product is another bra. The ket $X|\alpha\rangle$ and the bra $\langle\alpha|X$ are, in general, *not* dual to each other. We define the symbol $X^\dagger$ as

$$X|\alpha\rangle \overset{\text{DC}}{\leftrightarrow} \langle\alpha|X^\dagger. \tag{1.2.24}$$

The operator $X^\dagger$ is called the **Hermitian adjoint**, or simply the adjoint, of X. An operator X is said to be Hermitian if

$$X = X^\dagger. \tag{1.2.25}$$

Multiplication

Operators X and Y can be multiplied. Multiplication operations are, in general, *noncommutative*, that is,

$$XY \neq YX. \tag{1.2.26}$$

Multiplication operations are, however, associative:

$$X(YZ) = (XY)Z = XYZ. \tag{1.2.27}$$

We also have

$$X(Y|\alpha\rangle) = (XY)|\alpha\rangle = XY|\alpha\rangle, \qquad (\langle\beta|X)Y = \langle\beta|(XY) = \langle\beta|XY. \tag{1.2.28}$$

Notice that

$$(XY)^\dagger = Y^\dagger X^\dagger \tag{1.2.29}$$

because

$$XY|\alpha\rangle = X(Y|\alpha\rangle) \overset{\text{DC}}{\leftrightarrow} (\langle\alpha|Y^\dagger)X^\dagger = \langle\alpha|Y^\dagger X^\dagger. \tag{1.2.30}$$

So far, we have considered the following products: $\langle\beta|\alpha\rangle$, $X|\alpha\rangle$, $\langle\alpha|X$, and XY. Are there other products we are allowed to form? Let us multiply $|\beta\rangle$ and $\langle\alpha|$, in that order. The resulting product

$$(|\beta\rangle)\cdot(\langle\alpha|) = |\beta\rangle\langle\alpha| \tag{1.2.31}$$

is known as the **outer product** of $|\beta\rangle$ and $\langle\alpha|$. We will emphasize in a moment that $|\beta\rangle\langle\alpha|$ is to be regarded as an operator; hence it is fundamentally different from the inner product $\langle\beta|\alpha\rangle$, which is just a number.

There are also "illegal products." We have already mentioned that an operator must stand on the left of a ket or on the right of a bra. In other words, $|\alpha\rangle X$ and $X\langle\alpha|$ are examples of illegal products. They are neither kets, nor bras, nor operators; they are simply nonsensical. Products like $|\alpha\rangle|\beta\rangle$ and $\langle\alpha|\langle\beta|$ are also illegal when $|\alpha\rangle$ and $|\beta\rangle$ ($\langle\alpha|$ and $\langle\beta|$) are ket (bra) vectors belonging to the same ket (bra) space.*

The Associative Axiom

As is clear from (1.2.27), multiplication operations among operators are associative. Actually the associative property is postulated to hold quite generally as long as we are dealing with "legal" multiplications among kets, bras, and operators. Dirac calls this important postulate the **associative axiom of multiplication**.

To illustrate the power of this axiom let us first consider an outer product acting on a ket:

$$(|\beta\rangle\langle\alpha|)\cdot|\gamma\rangle. \tag{1.2.32}$$

Because of the associative axiom, we can regard this equally well as

$$|\beta\rangle\cdot(\langle\alpha|\gamma\rangle), \tag{1.2.33}$$

where $\langle\alpha|\gamma\rangle$ is just a number. So the outer product acting on a ket is just another ket; in other words, $|\beta\rangle\langle\alpha|$ can be regarded as an operator. Because (1.2.32) and (1.2.33) are equal, we may as well omit the dots and let $|\beta\rangle\langle\alpha|\gamma\rangle$ stand for the operator $|\beta\rangle\langle\alpha|$ acting on $|\gamma\rangle$ *or*, equivalently, the number $\langle\alpha|\gamma\rangle$ multiplying $|\beta\rangle$. (On the other hand, if (1.2.33) is written as $(\langle\alpha|\gamma\rangle)\cdot|\beta\rangle$, we cannot afford to omit the dot and brackets because the

*Later in the book we will encounter products like $|\alpha\rangle|\beta\rangle$, which are more appropriately written as $|\alpha\rangle\otimes|\beta\rangle$, but in such cases $|\alpha\rangle$ and $|\beta\rangle$ always refer to kets from *different* vector spaces. For instance, the first ket belongs to the vector space for electron spin, the second ket to the vector space for electron orbital angular momentum; or the first ket lies in the vector space of particle 1, the second ket in the vector space of particle 2, and so forth.

resulting expression would look illegal.) Notice that the operator $|\beta\rangle\langle\alpha|$ rotates $|\gamma\rangle$ into the direction of $|\beta\rangle$. It is easy to see that if

$$X = |\beta\rangle\langle\alpha|, \tag{1.2.34}$$

then

$$X^\dagger = |\alpha\rangle\langle\beta|, \tag{1.2.35}$$

which is left as an exercise.

In a second important illustration of the associative axiom, we note that

$$\underset{\text{bra}}{(\langle\beta|)} \cdot \underset{\text{ket}}{(X|\alpha\rangle)} = \underset{\text{bra}}{(\langle\beta|X)} \cdot \underset{\text{ket}}{(|\alpha\rangle)}. \tag{1.2.36}$$

Because the two sides are equal, we might as well use the more compact notation

$$\langle\beta|X|\alpha\rangle \tag{1.2.37}$$

to stand for either side of (1.2.36). Recall now that $\langle\alpha|X^\dagger$ is the bra that is dual to $X|\alpha\rangle$, so

$$\begin{aligned}\langle\beta|X|\alpha\rangle &= \langle\beta|\cdot(X|\alpha\rangle) \\ &= \{(\langle\alpha|X^\dagger)\cdot|\beta\rangle\}^* \\ &= \langle\alpha|X^\dagger|\beta\rangle^*, \end{aligned} \tag{1.2.38}$$

where, in addition to the associative axiom, we used the fundamental property of the inner product (1.2.12). For a *Hermitian* X we have

$$\langle\beta|X|\alpha\rangle = \langle\alpha|X|\beta\rangle^*. \tag{1.2.39}$$

1.3. BASE KETS AND MATRIX REPRESENTATIONS

Eigenkets of an Observable

Let us consider the eigenkets and eigenvalues of a Hermitian operator A. We use the symbol A, reserved earlier for an observable, because in quantum mechanics Hermitian operators of interest quite often turn out to be the operators representing some physical observables.

We begin by stating an important theorem:

Theorem. *The eigenvalues of a Hermitian operator A are real; the eigenkets of A corresponding to different eigenvalues are orthogonal.*

Proof. First, recall that

$$A|a'\rangle = a'|a'\rangle. \tag{1.3.1}$$

Because A is Hermitian, we also have

$$\langle a''|A = a''^*\langle a''|, \tag{1.3.2}$$

where $a', a'' \ldots$ are eigenvalues of A. If we multiply both sides of (1.3.1) by $\langle a''|$ on the left, both sides of (1.3.2) by $|a'\rangle$ on the right, and subtract, we obtain

$$(a' - a''^*)\langle a''|a'\rangle = 0. \tag{1.3.3}$$

Now a' and a'' can be taken to be either the same or different. Let us first choose them to be the same; we then deduce the reality condition (the first half of the theorem)

$$a' = a'^*, \tag{1.3.4}$$

where we have used the fact that $|a'\rangle$ is not a null ket. Let us now assume a' and a'' to be different. Because of the just-proved reality condition, the difference $a' - a''^*$ that appears in (1.3.3) is equal to $a' - a''$, which cannot vanish, by assumption. The inner product $\langle a''|a'\rangle$ must then vanish:

$$\langle a''|a'\rangle = 0, \quad (a' \neq a''), \tag{1.3.5}$$

which proves the orthogonality property (the second half of the theorem). □

We expect on physical grounds that an observable has real eigenvalues, a point that will become clearer in the next section, where measurements in quantum mechanics will be discussed. The theorem just proved guarantees the reality of eigenvalues whenever the operator is Hermitian. That is why we talk about Hermitian observables in quantum mechanics.

It is conventional to normalize $|a'\rangle$ so the $\{|a'\rangle\}$ form a **orthonormal** set:

$$\langle a''|a'\rangle = \delta_{a''a'}. \tag{1.3.6}$$

We may logically ask, Is this set of eigenkets complete? Since we started our discussion by asserting that the whole ket space is spanned by the eigenkets of A, the eigenkets of A must therefore form a complete set by *construction* of our ket space.*

Eigenkets as Base Kets

We have seen that the normalized eigenkets of A form a complete orthonormal set. An arbitrary ket in the ket space can be expanded in terms

* The astute reader, already familiar with wave mechanics, may point out that the completeness of eigenfunctions we use can be proved by applying the Sturm-Liouville theory to the Schrödinger wave equation. But to "derive" the Schrödinger wave equation from our fundamental postulates, the completeness of the position eigenkets must be assumed.

of the eigenkets of A. In other words, the eigenkets of A are to be used as base kets in much the same way as a set of mutually orthogonal unit vectors is used as base vectors in Euclidean space.

Given an arbitrary ket $|\alpha\rangle$ in the ket space spanned by the eigenkets of A, let us attempt to expand it as follows:

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle. \tag{1.3.7}$$

Multiplying $\langle a''|$ on the left and using the orthonormality property (1.3.6), we can immediately find the expansion coefficient,

$$c_{a'} = \langle a'|\alpha\rangle. \tag{1.3.8}$$

In other words, we have

$$|\alpha\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha\rangle, \tag{1.3.9}$$

which is analogous to an expansion of a vector $\mathbf{V}$ in (real) Euclidean space:

$$\mathbf{V} = \sum_{i} \hat{\mathbf{e}}_i(\hat{\mathbf{e}}_i \cdot \mathbf{V}), \tag{1.3.10}$$

where $\{\hat{\mathbf{e}}_i\}$ form an orthogonal set of unit vectors. We now recall the associative axiom of multiplication: $|a'\rangle\langle a'|\alpha\rangle$ can be regarded either as the number $\langle a'|\alpha\rangle$ multiplying $|a'\rangle$ or, equivalently, as the operator $|a'\rangle\langle a'|$ acting on $|\alpha\rangle$. Because $|\alpha\rangle$ in (1.3.9) is an arbitrary ket, we must have

$$\sum_{a'} |a'\rangle\langle a'| = 1, \tag{1.3.11}$$

where the 1 on the right-hand side is to be understood as the identity *operator*. Equation (1.3.11) is known as the **completeness relation** or **closure**.

It is difficult to overestimate the usefulness of (1.3.11). Given a chain of kets, operators, or bras multiplied in legal orders, we can insert, in any place at our convenience, the identity operator written in form (1.3.11). Consider, for example $\langle\alpha|\alpha\rangle$; by inserting the identity operator between $\langle\alpha|$ and $|\alpha\rangle$, we obtain

$$\begin{aligned}\langle\alpha|\alpha\rangle &= \langle\alpha| \cdot \left(\sum_{a'} |a'\rangle\langle a'|\right) \cdot |\alpha\rangle \\ &= \sum_{a'} |\langle a'|\alpha\rangle|^2 \end{aligned} \tag{1.3.12}$$

This, incidentally, shows that if $|\alpha\rangle$ is normalized, then the expansion coefficients in (1.3.7) must satisfy

$$\sum_{a'} |c_{a'}|^2 = \sum_{a'} |\langle a'|\alpha\rangle|^2 = 1. \tag{1.3.13}$$

Let us now look at $|a'\rangle\langle a'|$ that appears in (1.3.11). Since this is an outer product, it must be an operator. Let it operate on $|\alpha\rangle$:

$$\big(|a'\rangle\langle a'|\big)\cdot|\alpha\rangle = |a'\rangle\langle a'|\alpha\rangle = c_{a'}|a'\rangle. \tag{1.3.14}$$

We see that $|a'\rangle\langle a'|$ selects that portion of the ket $|\alpha\rangle$ parallel to $|a'\rangle$, so $|a'\rangle\langle a'|$ is known as the **projection operator** along the base ket $|a'\rangle$ and is denoted by $\Lambda_{a'}$:

$$\Lambda_{a'} \equiv |a'\rangle\langle a'|. \tag{1.3.15}$$

The completeness relation (1.3.11) can now be written as

$$\sum_{a'} \Lambda_{a'} = 1. \tag{1.3.16}$$

Matrix Representations

Having specified the base kets, we now show how to represent an operator, say X, by a square matrix. First, using (1.3.11) twice, we write the operator X as

$$X = \sum_{a''}\sum_{a'} |a''\rangle\langle a''|X|a'\rangle\langle a'|. \tag{1.3.17}$$

There are altogether N^2 numbers of form $\langle a''|X|a'\rangle$, where N is the dimensionality of the ket space. We may arrange them into an $N \times N$ square matrix such that the column and row indices appear as follows:

$$\underset{\text{row}}{\langle a''|}X\ \underset{\text{column}}{|a'\rangle}\ . \tag{1.3.18}$$

Explicitly we may write the matrix as

$$X \doteq \begin{pmatrix} \langle a^{(1)}|X|a^{(1)}\rangle & \langle a^{(1)}|X|a^{(2)}\rangle & \cdots \\ \langle a^{(2)}|X|a^{(1)}\rangle & \langle a^{(2)}|X|a^{(2)}\rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{1.3.19}$$

where the symbol $\doteq$ stands for "is represented by."*

Using (1.2.38), we can write

$$\langle a''|X|a'\rangle = \langle a'|X^\dagger|a''\rangle^*. \tag{1.3.20}$$

At last, the Hermitian adjoint operation, originally defined by (1.2.24), has been related to the (perhaps more familiar) concept of *complex conjugate transposed*. If an operator B is Hermitian, we have

$$\langle a''|B|a'\rangle = \langle a'|B|a''\rangle^*. \tag{1.3.21}$$

---

*We do not use the equality sign here because the particular form of a matrix representation depends on the particular choice of base kets used. The operator is different from a representation of the operator just as the actress is different from a poster of the actress.

The way we arranged $\langle a''|X|a'\rangle$ into a square matrix is in conformity with the usual rule of matrix multiplication. To see this just note that the matrix representation of the operator relation

$$Z = XY \tag{1.3.22}$$

reads

$$\begin{aligned}\langle a''|Z|a'\rangle &= \langle a''|XY|a'\rangle \\ &= \sum_{a'''} \langle a''|X|a'''\rangle\langle a'''|Y|a'\rangle.\end{aligned} \tag{1.3.23}$$

Again, all we have done is to insert the identity operator, written in form (1.3.11), between X and Y!

Let us now examine how the ket relation

$$|\gamma\rangle = X|\alpha\rangle \tag{1.3.24}$$

can be represented using our base kets. The expansion coefficients of $|\gamma\rangle$ can be obtained by multiplying $\langle a'|$ on the left:

$$\begin{aligned}\langle a'|\gamma\rangle &= \langle a'|X|\alpha\rangle \\ &= \sum_{a''}\langle a'|X|a''\rangle\langle a''|\alpha\rangle.\end{aligned} \tag{1.3.25}$$

But this can be seen as an application of the rule for multiplying a square matrix with a column matrix representing once the expansion coefficients of $|\alpha\rangle$ and $|\gamma\rangle$ arrange themselves to form column matrices as follows:

$$|\alpha\rangle \doteq \begin{pmatrix} \langle a^{(1)}|\alpha\rangle \\ \langle a^{(2)}|\alpha\rangle \\ \langle a^{(3)}|\alpha\rangle \\ \vdots \end{pmatrix}, \qquad |\gamma\rangle \doteq \begin{pmatrix} \langle a^{(1)}|\gamma\rangle \\ \langle a^{(2)}|\gamma\rangle \\ \langle a^{(3)}|\gamma\rangle \\ \vdots \end{pmatrix}. \tag{1.3.26}$$

Likewise, given

$$\langle\gamma| = \langle\alpha|X, \tag{1.3.27}$$

we can regard

$$\langle\gamma|a'\rangle = \sum_{a''}\langle\alpha|a''\rangle\langle a''|X|a'\rangle. \tag{1.3.28}$$

So a bra is represented by a row matrix as follows:

$$\begin{aligned}\langle\gamma| &\doteq (\langle\gamma|a^{(1)}\rangle, \langle\gamma|a^{(2)}\rangle, \langle\gamma|a^{(3)}\rangle, \ldots) \\ &= (\langle a^{(1)}|\gamma\rangle^*, \langle a^{(2)}|\gamma\rangle^*, \langle a^{(3)}|\gamma\rangle^*, \ldots).\end{aligned} \tag{1.3.29}$$

Note the appearance of complex conjugation when the elements of the column matrix are written as in (1.3.29). The inner product $\langle\beta|\alpha\rangle$ can be written as

the product of the row matrix representing $\langle\beta|$ with the column matrix representing $|\alpha\rangle$:

$$\begin{aligned}\langle\beta|\alpha\rangle &= \sum_{a'}\langle\beta|a'\rangle\langle a'|\alpha\rangle \\ &= \left(\langle a^{(1)}|\beta\rangle^*, \langle a^{(2)}|\beta\rangle^*, \ldots\right)\begin{pmatrix}\langle a^{(1)}|\alpha\rangle \\ \langle a^{(2)}|\alpha\rangle \\ \vdots\end{pmatrix}\end{aligned} \tag{1.3.30}$$

If we multiply the row matrix representing $\langle\alpha|$ with the column matrix representing $|\beta\rangle$, then we obtain just the complex conjugate of the preceding expression, which is consistent with the fundamental property of the inner product (1.2.12). Finally, the matrix representation of the outer product $|\beta\rangle\langle\alpha|$ is easily seen to be

$$|\beta\rangle\langle\alpha| \doteq \begin{pmatrix}\langle a^{(1)}|\beta\rangle\langle a^{(1)}|\alpha\rangle^* & \langle a^{(1)}|\beta\rangle\langle a^{(2)}|\alpha\rangle^* & \cdots \\ \langle a^{(2)}|\beta\rangle\langle a^{(1)}|\alpha\rangle^* & \langle a^{(2)}|\beta\rangle\langle a^{(2)}|\alpha\rangle^* & \cdots \\ \vdots & \vdots & \ddots\end{pmatrix}. \tag{1.3.31}$$

The matrix representation of an observable A becomes particularly simple if the eigenkets of A themselves are used as the base kets. First, we have

$$A = \sum_{a''}\sum_{a'}|a''\rangle\langle a''|A|a'\rangle\langle a'|. \tag{1.3.32}$$

But the square matrix $\langle a''|A|a'\rangle$ is obviously diagonal,

$$\langle a''|A|a'\rangle = \langle a'|A|a'\rangle\delta_{a'a''} = a'\delta_{a'a''}, \tag{1.3.33}$$

so

$$\begin{aligned}A &= \sum_{a'} a'|a'\rangle\langle a'| \\ &= \sum_{a'} a'\Lambda_{a'}.\end{aligned} \tag{1.3.34}$$

Spin $\frac{1}{2}$ Systems

It is here instructive to consider the special case of spin $\frac{1}{2}$ systems. The base kets used are $|S_z;\pm\rangle$, which we denote, for brevity, as $|\pm\rangle$. The simplest operator in the ket space spanned by $|\pm\rangle$ is the identity operator, which, according to (1.3.11), can be written as

$$1 = |+\rangle\langle+| + |-\rangle\langle-|. \tag{1.3.35}$$

According to (1.3.34), we must be able to write S_z as

$$S_z = (\hbar/2)\left[(|+\rangle\langle+|) - (|-\rangle\langle-|)\right]. \tag{1.3.36}$$

The eigenket-eigenvalue relation

$$S_z|\pm\rangle = \pm(\hbar/2)|\pm\rangle \tag{1.3.37}$$

immediately follows from the orthonormality property of $|\pm\rangle$.

It is also instructive to look at two other operators,

$$S_+ \equiv \hbar|+\rangle\langle -|, \qquad S_- \equiv \hbar|-\rangle\langle +|, \tag{1.3.38}$$

which are both seen to be *non*-Hermitian. The operator S_+, acting on the spin-down ket $|-\rangle$, turns $|-\rangle$ into the spin-up ket $|+\rangle$ multiplied by $\hbar$. On the other hand, the spin-up ket $|+\rangle$, when acted upon by S_+, becomes a null ket. So the physical interpretation of S_+ is that it raises the spin component by one unit of $\hbar$; if the spin component cannot be raised any further, we automatically get a null state. Likewise, S_- can be interpreted as an operator that lowers the spin component by one unit of $\hbar$. Later we will show that $S_\pm$ can be written as $S_x \pm iS_y$.

In constructing the matrix representations of the angular momentum operators, it is customary to label the column (row) indices in *descending* order of angular momentum components, that is, the first entry corresponds to the maximum angular momentum component, the second, the next highest, and so forth. In our particular case of spin $\frac{1}{2}$ systems, we have

$$|+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{1.3.39a}$$

$$S_z \doteq \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad S_+ \doteq \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad S_- \doteq \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{1.3.39b}$$

We will come back to these explicit expressions when we discuss the Pauli two-component formalism in Chapter 3.

1.4. MEASUREMENTS, OBSERVABLES, AND THE UNCERTAINTY RELATIONS

Measurements

Having developed the mathematics of ket spaces, we are now in a position to discuss the quantum theory of measurement processes. This is not a particularly easy subject for beginners, so we first turn to the words of the great master, P. A. M. Dirac, for guidance (Dirac 1958, 36): "A measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured." What does all this mean? We interpret Dirac's words as follows: Before a measurement of observable A is

made, the system is assumed to be represented by some linear combination

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle = \sum_{a'}|a'\rangle\langle a'|\alpha\rangle. \tag{1.4.1}$$

When the measurement is performed, the system is "thrown into" one of the eigenstates, say $|a'\rangle$ of observable A. In other words,

$$|\alpha\rangle \xrightarrow{A \text{ measurement}} |a'\rangle. \tag{1.4.2}$$

For example, a silver atom with an arbitrary spin orientation will change into either $|S_z; +\rangle$ or $|S_z; -\rangle$ when subjected to a SG apparatus of type SG$\hat{z}$. Thus *a measurement usually changes the state*. The only exception is when the state is already in one of the eigenstates of the observable being measured, in which case

$$|a'\rangle \xrightarrow{A \text{ measurement}} |a'\rangle \tag{1.4.3}$$

with certainty, as will be discussed further. When the measurement causes $|\alpha\rangle$ to change into $|a'\rangle$, it is said that A is measured to be a'. It is in this sense that the result of a measurement yields one of the eigenvalues of the observable being measured.

Given (1.4.1), which is the state ket of a physical system before the measurement, we do not know in advance into which of the various $|a'\rangle$'s the system will be thrown as the result of the measurement. We do postulate, however, that the probability for jumping into some particular $|a'\rangle$ is given by

$$\text{Probability for } a' = |\langle a'|\alpha\rangle|^2, \tag{1.4.4}$$

provided that $|\alpha\rangle$ is normalized.

Although we have been talking about a single physical system, to determine probability (1.4.4) empirically, we must consider a great number of measurements performed on an ensemble—that is, a collection—of identically prepared physical systems, all characterized by the same ket $|\alpha\rangle$. Such an ensemble is known as a **pure ensemble**. (We will say more about ensembles in Chapter 3.) As an example, a beam of silver atoms which survive the first SG$\hat{z}$ apparatus of Figure 1.3 with the S_z- component blocked is an example of a pure ensemble because every member atom of the ensemble is characterized by $|S_z; +\rangle$.

The probabilistic interpretation (1.4.4) for the squared inner product $|\langle a'|\alpha\rangle|^2$ is one of the fundamental postulates of quantum mechanics, so it cannot be proven. Let us note, however, that it makes good sense in extreme cases. Suppose the state ket is $|a'\rangle$ itself even before a measurement is made; then according to (1.4.4), the probability for getting a'—or, more precisely, for being thrown into $|a'\rangle$—as the result of the measurement is predicted to be 1, which is just what we expect. By measuring A once again,

we, of course, get $|a'\rangle$ only; quite generally, repeated measurements of the same observable in succession yield the same result.* If, on the other hand, we are interested in the probability for the system initially characterized by $|a'\rangle$ to be thrown into some other eigenket $|a''\rangle$ with $a'' \neq a'$, then (1.4.4) gives zero because of the orthogonality between $|a'\rangle$ and $|a''\rangle$. From the point of view of measurement theory, orthogonal kets correspond to mutually exclusive alternatives; for example, if a spin $\frac{1}{2}$ system is in $|S_z; +\rangle$, it is not in $|S_z; -\rangle$ with certainty.

Quite generally, the probability for anything must be nonnegative. Furthermore, the probabilities for the various alternative possibilities must add up to unity. Both of these expectations are met by our probability postulate (1.4.4).

We define the **expectation value** of A taken with respect to state $|\alpha\rangle$ as

$$\langle A \rangle \equiv \langle \alpha | A | \alpha \rangle. \tag{1.4.5}$$

To make sure that we are referring to state $|\alpha\rangle$, the notation $\langle A \rangle_\alpha$ is sometimes used. Equation (1.4.5) is a definition; however, it agrees with our intuitive notion of *average measured value* because it can be written as

$$\begin{aligned} \langle A \rangle &= \sum_{a'} \sum_{a''} \langle \alpha | a'' \rangle \langle a'' | A | a' \rangle \langle a' | \alpha \rangle \\ &= \sum_{a'} \underset{\substack{\uparrow \\ \text{measured value } a'}}{a'} \underbrace{|\langle a' | \alpha \rangle|^2}_{\text{probability for obtaining } a'} \end{aligned} \tag{1.4.6}$$

It is very important not to confuse eigenvalues with expectation values. For example, the expectation value of S_z for spin $\frac{1}{2}$ systems can assume *any* real value between $-\hbar/2$ and $+\hbar/2$, say $0.273\hbar$; in contrast, the eigenvalue of S_z assumes only two values, $\hbar/2$ and $-\hbar/2$.

To clarify further the meaning of measurements in quantum mechanics we introduce the notion of a **selective measurement**, or *filtration*. In Section 1.1 we considered a Stern-Gerlach arrangement where we let only one of the spin components pass out of the apparatus while we completely blocked the other component. More generally, we imagine a measurement process with a device that selects only one of the eigenkets of A, say $|a'\rangle$, and rejects all others; see Figure 1.6. This is what we mean by a selective measurement; it is also called filtration because only one of the A eigenkets filters through the ordeal. Mathematically we can say that such a selective

*Here successive measurements must be carried out immediately afterward. This point will become clear when we discuss the time evolution of a state ket in Chapter 2.

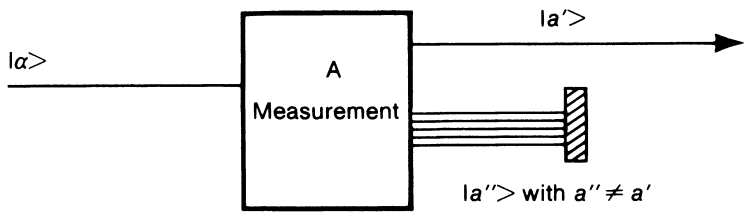


FIGURE 1.6. Selective measurement.

measurement amounts to applying the projection operator $\Lambda_{a'}$ to $|\alpha\rangle$:

$$\Lambda_{a'}|\alpha\rangle = |a'\rangle\langle a'|\alpha\rangle. \tag{1.4.7}$$

J. Schwinger has developed a formalism of quantum mechanics based on a thorough examination of selective measurements. He introduces a measurement symbol $M(a')$ in the beginning, which is identical to $\Lambda_{a'}$ or $|a'\rangle\langle a'|$ in our notation, and deduces a number of properties of $M(a')$ (and also of $M(b', a')$ which amount to $|b'\rangle\langle a'|$) by studying the outcome of various Stern-Gerlach–type experiments. In this way he motivates the entire mathematics of kets, bras, and operators. In this book we do not follow Schwinger's path; the interested reader may consult Gottfried's book. (Gottfried 1966, 192–9).

Spin $\frac{1}{2}$ Systems, Once Again

Before proceeding with a general discussion of observables, we once again consider spin $\frac{1}{2}$ systems. This time we show that the results of sequential Stern-Gerlach experiments, when combined with the postulates of quantum mechanics discussed so far, are sufficient to determine not only the $S_{x,y}$ eigenkets, $|S_x; \pm\rangle$ and $|S_y; \pm\rangle$, but also the operators S_x and S_y themselves.

First, we recall that when the $S_x +$ beam is subjected to an apparatus of type SG$\hat{z}$, the beam splits into two components with equal intensities. This means that the probability for the $S_x +$ state to be thrown into $|S_z; \pm\rangle$, simply denoted as $|\pm\rangle$, is $\frac{1}{2}$ each; hence,

$$|\langle + |S_x; +\rangle| = |\langle - |S_x; +\rangle| = \frac{1}{\sqrt{2}}. \tag{1.4.8}$$

We can therefore construct the $S_x +$ ket as follows:

$$|S_x; +\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}e^{i\delta_1}|-\rangle, \tag{1.4.9}$$

with δ_1 real. In writing (1.4.9) we have used the fact that the *overall* phase (common to both $|+\rangle$ and $|-\rangle$) of a state ket is immaterial; the coefficient

of $|+\rangle$ can be chosen to be real and positive by convention. The S_x- ket must be orthogonal to the S_x+ ket because the S_x+ alternative and S_x- alternative are mutually exclusive. This orthogonality requirement leads to

$$|S_x;-\rangle = \frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}e^{i\delta_1}|-\rangle, \tag{1.4.10}$$

where we have, again, chosen the coefficient of $|+\rangle$ to be real and positive by convention. We can now construct the operator S_x using (1.3.34) as follows:

$$\begin{aligned} S_x &= \frac{\hbar}{2}\left[(|S_x;+\rangle\langle S_x;+|) - (|S_x;-\rangle\langle S_x;-|)\right] \\ &= \frac{\hbar}{2}\left[e^{-i\delta_1}(|+\rangle\langle -|) + e^{i\delta_1}(|-\rangle\langle +|)\right]. \end{aligned} \tag{1.4.11}$$

Notice that the S_x we have constructed is Hermitian, just as it must be. A similar argument with S_x replaced by S_y leads to

$$|S_y;\pm\rangle = \frac{1}{\sqrt{2}}|+\rangle \pm \frac{1}{\sqrt{2}}e^{i\delta_2}|-\rangle, \tag{1.4.12}$$

$$S_y = \frac{\hbar}{2}\left[e^{-i\delta_2}(|+\rangle\langle -|) + e^{i\delta_2}(|-\rangle\langle +|)\right]. \tag{1.4.13}$$

Is there any way of determining δ_1 and δ_2? Actually there is one piece of information we have not yet used. Suppose we have a beam of spin $\frac{1}{2}$ atoms moving in the z-direction. We can consider a sequential Stern-Gerlach experiment with SG$\hat{x}$ followed by SG$\hat{y}$. The results of such an experiment are completely analogous to the earlier case leading to (1.4.8):

$$|\langle S_y;\pm|S_x;+\rangle| = |\langle S_y;\pm|S_x;-\rangle| = \frac{1}{\sqrt{2}}, \tag{1.4.14}$$

which is not surprising in view of the invariance of physical systems under rotations. Inserting (1.4.10) and (1.4.12) into (1.4.14), we obtain

$$\frac{1}{2}|1 \pm e^{i(\delta_1-\delta_2)}| = \frac{1}{\sqrt{2}}, \tag{1.4.15}$$

which is satisfied only if

$$\delta_2 - \delta_1 = \pi/2 \quad \text{or} \quad -\pi/2. \tag{1.4.16}$$

We thus see that the matrix elements of S_x and S_y cannot all be real. If the S_x matrix elements are real, the S_y matrix elements must be purely imaginary (and vice versa). Just from this extremely simple example, the introduction of complex numbers is seen to be an essential feature in quantum mechanics. It is convenient to take the S_x matrix elements to be real* and

*This can always be done by adjusting arbitrary phase factors in the definition of $|+\rangle$ and $|-\rangle$. This point will become clearer in Chapter 3, where the behavior of $|\pm\rangle$ under rotations will be discussed.

set $\delta_1 = 0$; if we were to choose $\delta_1 = \pi$, the positive x-axis would be oriented in the opposite direction. The second phase angle δ_2 must then be $-\pi/2$ or $\pi/2$. The fact that there is still an ambiguity of this kind is not surprising. We have not yet specified whether the coordinate system we are using is right-handed or left-handed; given the x- and the z-axes there is still a twofold ambiguity in the choice of the positive y-axis. Later we will discuss angular momentum as a generator of rotations using the right-handed coordinate system; it can then be shown that $\delta_2 = \pi/2$ is the correct choice.

To summarize, we have

$$|S_x; \pm\rangle = \frac{1}{\sqrt{2}}|+\rangle \pm \frac{1}{\sqrt{2}}|-\rangle, \tag{1.4.17a}$$

$$|S_y; \pm\rangle = \frac{1}{\sqrt{2}}|+\rangle \pm \frac{i}{\sqrt{2}}|-\rangle, \tag{1.4.17b}$$

and

$$S_x = \frac{\hbar}{2}\left[(|+\rangle\langle -|) + (|-\rangle\langle +|)\right], \tag{1.4.18a}$$

$$S_y = \frac{\hbar}{2}\left[-i(|+\rangle\langle -|) + i(|-\rangle\langle +|)\right]. \tag{1.4.18b}$$

The $S_x \pm$ and $S_y \pm$ eigenkets given here are seen to be in agreement with our earlier guesses (1.1.9) and (1.1.14) based on an analogy with linearly and circularly polarized light. (Note, in this comparison, that only the relative phase between the $|+\rangle$ and $|-\rangle$ components is of physical significance.) Furthermore, the non-Hermitian $S_\pm$ operators defined by (1.3.38) can now be written as

$$S_\pm = S_x \pm iS_y. \tag{1.4.19}$$

The operators S_x and S_y, together with S_z given earlier, can be readily shown to satisfy the commutation relations

$$\left[S_i, S_j\right] = i\epsilon_{ijk}\hbar S_k, \tag{1.4.20}$$

and the anticommutation relations

$$\left\{S_i, S_j\right\} = \frac{1}{2}\hbar^2\delta_{ij}, \tag{1.4.21}$$

where the commutator [,] and the anticommutator { , } are defined by

$$[A, B] \equiv AB - BA, \tag{1.4.22a}$$

$$\{A, B\} \equiv AB + BA. \tag{1.4.22b}$$

The commutation relations in (1.4.20) will be recognized as the simplest realization of the angular momentum commutation relations, whose significance will be discussed in detail in Chapter 3. In contrast, the anticommutation relations in (1.4.21) turn out to be a *special* property of spin $\frac{1}{2}$ systems.

We can also define the operator $\mathbf{S}\cdot\mathbf{S}$, or $\mathbf{S}^2$ for short, as follows:

$$\mathbf{S}^2 \equiv S_x^2 + S_y^2 + S_z^2. \tag{1.4.23}$$

Because of (1.4.21), this operator turns out to be just a constant multiple of the identity operator

$$\mathbf{S}^2 = \left(\frac{3}{4}\right)\hbar^2. \tag{1.4.24}$$

We obviously have

$$[\mathbf{S}^2, S_i] = 0. \tag{1.4.25}$$

As will be shown in Chapter 3, for spins higher than $\frac{1}{2}$, $\mathbf{S}^2$ is no longer a multiple of the identity operator; however, (1.4.25) still holds.

Compatible Observables

Returning now to the general formalism, we will discuss compatible versus incompatible observables. Observables A and B are defined to be **compatible** when the corresponding operators commute,

$$[A, B] = 0, \tag{1.4.26}$$

and **incompatible** when

$$[A, B] \neq 0, \tag{1.4.27}$$

For example, $\mathbf{S}^2$ and S_z are compatible observables, while S_x and S_z are incompatible observables.

Let us first consider the case of compatible observables A and B. As usual, we assume that the ket space is spanned by the eigenkets of A. We may also regard the same ket space as being spanned by the eigenkets of B. We now ask, How are the A eigenkets related to the B eigenkets when A and B are compatible observables?

Before answering this question we must touch upon a very important point we have bypassed earlier—the concept of *degeneracy*. Suppose there are two (or more) linearly independent eigenkets of A having the same eigenvalue; then the eigenvalues of the two eigenkets are said to be **degenerate**. In such a case the notation $|a'\rangle$ that labels the eigenket by its eigenvalue alone does not give a complete description; furthermore, we may recall that our earlier theorem on the orthogonality of different eigenkets was proved under the assumption of no degeneracy. Even worse, the whole concept that the ket space is spanned by $\{|a'\rangle\}$ appears to run into difficulty when the dimensionality of the ket space is larger than the number of distinct eigenvalues of A. Fortunately, in practical applications in quantum mechanics, it is usually the case that in such a situation the eigenvalues of *some other* commuting observable, say B, can be used to label the degenerate eigenkets.

Now we are ready to state an important theorem.

Theorem. *Suppose that A and B are compatible observables, and the eigenvalues of A are nondegenerate. Then the matrix elements $\langle a''|B|a'\rangle$ are all diagonal. (Recall here that the matrix elements of A are already diagonal if $\{|a'\rangle\}$ are used as the base kets.)*

Proof. The proof of this important theorem is extremely simple. Using the definition (1.4.26) of compatible observables, we observe that

$$\langle a''|[A, B]|a'\rangle = (a'' - a')\langle a''|B|a'\rangle = 0. \tag{1.4.28}$$

So $\langle a''|B|a'\rangle$ must vanish unless $a' = a''$, which proves our assertion. □

We can write the matrix elements of B as

$$\langle a''|B|a'\rangle = \delta_{a'a''}\langle a'|B|a'\rangle. \tag{1.4.29}$$

So both A and B can be represented by diagonal matrices with the *same* set of base kets. Using (1.3.17) and (1.4.29) we can write B as

$$B = \sum_{a''} |a''\rangle\langle a''|B|a''\rangle\langle a''|. \tag{1.4.30}$$

Suppose that this operator acts on an eigenket of A:

$$B|a'\rangle = \sum_{a''} |a''\rangle\langle a''|B|a''\rangle\langle a''|a'\rangle = \big(\langle a'|B|a'\rangle\big)|a'\rangle. \tag{1.4.31}$$

But this is nothing other than the eigenvalue equation for the operator B with eigenvalue

$$b' \equiv \langle a'|B|a'\rangle. \tag{1.4.32}$$

The ket $|a'\rangle$ is therefore a **simultaneous eigenket** of A and B. Just to be impartial to both operators, we may use $|a', b'\rangle$ to characterize this simultaneous eigenket.

We have seen that compatible observables have simultaneous eigenkets. Even though the proof given is for the case where the A eigenkets are nondegenerate, the statement holds even if there is an n-fold degeneracy, that is,

$$A|a'^{(i)}\rangle = a'|a'^{(i)}\rangle \quad \text{for } i = 1, 2, \ldots, n \tag{1.4.33}$$

where $|a'^{(i)}\rangle$ are n mutually orthonormal eigenkets of A, all with the same eigenvalue a'. To see this, all we need to do is construct appropriate linear combinations of $|a'^{(i)}\rangle$ that diagonalize the B operator by following the diagonalization procedure to be discussed in Section 1.5.

A simultaneous eigenket of A and B, denoted by $|a', b'\rangle$, has the property

$$A|a', b'\rangle = a'|a', b'\rangle, \tag{1.4.34a}$$

$$B|a', b'\rangle = b'|a', b'\rangle. \tag{1.4.34b}$$

When there is no degeneracy, this notation is somewhat superfluous because it is clear from (1.4.32) that if we specify a', we necessarily know the b' that appears in $|a', b'\rangle$. The notation $|a', b'\rangle$ is much more powerful when there are degeneracies. A simple example may be used to illustrate this point.

Even though a complete discussion of orbital angular momentum will not appear in this book until Chapter 3, the reader may be familiar from his or her earlier training in elementary wave mechanics that the eigenvalues of $\mathbf{L}^2$ (orbital angular momentum squared) and L_z (the z-component of orbital angular momentum) are $\hbar^2 l(l+1)$ and $m_l\hbar$, respectively, with l an integer and $m_l = -l, -l+1, \ldots, +l$. To characterize an orbital angular momentum state completely, it is necessary to specify *both* l and m_l. For example, if we just say $l=1$, the m_l value can still be 0, $+1$, or -1; if we just say $m_l=1$, l can be 1, 2, 3, 4, and so on. Only by specifying *both* l and m_l do we succeed in uniquely characterizing the orbital angular momentum state in question. Quite often a **collective index** K' is used to stand for (a', b'), so that

$$|K'\rangle = |a', b'\rangle. \tag{1.4.35}$$

We can obviously generalize our considerations to a situation where there are several (more than two) mutually compatible observables, namely,

$$[A, B] = [B, C] = [A, C] = \cdots = 0. \tag{1.4.36}$$

Assume that we have found a **maximal** set of commuting observables; that is, we cannot add any more observables to our list without violating (1.4.36). The eigenvalues of individual operators $A, B, C, \ldots$ may have degeneracies, but if we specify a combination $(a', b', c', \ldots)$, then the corresponding simultaneous eigenket of $A, B, C, \ldots$ is uniquely specified. We can again use a collective index K' to stand for $(a', b', c', \ldots)$. The orthonormality relation for

$$|K'\rangle = |a', b', c', \ldots\rangle \tag{1.4.37}$$

reads

$$\langle K''|K'\rangle = \delta_{K'K''} = \delta_{aa'}\delta_{bb'}\delta_{cc'}\ldots, \tag{1.4.38}$$

while the completeness relation, or closure, can be written as

$$\sum_{K'} |K'\rangle\langle K'| = \sum_{a'}\sum_{b'}\sum_{c'}\ldots|a', b', c', \ldots\rangle\langle a', b', c', \ldots| = 1. \tag{1.4.39}$$

We now consider measurements of A and B when they are compatible observables. Suppose we measure A first and obtain result a'. Subsequently, we may measure B and get result b'. Finally we measure A again. It follows from our measurement formalism that the third measurement always gives a' with certainty, that is, the second (B) measurement does not destroy the previous information obtained in the first (A) measurement. This is rather obvious when the eigenvalues of A are nondegenerate:

$$|\alpha\rangle \xrightarrow{A \text{ measurement}} |a', b'\rangle \xrightarrow{B \text{ measurement}} |a', b'\rangle \xrightarrow{A \text{ measurement}} |a', b'\rangle. \tag{1.4.40}$$

When there is degeneracy, the argument goes as follows: After the first (A) measurement, which yields a', the system is thrown into some linear combination

$$\sum_i^n c_{a'}^{(i)} |a', b^{(i)}\rangle, \tag{1.4.41}$$

where n is the degree of degeneracy and the kets $|a', b^{(i)}\rangle$ all have the same eigenvalue a' as far as operator A is concerned. The second (B) measurement may select just one of the terms in the linear combination (1.4.41), say, $|a', b^{(j)}\rangle$, but the third (A) measurement applied to it still yields a'. Whether or not there is degeneracy, A measurements and B measurements do not interfere. The term *compatible* is indeed deemed appropriate.

Incompatible Observables

We now turn to incompatible observables, which are more nontrivial. The first point to be emphasized is that incompatible observables do not have a complete set of simultaneous eigenkets. To show this let us assume the converse to be true. There would then exist a set of simultaneous eigenkets with property (1.4.34a) and (1.4.34b). Clearly,

$$AB|a', b'\rangle = Ab'|a', b'\rangle = a'b'|a', b'\rangle. \tag{1.4.42}$$

Likewise,

$$BA|a', b'\rangle = Ba'|a', b'\rangle = a'b'|a', b'\rangle; \tag{1.4.43}$$

hence,

$$AB|a', b'\rangle = BA|a', b'\rangle, \tag{1.4.44}$$

and thus $[A, B] = 0$ in contradiction to the assumption. So in general, $|a', b'\rangle$ does not make sense for incompatible observables. There is, however, an interesting exception; it may happen that there exists a subspace of the ket space such that (1.4.44) holds for all elements of this subspace, even though A and B are incompatible. An example from the theory of orbital

angular momentum may be helpful here. Suppose we consider an $l = 0$ state (s-state). Even though L_x and L_z do *not* commute, this state *is* a simultaneous eigenstate of L_x and L_z (with eigenvalue zero for both operators). The subspace in this case is one-dimensional.

We already encountered some of the peculiarities associated with incompatible observables when we discussed sequential Stern-Gerlach experiments in Section 1.1. We now give a more general discussion of experiments of that type. Consider the sequence of selective measurements shown in Figure 1.7(a). The first (A) filter selects some particular $|a'\rangle$ and rejects all others, the second (B) filter selects some particular $|b'\rangle$ and rejects all others, and the third (C) filter selects some particular $|c'\rangle$ and rejects all others. We are interested in the probability of obtaining $|c'\rangle$ when the beam coming out of the first filter is normalized to unity. Because the probabilities are multiplicative, we obviously have

$$|\langle c'|b'\rangle|^2|\langle b'|a'\rangle|^2. \tag{1.4.45}$$

Now let us sum over b' to consider the total probability for going through all possible b' routes. Operationally this means that we first record the probability of obtaining c' with all but the first b' route blocked, then we repeat the procedure with all but the second b' blocked, and so on; then we sum the probabilities at the end and obtain

$$\sum_{b'} |\langle c'|b'\rangle|^2|\langle b'|a'\rangle|^2 = \sum_{b'} \langle c'|b'\rangle\langle b'|a'\rangle\langle a'|b'\rangle\langle b'|c'\rangle. \tag{1.4.46}$$

We now compare this with a different arrangement, where the B filter is absent (or not operative); see Figure 1.7b. Clearly, the probability is

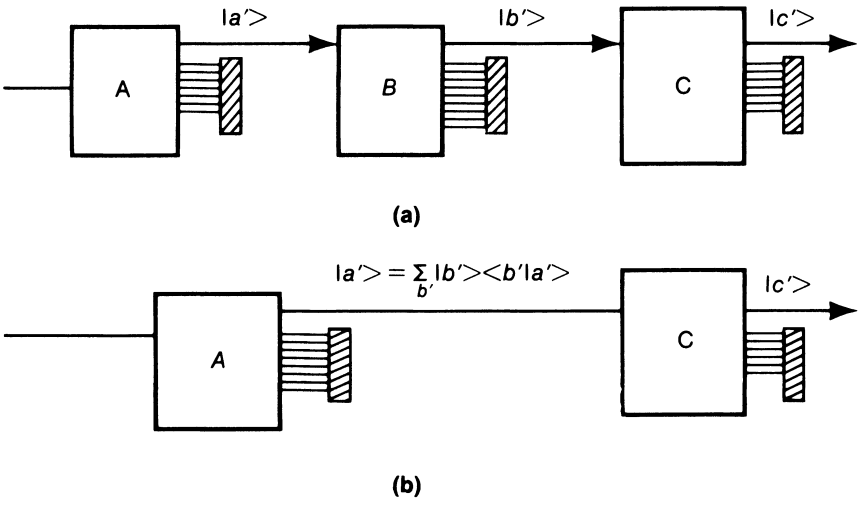


FIGURE 1.7. Sequential selective measurements.

just $|\langle c'|a'\rangle|^2$, which can also be written as follows:

$$|\langle c'|a'\rangle|^2 = |\sum_{b'} \langle c'|b'\rangle\langle b'|a'\rangle|^2 = \sum_{b'}\sum_{b''} \langle c'|b'\rangle\langle b'|a'\rangle\langle a'|b''\rangle\langle b''|c'\rangle. \tag{1.4.47}$$

Notice that expressions (1.4.46) and (1.4.47) are different! This is remarkable because in both cases the pure $|a'\rangle$ beam coming out of the first (A) filter can be regarded as being made up of the B eigenkets

$$|a'\rangle = \sum_{b'} |b'\rangle\langle b'|a'\rangle, \tag{1.4.48}$$

where the sum is over all possible values of b'. The crucial point to be noted is that the result coming out of the C filter depends on whether or not B measurements have actually been carried out. In the first case we experimentally ascertain which of the B eigenvalues are actually realized; in the second case, we merely imagine $|a'\rangle$ to be built up of the various $|b'\rangle$'s in the sense of (1.4.48). Put in another way, actually recording the probabilities of going through the various b' routes makes all the difference even though we sum over b' afterwards. Here lies the heart of quantum mechanics.

Under what conditions do the two expressions become equal? It is left as an exercise for the reader to show that for this to happen, in the absence of degeneracy, it is sufficient that

$$[A, B] = 0 \quad \text{or} \quad [B, C] = 0. \tag{1.4.49}$$

In other words, the peculiarity we have illustrated is characteristic of incompatible observables.

The Uncertainty Relation

The last topic to be discussed in this section is the uncertainty relation. Given an observable A, we define an **operator**

$$\Delta A \equiv A - \langle A\rangle, \tag{1.4.50}$$

where the expectation value is to be taken for a certain physical state under consideration. The expectation value of $(\Delta A)^2$ is known as the **dispersion** of A. Because we have

$$\langle(\Delta A)^2\rangle = \langle(A^2 - 2A\langle A\rangle + \langle A\rangle^2)\rangle = \langle A^2\rangle - \langle A\rangle^2, \tag{1.4.51}$$

the last line of (1.4.51) may be taken as an alternative definition of dispersion. Sometimes the terms **variance** and **mean square deviation** are used for the same quantity. Clearly, the dispersion vanishes when the state in question is an eigenstate of A. Roughly speaking, the dispersion of an observable characterizes "fuzziness." For example, for the $S_z +$ state of a

spin $\frac{1}{2}$ system, the dispersion of S_x can be computed to be

$$\langle S_x^2\rangle - \langle S_x\rangle^2 = \hbar^2/4. \tag{1.4.52}$$

In contrast the dispersion $\langle \Delta S_z)^2\rangle$ obviously vanishes for the $S_z +$ state. So, for the $S_z +$ state, S_z is "sharp"—a vanishing dispersion for S_z—while S_x is fuzzy.

We now state the uncertainty relation, which is the generalization of the well-known x-p uncertainty relation to be discussed in Section 1.6. Let A and B be observables. Then for any state we must have the following inequality:

$$\langle (\Delta A)^2\rangle\langle (\Delta B)^2\rangle \geq \frac{1}{4}|\langle [A, B]\rangle|^2. \tag{1.4.53}$$

To prove this we first state three lemmas.

Lemma 1. *The Schwarz inequality*

$$\langle \alpha|\alpha\rangle\langle \beta|\beta\rangle \geq |\langle \alpha|\beta\rangle|^2, \tag{1.4.54}$$

which is analogous to

$$|\mathbf{a}|^2|\mathbf{b}|^2 \geq |\mathbf{a}\cdot\mathbf{b}|^2 \tag{1.4.55}$$

in real Euclidian space.

Proof. First note

$$\big(\langle \alpha| + \lambda^*\langle \beta|\big)\cdot\big(|\alpha\rangle + \lambda|\beta\rangle\big) \geq 0, \tag{1.4.56}$$

where λ can be any complex number. This inequality must hold when λ is set equal to $-\langle \beta|\alpha\rangle/\langle \beta|\beta\rangle$:

$$\langle \alpha|\alpha\rangle\langle \beta|\beta\rangle - |\langle \alpha|\beta\rangle|^2 \geq 0, \tag{1.4.57}$$

which is the same as (1.4.54). □

Lemma 2. *The expectation value of a Hermitian operator is purely real.*

Proof. The proof is trivial—just use (1.3.21). □

Lemma 3. *The expectation value of an anti-Hermitian operator, defined by* $C = -C^\dagger$, *is purely imaginary.*

Proof. The proof is also trivial. □

Armed with these lemmas, we are in a position to prove the uncertainty relation (1.4.53). Using Lemma 1 with

$$\begin{aligned} |\alpha\rangle &= \Delta A|\ \rangle, \\ |\beta\rangle &= \Delta B|\ \rangle, \end{aligned} \tag{1.4.58}$$

where the blank ket | ⟩ emphasizes the fact that our consideration may be applied to *any* ket, we obtain

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geq |\langle \Delta A \Delta B\rangle|^2, \tag{1.4.59}$$

where the Hermiticity of ΔA and ΔB has been used. To evaluate the right-hand side of (1.4.59), we note

$$\Delta A \Delta B = \frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\}, \tag{1.4.60}$$

where the commutator $[\Delta A, \Delta B]$, which is equal to $[A, B]$, is clearly anti-Hermitian

$$([A, B])^\dagger = (AB - BA)^\dagger = BA - AB = -[A, B]. \tag{1.4.61}$$

In contrast, the anticommutator $\{\Delta A, \Delta B\}$ is obviously Hermitian, so

$$\langle \Delta A \Delta B\rangle = \frac{1}{2}\underbrace{\langle [A, B]\rangle}_{\substack{\text{purely}\\ \text{imaginary}}} + \frac{1}{2}\underbrace{\langle \{\Delta A, \Delta B\}\rangle}_{\substack{\text{purely}\\ \text{real}}}, \tag{1.4.62}$$

where Lemmas 2 and 3 have been used. The right-hand side of (1.4.59) now becomes

$$|\langle \Delta A \Delta B\rangle|^2 = \frac{1}{4}|\langle [A, B]\rangle|^2 + \frac{1}{4}|\langle \{\Delta A, \Delta B\}\rangle|^2. \tag{1.4.63}$$

The proof of (1.4.53) is now complete because the omission of the second (the anticommutator) term of (1.4.63) can only make the inequality relation stronger.*

Applications of the uncertainty relation to spin $\frac{1}{2}$ systems will be left as exercises. We come back to this topic when we discuss the fundamental x-p commutation relation in Section 1.6.

1.5. CHANGE OF BASIS

Transformation Operator

Suppose we have two incompatible observables A and B. The ket space in question can be viewed as being spanned either by the set $\{|a'\rangle\}$ or by the set $\{|b'\rangle\}$. For example, for spin $\frac{1}{2}$ systems $|S_z \pm\rangle$ may be used as our base kets; alternatively, $|S_x \pm\rangle$ may be used as our base kets. The two different sets of base kets, of course, span the same ket space. We are interested in finding out how the two descriptions are related. Changing the

*In the literature most authors use ΔA for our $\sqrt{\langle(\Delta A)^2\rangle}$ so the uncertainty relation is written as $\Delta A \Delta B \geq \frac{1}{2}|\langle [A, B]\rangle|$. In this book, however, ΔA and ΔB are to be understood as operators [see (1.4.50)], not numbers.

set of base kets is referred to as a **change of basis** or a **change of representation**. The basis in which the base eigenkets are given by $\{|a'\rangle\}$ is called the A representation or, sometimes, the A diagonal representation because the square matrix corresponding to A is diagonal in this basis.

Our basic task is to construct a transformation operator that connects the old orthonormal set $\{|a'\rangle\}$ and the new orthonormal set $\{|b'\rangle\}$. To this end, we first show the following.

Theorem. *Given two sets of base kets, both satisfying orthonormality and completeness, there exists a unitary operator U such that*

$$|b^{(1)}\rangle = U|a^{(1)}\rangle, |b^{(2)}\rangle = U|a^{(2)}\rangle, \ldots, |b^{(N)}\rangle = U|a^{(N)}\rangle. \quad (1.5.1)$$

By a **unitary operator** *we mean an operator fulfilling the conditions*

$$U^\dagger U = 1 \quad (1.5.2)$$

as well as

$$UU^\dagger = 1. \quad (1.5.3)$$

Proof. We prove this theorem by explicit construction. We assert that the operator

$$U = \sum_k |b^{(k)}\rangle\langle a^{(k)}| \quad (1.5.4)$$

will do the job and we apply this U to $|a^{(l)}\rangle$. Clearly,

$$U|a^{(l)}\rangle = |b^{(l)}\rangle \quad (1.5.5)$$

is guaranteed by the orthonormality of $\{|a'\rangle\}$. Furthermore, U is unitary:

$$U^\dagger U = \sum_k \sum_l |a^{(l)}\rangle\langle b^{(l)}|b^{(k)}\rangle\langle a^{(k)}| = \sum_k |a^{(k)}\rangle\langle a^{(k)}| = 1, \quad (1.5.6)$$

where we have used the orthonormality of $\{|b'\rangle\}$ and the completeness of $\{|a'\rangle\}$. We obtain relation (1.5.3) in an analogous manner. □

Transformation Matrix

It is instructive to study the matrix representation of the U operator in the old $\{|a'\rangle\}$ basis. We have

$$\langle a^{(k)}|U|a^{(l)}\rangle = \langle a^{(k)}|b^{(l)}\rangle, \quad (1.5.7)$$

which is obvious from (1.5.5). In other words, the matrix elements of the U operator are built up of the inner products of old base bras and new base kets. We recall that the rotation matrix in three dimensions that changes one set of unit base vectors $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ into another set $(\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{z}}')$ can be written as

(Goldstein 1980, 128–37, for example)

$$R = \begin{pmatrix} \hat{\mathbf{x}}\cdot\hat{\mathbf{x}}' & \hat{\mathbf{x}}\cdot\hat{\mathbf{y}}' & \hat{\mathbf{x}}\cdot\hat{\mathbf{z}}' \\ \hat{\mathbf{y}}\cdot\hat{\mathbf{x}}' & \hat{\mathbf{y}}\cdot\hat{\mathbf{y}}' & \hat{\mathbf{y}}\cdot\hat{\mathbf{z}}' \\ \hat{\mathbf{z}}\cdot\hat{\mathbf{x}}' & \hat{\mathbf{z}}\cdot\hat{\mathbf{y}}' & \hat{\mathbf{z}}\cdot\hat{\mathbf{z}}' \end{pmatrix}. \tag{1.5.8}$$

The square matrix made up of $\langle a^{(k)}|U|a^{(l)}\rangle$ is referred to as the **transformation matrix** from the $\{|a'\rangle\}$ basis to the $\{|b'\rangle\}$ basis.

Given an arbitrary ket $|\alpha\rangle$ whose expansion coefficients $\langle a'|\alpha\rangle$ are known in the old basis,

$$|\alpha\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha\rangle, \tag{1.5.9}$$

how can we obtain $\langle b'|\alpha\rangle$, the expansion coefficients in the new basis? The answer is very simple: Just multiply (1.5.9) (with a' replaced by $a^{(l)}$ to avoid confusion) by $\langle b^{(k)}|$

$$\langle b^{(k)}|\alpha\rangle = \sum_{l} \langle b^{(k)}|a^{(l)}\rangle\langle a^{(l)}|\alpha\rangle = \sum_{l} \langle a^{(k)}|U^{\dagger}|a^{(l)}\rangle\langle a^{(l)}|\alpha\rangle. \tag{1.5.10}$$

In matrix notation, (1.5.10) states that the column matrix for $|\alpha\rangle$ in the new basis can be obtained just by applying the square matrix $U^{\dagger}$ to the column matrix in the old basis:

$$(\text{New}) = (U^{\dagger})(\text{old}). \tag{1.5.11}$$

The relationships between the old matrix elements and the new matrix elements are also easy to obtain:

$$\begin{aligned} \langle b^{(k)}|X|b^{(l)}\rangle &= \sum_{m}\sum_{n} \langle b^{(k)}|a^{(m)}\rangle\langle a^{(m)}|X|a^{(n)}\rangle\langle a^{(n)}|b^{(l)}\rangle \\ &= \sum_{m}\sum_{n} \langle a^{(k)}|U^{\dagger}|a^{(m)}\rangle\langle a^{(m)}|X|a^{(n)}\rangle\langle a^{(n)}|U|a^{(l)}\rangle. \end{aligned} \tag{1.5.12}$$

This is simply the well-known formula for a **similarity transformation** in matrix algebra,

$$X' = U^{\dagger}XU. \tag{1.5.13}$$

The **trace** of an operator X is defined as the sum of diagonal elements:

$$\text{tr}(X) = \sum_{a'} \langle a'|X|a'\rangle. \tag{1.5.14}$$

Even though a particular set of base kets is used in the definition, $\text{tr}(X)$

turns out to be independent of representation, as shown:

$$\sum_{a'}\langle a'|X|a'\rangle = \sum_{a'}\sum_{b'}\sum_{b''}\langle a'|b'\rangle\langle b'|X|b''\rangle\langle b''|a'\rangle$$
$$= \sum_{b'}\sum_{b''}\langle b''|b'\rangle\langle b'|X|b''\rangle$$
$$= \sum_{b'}\langle b'|X|b'\rangle. \tag{1.5.15}$$

We can also prove

$$\mathrm{tr}(XY) = \mathrm{tr}(YX), \tag{1.5.16a}$$
$$\mathrm{tr}(U^{\dagger}XU) = \mathrm{tr}(X), \tag{1.5.16b}$$
$$\mathrm{tr}(|a'\rangle\langle a''|) = \delta_{a'a''}, \tag{1.5.16c}$$
$$\mathrm{tr}(|b'\rangle\langle a'|) = \langle a'|b'\rangle. \tag{1.5.16d}$$

Diagonalization

So far we have not discussed how to find the eigenvalues and eigenkets of an operator B whose matrix elements in the old $\{|a'\rangle\}$ basis are assumed to be known. This problem turns out to be equivalent to that of finding the unitary matrix that diagonalizes B. Even though the reader may already be familiar with the diagonalization procedure in matrix algebra, it is worth working out this problem using the Dirac bra-ket notation.

We are interested in obtaining the eigenvalue b' and the eigenket $|b'\rangle$ with the property

$$B|b'\rangle = b'|b'\rangle. \tag{1.5.17}$$

First, we rewrite this as

$$\sum_{a'}\langle a''|B|a'\rangle\langle a'|b'\rangle = b'\langle a''|b'\rangle. \tag{1.5.18}$$

When $|b'\rangle$ in (1.5.17) stands for the lth eigenket of operator B, we can write (1.5.18) in matrix notation as follows:

$$\begin{pmatrix} B_{11} & B_{12} & B_{13} & \cdots \\ B_{21} & B_{22} & B_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} C_1^{(l)} \\ C_2^{(l)} \\ \vdots \end{pmatrix} = b^{(l)}\begin{pmatrix} C_1^{(l)} \\ C_2^{(l)} \\ \vdots \end{pmatrix}, \tag{1.5.19}$$

with

$$B_{ij} = \langle a^{(i)}|B|a^{(j)}\rangle, \tag{1.5.20a}$$

and

$$C_k^{(l)} = \langle a^{(k)}|b^{(l)}\rangle, \tag{1.5.20b}$$

where i, j, k run up to N, the dimensionality of the ket space. As we know

from linear algebra, nontrivial solutions for $C_k^{(l)}$ are possible only if the characteristic equation

$$\det(B - \lambda 1) = 0 \tag{1.5.21}$$

is satisfied. This is an Nth order algebraic equation for λ, and the N roots obtained are to be identified with the various $b^{(l)}$'s we are trying to determine. Knowing $b^{(l)}$ we can solve for the corresponding $C_k^{(l)}$'s up to an overall constant to be determined from the normalization condition. Comparing (1.5.20b) with (1.5.7), we see that the $C_k^{(l)}$'s are just the elements of the unitary matrix involved in the change of basis $\{|a'\rangle\} \rightarrow \{|b'\rangle\}$.

For this procedure the Hermiticity of B is important. For example, consider S_+ defined by (1.3.38) or (1.4.19). This operator is obviously non-Hermitian. The corresponding matrix, which reads in the S_z basis as

$$S_+ \doteq \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{1.5.22}$$

cannot be diagonalized by any unitary matrix. In Chapter 2 we will encounter eigenkets of a non-Hermitian operator in connection with a coherent state of a simple harmonic oscillator. Such eigenkets, however, are known *not* to form a complete orthonormal set, and the formalism we have developed in this section cannot be immediately applied.

Unitary Equivalent Observables

We conclude this section by discussing a remarkable theorem on the unitary transform of an observable.

Theorem. *Consider again two sets of orthonormal basis* $\{|a'\rangle\}$ *and* $\{|b'\rangle\}$ *connected by the U operator* (1.5.4). *Knowing U, we may construct a* **unitary transform** *of A,* UAU^{-1}*; then A and* UAU^{-1} *are said to be* **unitary equivalent observables.** *The eigenvalue equation for A,*

$$A|a^{(l)}\rangle = a^{(l)}|a^{(l)}\rangle, \tag{1.5.23}$$

clearly implies that

$$UAU^{-1}U|a^{(l)}\rangle = a^{(l)}U|a^{(l)}\rangle. \tag{1.5.24}$$

But this can be rewritten as

$$(UAU^{-1})|b^{(l)}\rangle = a^{(l)}|b^{(l)}\rangle. \tag{1.5.25}$$

This deceptively simple result is quite profound. It tells us that the $|b'\rangle$'s are eigenkets of UAU^{-1} with *exactly the same eigenvalues* as the A

eigenvalues. In other words, *unitary equivalent observables have identical spectra.*

The eigenket $|b^{(l)}\rangle$, by definition, satisfies the relationship

$$B|b^{(l)}\rangle = b^{(l)}|b^{(l)}\rangle. \tag{1.5.26}$$

Comparing (1.5.25) and (1.5.26), we infer that B and UAU^{-1} are simultaneously diagonalizable. A natural question is, is UAU^{-1} the same as B itself? The answer quite often is yes in cases of physical interest. Take, for example, S_x and S_z. They are related by a unitary operator, which, as we will discuss in Chapter 3, is actually the rotation operator around the y-axis by angle $\pi/2$. In this case S_x itself is the unitary transform of S_z. Because we know that S_x and S_z exhibit the same set of eigenvalues—namely, $+\hbar/2$ and $-\hbar/2$—we see that our theorem holds in this particular example.

1.6. POSITION, MOMENTUM, AND TRANSLATION

Continuous Spectra

The observables considered so far have all been assumed to exhibit discrete eigenvalue spectra. In quantum mechanics, however, there are observables with continuous eigenvalues. Take, for instance, p_z, the z-component of momentum. In quantum mechanics this is again represented by a Hermitian operator. In contrast to S_z, however, the eigenvalues of p_z (in appropriate units) can assume any real value between $-\infty$ and ∞.

The rigorous mathematics of a vector space spanned by eigenkets that exhibit a continuous spectrum is rather treacherous. The dimensionality of such a space is obviously infinite. Fortunately, many of the results we worked out for a finite-dimensional vector space with discrete eigenvalues can immediately be generalized. In places where straightforward generalizations do not hold, we indicate danger signals.

We start with the analogue of eigenvalue equation (1.2.5), which, in the continuous-spectrum case, is written as

$$\xi|\xi'\rangle = \xi'|\xi'\rangle, \tag{1.6.1}$$

where ξ is an operator and ξ' is simply a number. The ket $|\xi'\rangle$ is, in other words, an eigenket of operator ξ with eigenvalue ξ', just as $|a'\rangle$ is an eigenket of operator A with eigenvalue a'.

In pursuing this analogy we replace the Kronecker symbol by Dirac's δ-function—a discrete sum over the eigenvalues $\{a'\}$ by an integral over the

continuous variable ξ'—so

$$\langle a'|a''\rangle = \delta_{a'a''} \rightarrow \langle \xi'|\xi''\rangle = \delta(\xi' - \xi''), \tag{1.6.2a}$$

$$\sum_{a'} |a'\rangle\langle a'| = 1 \rightarrow \int d\xi' |\xi'\rangle\langle\xi'| = 1, \tag{1.6.2b}$$

$$|\alpha\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha\rangle \rightarrow |\alpha\rangle = \int d\xi' |\xi'\rangle\langle\xi'|\alpha\rangle, \tag{1.6.2c}$$

$$\sum_{a'} |\langle a'|\alpha\rangle|^2 = 1 \rightarrow \int d\xi' |\langle\xi'|\alpha\rangle|^2 = 1, \tag{1.6.2d}$$

$$\langle\beta|\alpha\rangle = \sum_{a'} \langle\beta|a'\rangle\langle a'|\alpha\rangle \rightarrow \langle\beta|\alpha\rangle = \int d\xi' \, \langle\beta|\xi'\rangle\langle\xi'|\alpha\rangle, \tag{1.6.2e}$$

$$\langle a''|A|a'\rangle = a'\delta_{a'a''} \rightarrow \langle\xi''|\xi|\xi'\rangle = \xi'\delta(\xi'' - \xi'). \tag{1.6.2f}$$

Notice in particular how the completeness relation (1.6.2b) is used to obtain (1.6.2c) and (1.6.2e).

Position Eigenkets and Position Measurements

In Section 1.4 we emphasized that a measurement in quantum mechanics is essentially a filtering process. To extend this idea to measurements of observables exhibiting continuous spectra it is best to work with a specific example. To this end we consider the position (or coordinate) operator in one dimension.

The eigenkets $|x'\rangle$ of the position operator x satisfying

$$x|x'\rangle = x'|x'\rangle \tag{1.6.3}$$

are postulated to form a complete set. Here x' is just a number with the dimension of length 0.23 cm, for example, while x is an operator. The state ket for an arbitrary physical state can be expanded in terms of $\{|x'\rangle\}$:

$$|\alpha\rangle = \int_{-\infty}^{\infty} dx' |x'\rangle\langle x'|\alpha\rangle. \tag{1.6.4}$$

We now consider a highly idealized selective measurement of the position observable. Suppose we place a very tiny detector that clicks only when the particle is precisely at x' and nowhere else. Immediately after the detector clicks, we can say that the state in question is represented by $|x'\rangle$. In other words, when the detector clicks, $|\alpha\rangle$ abruptly "jumps into" $|x'\rangle$ in much the same way as an arbitrary spin state jumps into the S_z+ (or S_z-) state when subjected to an SG apparatus of the S_z type.

In practice the best the detector can do is to locate the particle within a narrow interval around x'. A realistic detector clicks when a particle is observed to be located within some narrow range $(x' - \Delta/2, x' + \Delta/2)$.

When a count is registered in such a detector, the state ket changes abruptly as follows:

$$|\alpha\rangle = \int_{-\infty}^{\infty} dx''|x''\rangle\langle x''|\alpha\rangle \xrightarrow{\text{measurement}} \int_{x'-\Delta/2}^{x'+\Delta/2} dx''|x''\rangle\langle x''|\alpha\rangle. \quad (1.6.5)$$

Assuming that $\langle x''|\alpha\rangle$ does not change appreciably within the narrow interval, the probability for the detector to click is given by

$$|\langle x'|\alpha\rangle|^2 dx', \quad (1.6.6)$$

where we have written dx' for Δ. This is analogous to $|\langle a'|\alpha\rangle|^2$ for the probability for $|\alpha\rangle$ to be thrown into $|a'\rangle$ when A is measured. The probability of recording the particle *somewhere* between $-\infty$ and ∞ is given by

$$\int_{-\infty}^{\infty} dx'|\langle x'|\alpha\rangle|^2, \quad (1.6.7)$$

which is normalized to unity if $|\alpha\rangle$ is normalized:

$$\langle\alpha|\alpha\rangle = 1 \Rightarrow \int_{-\infty}^{\infty} dx'\langle\alpha|x'\rangle\langle x'|\alpha\rangle = 1. \quad (1.6.8)$$

The reader familiar with wave mechanics may have recognized by this time that $\langle x'|\alpha\rangle$ is the wave function for the physical state represented by $|\alpha\rangle$. We will say more about this identification of the expansion coefficient with the x-representation of the wave function in Section 1.7.

The notion of a position eigenket can be extended to three dimensions. It is assumed in nonrelativistic quantum mechanics that the position eigenkets $|\mathbf{x}'\rangle$ are complete. The state ket for a particle with internal degrees of freedom, such as spin, ignored can therefore be expanded in terms of $\{|\mathbf{x}'\rangle\}$ as follows:

$$|\alpha\rangle = \int d^3x'|\mathbf{x}'\rangle\langle\mathbf{x}'|\alpha\rangle, \quad (1.6.9)$$

where $\mathbf{x}'$ stands for x', y', and z'; in other words, $|\mathbf{x}'\rangle$ is a *simultaneous* eigenket of the observables x, y, and z in the sense of Section 1.4:

$$|\mathbf{x}'\rangle \equiv |x', y', z'\rangle, \quad (1.6.10a)$$

$$x|\mathbf{x}'\rangle = x'|\mathbf{x}'\rangle, \qquad y|\mathbf{x}'\rangle = y'|\mathbf{x}'\rangle, \qquad z|\mathbf{x}'\rangle = z'|\mathbf{x}'\rangle. \quad (1.6.10b)$$

To be able to consider such a simultaneous eigenket at all, we are implicitly assuming that the three components of the position vector can be measured simultaneously to arbitrary degrees of accuracy; hence, we must have

$$[x_i, x_j] = 0, \quad (1.6.11)$$

where x_1, x_2, and x_3 stand for x, y, and z, respectively.

Translation

We now introduce the very important concept of translation, or spatial displacement. Suppose we start with a state that is well localized around $\mathbf{x}'$. Let us consider an operation that changes this state into another well-localized state, this time around $\mathbf{x}'+d\mathbf{x}'$ with everything else (for example, the spin direction) unchanged. Such an operation is defined to be an **infinitesimal translation** by $d\mathbf{x}'$, and the operator that does the job is denoted by $\mathcal{T}(d\mathbf{x}')$:

$$\mathcal{T}(d\mathbf{x}')|\mathbf{x}'\rangle = |\mathbf{x}'+d\mathbf{x}'\rangle, \tag{1.6.12}$$

where a possible arbitrary phase factor is set to unity by convention. Notice that the right-hand side of (1.6.12) is again a position eigenket, but this time with eigenvalue $\mathbf{x}'+\mathrm{d}\mathbf{x}'$. Obviously $|\mathbf{x}'\rangle$ is *not* an eigenket of the infinitesimal translation operator.

By expanding an arbitrary state ket $|\alpha\rangle$ in terms of the position eigenkets we can examine the effect of infinitesimal translation on $|\alpha\rangle$:

$$|\alpha\rangle \rightarrow \mathcal{T}(d\mathbf{x}')|\alpha\rangle = \mathcal{T}(d\mathbf{x}')\int d^3x'|\mathbf{x}'\rangle\langle\mathbf{x}'|\alpha\rangle = \int d^3x'|\mathbf{x}'+d\mathbf{x}'\rangle\langle\mathbf{x}'|\alpha\rangle. \tag{1.6.13}$$

We also write the right-hand side of (1.6.13) as

$$\int d^3x'|\mathbf{x}'+d\mathbf{x}'\rangle\langle\mathbf{x}'|\alpha\rangle = \int d^3x'|\mathbf{x}'\rangle\langle\mathbf{x}'-d\mathbf{x}'|\alpha\rangle \tag{1.6.14}$$

because the integration is over all space and $\mathbf{x}'$ is just an integration variable. This shows that the wave function of the translated state $\mathcal{T}(d\mathbf{x}')|\alpha\rangle$ is obtained by substituting $\mathbf{x}'-d\mathbf{x}'$ for $\mathbf{x}'$ in $\langle\mathbf{x}'|\alpha\rangle$.

There is an equivalent approach to translation that is often treated in the literature. Instead of considering an infinitesimal translation of the physical system itself, we consider a change in the coordinate system being used such that the origin is shifted in the *opposite* direction, $-d\mathbf{x}'$. Physically, in this alternative approach we are asking how the *same* state ket would look to another observer whose coordinate system is shifted by $-d\mathbf{x}'$. In this book we try not to use this approach. Obviously it is important that we do not mix the two approaches!

We now list the properties of the infinitesimal translation operator $\mathcal{T}(d\mathbf{x}')$. The first property we demand is the unitarity property imposed by probability conservation. It is reasonable to require that if the ket $|\alpha\rangle$ is normalized to unity, the translated ket $\mathcal{T}(d\mathbf{x}')|\alpha\rangle$ also be normalized to unity, so

$$\langle\alpha|\alpha\rangle = \langle\alpha|\mathcal{T}^\dagger(d\mathbf{x}')\mathcal{T}(d\mathbf{x}')|\alpha\rangle. \tag{1.6.15}$$

This condition is guaranteed by demanding that the infinitesimal translation

be unitary:

$$\mathscr{T}^{\dagger}(d\mathbf{x}')\mathscr{T}(d\mathbf{x}')=1. \tag{1.6.16}$$

Quite generally, the norm of a ket is preserved under unitary transformations. For the second property, suppose we consider two successive infinitesimal translations—first by $d\mathbf{x}'$ and subsequently by $d\mathbf{x}''$, where $d\mathbf{x}'$ and $d\mathbf{x}''$ need not be in the same direction. We expect the net result to be just a *single* translation operation by the vector sum $d\mathbf{x}'+d\mathbf{x}''$, so we demand that

$$\mathscr{T}(d\mathbf{x}'')\mathscr{T}(d\mathbf{x}')=\mathscr{T}(d\mathbf{x}'+d\mathbf{x}''). \tag{1.6.17}$$

For the third property, suppose we consider a translation in the opposite direction; we expect the opposite-direction translation to be the same as the inverse of the original translation:

$$\mathscr{T}(-d\mathbf{x}')=\mathscr{T}^{-1}(d\mathbf{x}'). \tag{1.6.18}$$

For the fourth property, we demand that as $d\mathbf{x}'\to 0$, the translation operation reduce to the identity operation

$$\lim_{d\mathbf{x}'\to 0}\mathscr{T}(d\mathbf{x}')=1 \tag{1.6.19}$$

and that the difference between $\mathscr{T}(d\mathbf{x}')$ and the identity operator be of first order in $d\mathbf{x}'$.

We now demonstrate that if we take the infinitesimal translation operator to be

$$\mathscr{T}(d\mathbf{x}')=1-i\mathbf{K}\cdot d\mathbf{x}', \tag{1.6.20}$$

where the components of $\mathbf{K}$, K_x, K_y, and K_z, are **Hermitian operators**, then all the properties listed are satisfied. The first property, the unitarity of $\mathscr{T}(\mathrm{d}\mathbf{x}')$, is checked as follows:

$$\begin{aligned}\mathscr{T}^{\dagger}(d\mathbf{x}')\mathscr{T}(d\mathbf{x}')&=(1+i\mathbf{K}^{\dagger}\cdot d\mathbf{x}')(1-i\mathbf{K}\cdot d\mathbf{x}')\\&=1-i(\mathbf{K}-\mathbf{K}^{\dagger})\cdot d\mathbf{x}'+0\left[(d\mathbf{x}')^2\right]\\&\simeq 1,\end{aligned} \tag{1.6.21}$$

where terms of second order in $d\mathbf{x}'$ have been ignored for an infinitesimal translation. The second property [(1.6.17)] can also be proved as follows:

$$\begin{aligned}\mathscr{T}(d\mathbf{x}'')\mathscr{T}(d\mathbf{x}')&=(1-i\mathbf{K}\cdot d\mathbf{x}'')(1-i\mathbf{K}\cdot d\mathbf{x}')\\&\simeq 1-i\mathbf{K}\cdot(d\mathbf{x}'+d\mathbf{x}'')\\&=\mathscr{T}(d\mathbf{x}'+d\mathbf{x}'').\end{aligned} \tag{1.6.22}$$

The third and fourth properties are obviously satisfied by (1.6.20).

Accepting (1.6.20) to be the correct form for $\mathscr{T}(d\mathbf{x}')$, we are in a position to derive an extremely fundamental relation between the $\mathbf{K}$ oper-

ator and the $\mathbf{x}$ operator. First, note that

$$\mathbf{x}\mathscr{T}(d\mathbf{x}')|\mathbf{x}'\rangle = \mathbf{x}|\mathbf{x}'+d\mathbf{x}'\rangle = (\mathbf{x}'+d\mathbf{x}')|\mathbf{x}'+d\mathbf{x}'\rangle \quad (1.6.23a)$$

and

$$\mathscr{T}(d\mathbf{x}')\mathbf{x}|\mathbf{x}'\rangle = \mathbf{x}'\mathscr{T}(d\mathbf{x}')|\mathbf{x}'\rangle = \mathbf{x}'|\mathbf{x}'+d\mathbf{x}'\rangle; \quad (1.6.23b)$$

hence,

$$\left[\mathbf{x}, \mathscr{T}(d\mathbf{x}')\right]|\mathbf{x}'\rangle = d\mathbf{x}'|\mathbf{x}'+d\mathbf{x}'\rangle \simeq d\mathbf{x}'|\mathbf{x}'\rangle, \quad (1.6.24)$$

where the error made in writing the last line of (1.6.24) is of second order in $d\mathbf{x}'$. Now $|\mathbf{x}'\rangle$ can be *any* position eigenket, and the position eigenkets are known to form a complete set. We must therefore have an **operator identity**

$$\left[\mathbf{x}, \mathscr{T}(d\mathbf{x}')\right] = d\mathbf{x}', \quad (1.6.25)$$

or

$$-i\mathbf{x}\mathbf{K}\cdot d\mathbf{x}' + i\mathbf{K}\cdot d\mathbf{x}'\mathbf{x} = d\mathbf{x}', \quad (1.6.26)$$

where on the right-hand sides of (1.6.25) and (1.6.26) $d\mathbf{x}'$ is understood to be the number $d\mathbf{x}'$ multiplied by the identity operator in the ket space spanned by $|\mathbf{x}'\rangle$. By choosing $d\mathbf{x}'$ in the direction of $\hat{\mathbf{x}}_j$ and forming the scalar product with $\hat{\mathbf{x}}_i$, we obtain

$$\left[x_i, K_j\right] = i\delta_{ij}, \quad (1.6.27)$$

where again δ_{ij} is understood to be multiplied by the identity operator.

Momentum as a Generator of Translation

Equation (1.6.27) is the fundamental commutation relation between the position operators x, y, z and the K operators K_x, K_y, K_z. Remember that so far the K operator is *defined* in terms of the infinitesimal translation operator by (1.6.20). What is the physical significance we can attach to $\mathbf{K}$?

J. Schwinger, lecturing on quantum mechanics, once remarked, "... for fundamental properties we will borrow only names from classical physics." In the present case we would like to borrow from classical mechanics the notion that momentum is the generator of an infinitesimal translation. An infinitesimal translation in classical mechanics can be regarded as a canonical transformation,

$$\mathbf{x}_{\text{new}} \equiv \mathbf{X} = \mathbf{x} + d\mathbf{x}, \qquad \mathbf{p}_{\text{new}} \equiv \mathbf{P} = \mathbf{p}, \quad (1.6.28)$$

obtainable from the generating function (Goldstein 1980, 395 and 411)

$$F(\mathbf{x},\mathbf{P}) = \mathbf{x}\cdot\mathbf{P} + \mathbf{p}\cdot d\mathbf{x}, \quad (1.6.29)$$

where $\mathbf{p}$ and $\mathbf{P}$ refer to the corresponding momenta.

This equation has a striking similarity to the infinitesimal translation operator (1.6.20) in quantum mechanics, particularly if we recall that $\mathbf{x}\cdot\mathbf{P}$ in

(1.6.29) is the generating function for the identity transformation ($\mathbf{X} = \mathbf{x}, \mathbf{P} = \mathbf{p}$). We are therefore led to speculate that the operator $\mathbf{K}$ is in some sense related to the momentum operator in quantum mechanics.

Can the K operator be identified with the momentum operator itself? Unfortunately the dimension is all wrong; the K operator has the dimension of 1/length because $\mathbf{K}\cdot d\mathbf{x}'$ must be dimensionless. But it appears legitimate to set

$$\mathbf{K} = \frac{\mathbf{p}}{\text{universal constant with the dimension of action}}. \tag{1.6.30}$$

From the fundamental postulates of quantum mechanics there is no way to determine the actual numerical value of the universal constant. Rather, this constant is needed here because, historically, classical physics was developed before quantum mechanics using units convenient for describing macroscopic quantities—the circumference of the earth, the mass of 1 cc of water, the duration of a mean solar day, and so forth. Had microscopic physics been formulated before macroscopic physics, the physicists would have almost certainly chosen the basic units in such a way that the universal constant appearing in (1.6.30) would be unity.

An analogy from electrostatics may be helpful here. The interaction energy between two particles of charge e separated at a distance r is proportional to e^2/r; in unrationalized Gaussian units, the proportionality factor is just 1, but in rationalized mks units, which may be more convenient for electrical engineers, the proportionality factor is $1/4\pi\varepsilon_o$.

The universal constant that appears in (1.6.30) turns out to be the same as the constant $\hbar$ that appears in L. de Broglie's relation, written in 1924,

$$\frac{2\pi}{\lambda} = \frac{p}{\hbar}, \tag{1.6.31}$$

where λ is the wavelength of a "particle wave." In other words, the K operator is the quantum mechanical operator that corresponds to the wave number—that is, 2π times the reciprocal wavelength, usually denoted by k. With this identification the infinitesimal translation operator $\mathscr{T}(d\mathbf{x}')$ reads

$$\mathscr{T}(d\mathbf{x}') = 1 - i\mathbf{p}\cdot d\mathbf{x}'/\hbar, \tag{1.6.32}$$

where $\mathbf{p}$ is the momentum operator. The commutation relation (1.6.27) now becomes

$$[x_i, p_j] = i\hbar\delta_{ij}. \tag{1.6.33}$$

The commutation relations (1.6.33) imply, for example, that x and p_x (but not x and p_y) are incompatible observables. It is therefore impossible to find simultaneous eigenkets of x and p_x. The general formalism of

Section 1.4 can be applied here to obtain the **position-momentum uncertainty relation** of W. Heisenberg:

$$\langle(\Delta x)^2\rangle\langle(\Delta p_x)^2\rangle \geq \hbar^2/4. \tag{1.6.34}$$

Some applications of (1.6.34) will appear in Section 1.7.

So far we have concerned ourselves with infinitesimal translations. A finite translation—that is, a spatial displacement by a finite amount—can be obtained by successively compounding infinitesimal translations. Let us consider a finite translation in the x-direction by an amount $\Delta x'$:

$$\mathscr{T}(\Delta x'\hat{\mathbf{x}})|\mathbf{x}'\rangle = |\mathbf{x}' + \Delta x'\hat{\mathbf{x}}\rangle. \tag{1.6.35}$$

By compounding N infinitesimal translations, each of which is characterized by a spatial displacement $\Delta x'/N$ in the x-direction, and letting $N \to \infty$, we obtain

$$\begin{aligned}\mathscr{T}(\Delta x'\hat{\mathbf{x}}) &= \lim_{N\to\infty}\left(1 - \frac{ip_x\Delta x'}{N\hbar}\right)^N \\ &= \exp\left(-\frac{ip_x\Delta x'}{\hbar}\right).\end{aligned} \tag{1.6.36}$$

Here $\exp(-ip_x\Delta x'/\hbar)$ is understood to be a function of the *operator* p_x; generally, for any operator X we have

$$\exp(X) \equiv 1 + X + \frac{X^2}{2!} + \cdots. \tag{1.6.37}$$

A fundamental property of translations is that successive translations in different directions, say in the x- and y-directions, commute. We see this clearly in Figure 1.8; in shifting from A and B it does not matter whether we go via C or via D. Mathematically,

$$\begin{aligned}\mathscr{T}(\Delta y'\hat{\mathbf{y}})\mathscr{T}(\Delta x'\hat{\mathbf{x}}) &= \mathscr{T}(\Delta x'\hat{\mathbf{x}} + \Delta y'\hat{\mathbf{y}}), \\ \mathscr{T}(\Delta x'\hat{\mathbf{x}})\mathscr{T}(\Delta y'\hat{\mathbf{y}}) &= \mathscr{T}(\Delta x'\hat{\mathbf{x}} + \Delta y'\hat{\mathbf{y}}).\end{aligned} \tag{1.6.38}$$

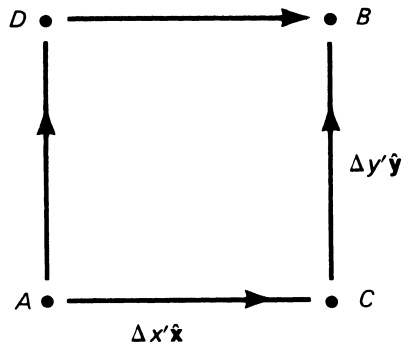


FIGURE 1.8. Successive translations in different directions.

This point is not so trivial as it may appear; we will show in Chapter 3 that rotations about different axes do *not* commute. Treating $\Delta x'$ and $\Delta y'$ up to second order, we obtain

$$\begin{aligned}\left[\mathscr{T}(\Delta y'\hat{\mathbf{y}}), \mathscr{T}(\Delta x'\hat{\mathbf{x}})\right] = &\left[\left(1 - \frac{ip_y \Delta y'}{\hbar} - \frac{p_y^2(\Delta y')^2}{2\hbar^2} + \cdots\right),\right. \\ &\left.\left(1 - \frac{ip_x \Delta x'}{\hbar} - \frac{p_x^2(\Delta x')^2}{2\hbar^2} + \cdots\right)\right] \\ \simeq &-\frac{(\Delta x')(\Delta y')[p_y, p_x]}{\hbar^2}. \end{aligned} \tag{1.6.39}$$

Because $\Delta x'$ and $\Delta y'$ are arbitrary, requirement (1.6.38), or

$$\left[\mathscr{T}(\Delta y'\hat{\mathbf{y}}), \mathscr{T}(\Delta x'\hat{\mathbf{x}})\right] = 0, \tag{1.6.40}$$

immediately leads to

$$[p_x, p_y] = 0, \tag{1.6.41}$$

or, more generally,

$$[p_i, p_j] = 0. \tag{1.6.42}$$

This commutation relation is a direct consequence of the fact that translations in different directions commute. Whenever the generators of transformations commute, the corresponding group is said to be **Abelian**. The translation group in three dimensions is Abelian.

Equation (1.6.42) implies that p_x, p_y, and p_z are mutually compatible observables. We can therefore conceive of a simultaneous eigenket of p_x, p_y, p_z, namely,

$$|\mathbf{p}'\rangle \equiv |p_x', p_y', p_z'\rangle, \tag{1.6.43a}$$

$$p_x|\mathbf{p}'\rangle = p_x'|\mathbf{p}'\rangle,\ p_y|\mathbf{p}'\rangle = p_y'|\mathbf{p}'\rangle,\ p_z|\mathbf{p}'\rangle = p_z'|\mathbf{p}'\rangle. \tag{1.6.43b}$$

It is instructive to work out the effect of $\mathscr{T}(d\mathbf{x}')$ on such a momentum eigenket:

$$\mathscr{T}(d\mathbf{x}')|\mathbf{p}'\rangle = \left(1 - \frac{i\mathbf{p}\cdot d\mathbf{x}'}{\hbar}\right)|\mathbf{p}'\rangle = \left(1 - \frac{i\mathbf{p}'\cdot d\mathbf{x}'}{\hbar}\right)|\mathbf{p}'\rangle. \tag{1.6.44}$$

We see that the momentum eigenket remains the same even though it suffers a slight phase change, so unlike $|\mathbf{x}'\rangle$, $|\mathbf{p}'\rangle$ *is* an eigenket of $\mathscr{T}(d\mathbf{x}')$, which we anticipated because

$$\left[\mathbf{p}, \mathscr{T}(d\mathbf{x}')\right] = 0. \tag{1.6.45}$$

Notice, however, that the eigenvalue of $\mathscr{T}(d\mathbf{x}')$ is complex; we do not

expect a real eigenvalue here because $\mathscr{T}(d\mathbf{x}')$, though unitary, is not Hermitian.

The Canonical Commutation Relations

We summarize the commutator relations we inferred by studying the properties of translation:

$$[x_i, x_j] = 0, \qquad [p_i, p_j] = 0, \qquad [x_i, p_j] = i\hbar\delta_{ij}. \tag{1.6.46}$$

These relations form the cornerstone of quantum mechanics; in his book, P. A. M. Dirac calls them the "fundamental quantum conditions." More often they are known as the **canonical commutation relations**, or the **fundamental commutation relations**.

Historically it was W. Heisenberg who, in 1925, showed that the combination rule for atomic transition lines known at that time could best be understood if one associated arrays of numbers obeying certain multiplication rules with these frequencies. Immediately afterward M. Born and P. Jordan pointed out that Heisenberg's multiplication rules are essentially those of matrix algebra, and a theory was developed based on the matrix analogues of (1.6.46), which is now known as **matrix mechanics**.*

Also in 1925, P. A. M. Dirac observed that the various quantum-mechanical relations can be obtained from the corresponding classical relations just by replacing classical Poisson brackets by commutators, as follows:

$$[\ ,\ ]_{\text{classical}} \to \frac{[\ ,\ ]}{i\hbar}, \tag{1.6.47}$$

where we may recall that the classical Poisson brackets are defined for functions of q's and p's as

$$[A(q,p), B(q,p)]_{\text{classical}} \equiv \sum_s \left(\frac{\partial A}{\partial q_s}\frac{\partial B}{\partial p_s} - \frac{\partial A}{\partial p_s}\frac{\partial B}{\partial q_s} \right). \tag{1.6.48}$$

For example, in classical mechanics, we have

$$[x_i, p_j]_{\text{classical}} = \delta_{ij}, \tag{1.6.49}$$

which in quantum mechanics turns into (1.6.33).

Dirac's rule (1.6.47) is plausible because the classical Poisson brackets and quantum-mechanical commutators satisfy similar algebraic properties. In particular, the following relations can be proved regardless of whether [,] is understood as a classical Poisson bracket or as a quantum-

*Appropriately, $pq - qp = h/2\pi i$ is inscribed on the gravestone of M. Born in Göttingen.

mechanical commutator:

$$[A, A] = 0 \tag{1.6.50a}$$

$$[A, B] = -[B, A] \tag{1.6.50b}$$

$$[A, c] = 0 \qquad (c \text{ is just a number}) \tag{1.6.50c}$$

$$[A + B, C] = [A, C] + [B, C] \tag{1.6.50d}$$

$$[A, BC] = [A, B]C + B[A, C] \tag{1.6.50e}$$

$$[A,[B, C]] + [B,[C, A]] + [C,[A, B]] = 0, \tag{1.6.50f}$$

where the last relation is known as the **Jacobi identity.*** However, there are important differences. First, the dimension of the classical Poisson bracket differs from that of the quantum-mechanical commutator because of the differentiations with respect to q and p appearing in (1.6.48). Second, the Poisson bracket of real functions of q's and p's is purely real, while the commutator of two Hermitian operators is anti-Hermitian (see Lemma 3 of Section 1.4). To take care of these differences the factor $i\hbar$ is inserted in (1.6.47).

We have deliberately avoided exploiting Dirac's analogy in obtaining the canonical commutation relations. Our approach to the commutation relations is based solely on (1) the properties of translations and (2) the identification of the generator of translation with the momentum operator modulo a universal constant with the dimension of action. We believe that this approach is more powerful because it can be generalized to situations where observables have no classical analogues. For example, the spin-angular-momentum components we encountered in Section 1.4 having nothing to do with the p's and q's of classical mechanics; yet, as we will show in Chapter 3, the spin-angular-momentum commutation relations can be derived using the properties of rotations just as we derived the canonical commutation relations using the properties of translations.

1.7. WAVE FUNCTIONS IN POSITION AND MOMENTUM SPACE

Position-Space Wave Function

In this section we present a systematic study of the properties of wave functions in both position and momentum space. For simplicity let us return to the one-dimensional case. The base kets used are the position kets

*It is amusing that the Jacobi identity in quantum mechanics is much easier to prove than its classical analogue.

satisfying

$$x|x'\rangle = x'|x'\rangle, \tag{1.7.1}$$

normalized in such a way that the orthogonality condition reads

$$\langle x''|x'\rangle = \delta(x'' - x'). \tag{1.7.2}$$

We have already remarked that the ket representing a physical state can be expanded in terms of $|x'\rangle$,

$$|\alpha\rangle = \int dx' |x'\rangle\langle x'|\alpha\rangle, \tag{1.7.3}$$

and that the expansion coefficient $\langle x'|\alpha\rangle$ is interpreted in such a way that

$$|\langle x'|\alpha\rangle|^2 dx' \tag{1.7.4}$$

is the probability for the particle to be found in a narrow interval dx' around x'. In our formalism the inner product $\langle x'|\alpha\rangle$ is what is usually referred to as the **wave function** $\psi_\alpha(x')$ for state $|\alpha\rangle$:

$$\langle x'|\alpha\rangle = \psi_\alpha(x'). \tag{1.7.5}$$

In elementary wave mechanics the probabilistic interpretations for the expansion coefficient $c_{a'}$ $(= \langle a'|\alpha\rangle)$ and for the wave function $\psi_\alpha(x')$ $(= \langle x'|\alpha\rangle)$ are often presented as separate postulates. One of the major advantages of our formalism, originally due to Dirac, is that the two kinds of probabilistic interpretations are unified; $\psi_\alpha(x')$ *is* an expansion coefficient [see (1.7.3)] in much the same way as $c_{a'}$ is. By following the footsteps of Dirac we come to appreciate the unity of quantum mechanics.

Consider the inner product $\langle\beta|\alpha\rangle$. Using the completeness of $|x'\rangle$, we have

$$\begin{aligned}\langle\beta|\alpha\rangle &= \int dx' \langle\beta|x'\rangle\langle x'|\alpha\rangle \\ &= \int dx' \psi_\beta^*(x')\psi_\alpha(x'),\end{aligned} \tag{1.7.6}$$

so $\langle\beta|\alpha\rangle$ characterizes the overlap between the two wave functions. Note that we are not defining $\langle\beta|\alpha\rangle$ as the overlap integral; the identification of $\langle\beta|\alpha\rangle$ with the overlap integral *follows* from our completeness postulate for $|x'\rangle$. The more general interpretation of $\langle\beta|\alpha\rangle$, *independent of representations*, is that it represents the probability amplitude for state $|\alpha\rangle$ to be found in state $|\beta\rangle$.

This time let us interpret the expansion

$$|\alpha\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha\rangle \tag{1.7.7}$$

using the language of wave functions. We just multiply both sides of (1.7.7)

by the position eigenbra $\langle x'|$ on the left. Thus

$$\langle x'|\alpha\rangle = \sum_{a'} \langle x'|a'\rangle\langle a'|\alpha\rangle. \tag{1.7.8}$$

In the usual notation of wave mechanics this is recognized as

$$\psi_\alpha(x') = \sum_{a'} c_{a'} u_{a'}(x'),$$

where we have introduced an **eigenfunction** of operator A with eigenvalue a':

$$u_{a'}(x') = \langle x'|a'\rangle. \tag{1.7.9}$$

Let us now examine how $\langle\beta|A|\alpha\rangle$ can be written using the wave functions for $|\alpha\rangle$ and $|\beta\rangle$. Clearly, we have

$$\begin{aligned}\langle\beta|A|\alpha\rangle &= \int dx' \int dx'' \langle\beta|x'\rangle\langle x'|A|x''\rangle\langle x''|\alpha\rangle \\ &= \int dx' \int dx'' \psi_\beta^*(x')\langle x'|A|x''\rangle\psi_\alpha(x'').\end{aligned} \tag{1.7.10}$$

So to be able to evaluate $\langle\beta|A|\alpha\rangle$, we must know the matrix element $\langle x'|A|x''\rangle$, which is, in general, a function of the two variables x' and x''.

An enormous simplification takes place if observable A is a function of the position operator x. In particular, consider

$$A = x^2, \tag{1.7.11}$$

which actually appears in the Hamiltonian for the simple harmonic oscillator problem to be discussed in Chapter 2. We have

$$\langle x'|x^2|x''\rangle = \big(\langle x'|\big)\cdot\big(x''^2|x''\rangle\big) = x'^2\delta(x'-x''), \tag{1.7.12}$$

where we have used (1.7.1) and (1.7.2). The double integral (1.7.10) is now reduced to a *single* integral:

$$\begin{aligned}\langle\beta|x^2|\alpha\rangle &= \int dx' \langle\beta|x'\rangle x'^2\langle x'|\alpha\rangle \\ &= \int dx' \psi_\beta^*(x') x'^2 \psi_\alpha(x').\end{aligned} \tag{1.7.13}$$

In general,

$$\langle\beta|f(x)|\alpha\rangle = \int dx' \psi_\beta^*(x') f(x') \psi_\alpha(x'). \tag{1.7.14}$$

Note that the $f(x)$ on the left-hand side of (1.7.14) is an operator, while the $f(x')$ on the right-hand side is not an operator.

Momentum Operator in the Position Basis

We now examine how the momentum operator may look in the x-basis—that is, in the representation where the position eigenkets are used as base kets. Our starting point is the definition of momentum as the generator of infinitesimal translations:

$$\left(1-\frac{ip\Delta x'}{\hbar}\right)|\alpha\rangle = \int dx'\mathscr{T}(\Delta x')|x'\rangle\langle x'|\alpha\rangle$$

$$= \int dx'|x'+\Delta x'\rangle\langle x'|\alpha\rangle$$

$$= \int dx'|x'\rangle\langle x'-\Delta x'|\alpha\rangle$$

$$= \int dx'|x'\rangle\left(\langle x'|\alpha\rangle - \Delta x'\frac{\partial}{\partial x'}\langle x'|\alpha\rangle\right). \quad (1.7.15)$$

Comparison of both sides yields

$$p|\alpha\rangle = \int dx'|x'\rangle\left(-i\hbar\frac{\partial}{\partial x'}\langle x'|\alpha\rangle\right) \quad (1.7.16)$$

or

$$\langle x'|p|\alpha\rangle = -i\hbar\frac{\partial}{\partial x'}\langle x'|\alpha\rangle, \quad (1.7.17)$$

where we have used the orthogonality property (1.7.2). For the matrix element p in the x-representation, we obtain

$$\langle x'|p|x''\rangle = -i\hbar\frac{\partial}{\partial x'}\delta(x'-x''). \quad (1.7.18)$$

From (1.7.16) we get a very important identity:

$$\langle\beta|p|\alpha\rangle = \int dx'\langle\beta|x'\rangle\left(-i\hbar\frac{\partial}{\partial x'}\langle x'|\alpha\rangle\right)$$

$$= \int dx'\psi_\beta^*(x')\left(-i\hbar\frac{\partial}{\partial x'}\right)\psi_\alpha(x'). \quad (1.7.19)$$

In our formalism (1.7.19) is not a postulate; rather, it has been *derived* using the basic properties of momentum. By repeatedly applying (1.7.17), we can also obtain

$$\langle x'|p^n|\alpha\rangle = (-i\hbar)^n\frac{\partial^n}{\partial x'^n}\langle x'|\alpha\rangle, \quad (1.7.20)$$

$$\langle\beta|p^n|\alpha\rangle = \int dx'\psi_\beta^*(x')(-i\hbar)^n\frac{\partial^n}{\partial x'^n}\psi_\alpha(x'). \quad (1.7.21)$$

Momentum-Space Wave Function

So far we have worked exclusively in the x-basis. There is actually a complete symmetry between x and p—apart from occasional minus signs—which we can infer from the canonical commutation relations. Let us now work in the p-basis, that is, in the momentum representation.

For simplicity we continue working in one-space. The base eigenkets in the p-basis specify

$$p|p'\rangle = p'|p'\rangle \tag{1.7.22}$$

and

$$\langle p'|p''\rangle = \delta(p' - p''). \tag{1.7.23}$$

The momentum eigenkets $\{|p'\rangle\}$ span the ket space in much the same way as the position eigenkets $\{|x'\rangle\}$. An arbitrary state ket $|\alpha\rangle$ can therefore be expanded as follows:

$$|\alpha\rangle = \int dp' |p'\rangle\langle p'|\alpha\rangle. \tag{1.7.24}$$

We can give a probabilistic interpretation for the expansion coefficient $\langle p'|\alpha\rangle$; the probability that a measurement of p gives eigenvalue p' within a narrow interval dp' is $|\langle p'|\alpha\rangle|^2 dp'$. It is customary to call $\langle p'|\alpha\rangle$ the **momentum-space wave function**; the notation $\phi_\alpha(p')$ is often used:

$$\langle p'|\alpha\rangle = \phi_\alpha(p'). \tag{1.7.25}$$

If $|\alpha\rangle$ is normalized, we obtain

$$\int dp' \langle\alpha|p'\rangle\langle p'|\alpha\rangle = \int dp' |\phi_\alpha(p')|^2 = 1. \tag{1.7.26}$$

Let us now establish the connection between the x-representation and the p-representation. We recall that in the case of the discrete spectra, the change of basis from the old set $\{|a'\rangle\}$ to the new set $\{|b'\rangle\}$ is characterized by the transformation matrix (1.5.7). Likewise, we expect that the desired information is contained in $\langle x'|p'\rangle$, which is a function of x' and p', usually called the **transformation function** from the x-representation to the p-representation. To derive the explicit form of $\langle x'|p'\rangle$, first recall (1.7.17); letting $|\alpha\rangle$ be the momentum eigenket $|p'\rangle$, we obtain

$$\langle x'|p|p'\rangle = -i\hbar \frac{\partial}{\partial x'} \langle x'|p'\rangle \tag{1.7.27}$$

or

$$p'\langle x'|p'\rangle = -i\hbar \frac{\partial}{\partial x'} \langle x'|p'\rangle. \tag{1.7.28}$$

The solution to this differential equation for $\langle x'|p'\rangle$ is

$$\langle x'|p'\rangle = N\exp\left(\frac{ip'x'}{\hbar}\right), \tag{1.7.29}$$

where N is the normalization constant to be determined in a moment. Even though the transformation function $\langle x'|p'\rangle$ is a function of two variables, x' and p', we can temporarily regard it as a function of x' with p' fixed. It can then be viewed as the probability amplitude for the momentum eigenstate specified by p' to be found at position x'; in other words, it is just the wave function for the momentum eigenstate $|p'\rangle$, often referred to as the momentum eigenfunction (still in the x-space). So (1.7.29) simply says that the wave function of a momentum eigenstate is a plane wave. It is amusing that we have obtained this plane-wave solution without solving the Schrödinger equation (which we have not yet written down).

To get the normalization constant N let us first consider

$$\langle x'|x''\rangle = \int dp'\langle x'|p'\rangle\langle p'|x''\rangle. \tag{1.7.30}$$

The left-hand side is just $\delta(x'-x'')$; the right-hand side can be evaluated using the explicit form of $\langle x'|p'\rangle$:

$$\begin{aligned}\delta(x'-x'') &= |N|^2\int dp'\exp\left[\frac{ip'(x'-x'')}{\hbar}\right] \\ &= 2\pi\hbar|N|^2\delta(x'-x'').\end{aligned} \tag{1.7.31}$$

Choosing N to be purely real and positive by convention, we finally have

$$\langle x'|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp\left(\frac{ip'x'}{\hbar}\right). \tag{1.7.32}$$

We can now demonstrate how the position-space wave function is related to the momentum-space wave function. All we have to do is rewrite

$$\langle x'|\alpha\rangle = \int dp'\langle x'|p'\rangle\langle p'|\alpha\rangle \tag{1.7.33a}$$

and

$$\langle p'|\alpha\rangle = \int dx'\langle p'|x'\rangle\langle x'|\alpha\rangle \tag{1.7.33b}$$

as

$$\psi_\alpha(x') = \left[\frac{1}{\sqrt{2\pi\hbar}}\right]\int dp'\exp\left(\frac{ip'x'}{\hbar}\right)\phi_\alpha(p') \tag{1.7.34a}$$

and

$$\phi_\alpha(p') = \left[\frac{1}{\sqrt{2\pi\hbar}}\right]\int dx'\exp\left(\frac{-ip'x'}{\hbar}\right)\psi_\alpha(x'). \tag{1.7.34b}$$

This pair of equations is just what one expects from Fourier's inversion theorem. Apparently the mathematics we have developed somehow "knows" Fourier's work on integral transforms.

Gaussian Wave Packets

It is instructive to look at a physical example to illustrate our basic formalism. We consider what is known as a **Gaussian wave packet**, whose x-space wave function is given by

$$\langle x'|\alpha\rangle = \left[\frac{1}{\pi^{1/4}\sqrt{d}}\right]\exp\left[ikx' - \frac{x'^2}{2d^2}\right]. \tag{1.7.35}$$

This is a plane wave with wave number k modulated by a Gaussian profile centered on the origin. The probability of observing the particle vanishes very rapidly for $|x'| > d$; more quantitatively, the probability density $|\langle x'|\alpha\rangle|^2$ has a Gaussian shape with width d.

We now compute the expectation values of x, x^2, p, and p^2. The expectation value of x is clearly zero by symmetry:

$$\langle x\rangle = \int_{-\infty}^{\infty} dx' \langle \alpha|x'\rangle x' \langle x'|\alpha\rangle = \int_{-\infty}^{\infty} dx' |\langle x'|\alpha\rangle|^2 x' = 0. \tag{1.7.36}$$

For x^2 we obtain

$$\begin{aligned}\langle x^2\rangle &= \int_{-\infty}^{\infty} dx' x'^2 |\langle x'|\alpha\rangle|^2 \\ &= \left(\frac{1}{\sqrt{\pi}\,\mathrm{d}}\right)\int_{-\infty}^{\infty} dx' x'^2 \exp\left[\frac{-x'^2}{d^2}\right] \\ &= \frac{d^2}{2},\end{aligned} \tag{1.7.37}$$

which leads to

$$\left\langle (\Delta x)^2\right\rangle = \langle x^2\rangle - \langle x\rangle^2 = \frac{d^2}{2} \tag{1.7.38}$$

for the dispersion of the position operator. The expectation values of p and p^2 can also be computed as follows:

$$\langle p\rangle = \hbar k \tag{1.7.39a}$$

$$\langle p^2\rangle = \frac{\hbar^2}{2d^2} + \hbar^2 k^2, \tag{1.7.39b}$$

which is left as an exercise. The momentum dispersion is therefore given by

$$\left\langle (\Delta p)^2\right\rangle = \langle p^2\rangle - \langle p\rangle^2 = \frac{\hbar^2}{2d^2}. \tag{1.7.40}$$

Armed with (1.7.38) and (1.7.40), we can check the Heisenberg uncertainty relation (1.6.34); in this case the uncertainty product is given by

$$\langle (\Delta x)^2 \rangle \langle (\Delta p)^2 \rangle = \frac{\hbar^2}{4}, \tag{1.7.41}$$

independent of d, so for a Gaussian wave packet we actually have an *equality* relation rather than the more general inequality relation (1.6.34). For this reason a Gaussian wave packet is often called a *minimum uncertainty wave packet.*

We now go to momentum space. By a straightforward integration—just completing the square in the exponent—we obtain

$$\begin{aligned} \langle p'|\alpha \rangle &= \left(\frac{1}{\sqrt{2\pi\hbar}} \right) \left(\frac{1}{\pi^{1/4}\sqrt{d}} \right) \int_{-\infty}^{\infty} dx' \exp\left(\frac{-ip'x'}{\hbar} + ikx' - \frac{x'^2}{2d^2} \right) \\ &= \sqrt{\frac{d}{\hbar\sqrt{\pi}}} \exp\left[\frac{-(p' - \hbar k)^2 d^2}{2\hbar^2} \right]. \end{aligned} \tag{1.7.42}$$

This momentum-space wave function provides an alternative method for obtaining $\langle p \rangle$ and $\langle p^2 \rangle$, which is also left as an exercise.

The probability of finding the particle with momentum p' is Gaussian (in momentum space) centered on $\hbar k$, just as the probability of finding the particle at x' is Gaussian (in position space) centered on zero. Furthermore, the widths of the two Gaussians are inversely proportional to each other, which is just another way of expressing the constancy of the uncertainty product $\langle (\Delta x)^2 \rangle \langle \Delta p)^2 \rangle$ explicitly computed in (1.7.41). The wider the spread in the p-space, the narrower the spread in the x-space, and vice versa.

As an extreme example, suppose we let $d \to \infty$. The position-space wave function (1.7.35) then becomes a plane wave extending over all space; the probability of finding the particle is just constant, independent of x'. In contrast, the momentum-space wave function is δ-function-like and is sharply peaked at $\hbar k$. In the opposite extreme, by letting $d \to 0$, we obtain a position-space wave function localized like the δ-function, but the momentum-space wave function (1.7.42) is just constant, independent of p'.

We have seen that an extremely well localized (in the x-space) state is to be regarded as a superposition of momentum eigenstates with all possible values of momenta. Even those momentum eigenstates whose momenta are comparable to or exceed mc must be included in the superposition. However, at such high values of momentum, a description based on nonrelativistic quantum mechanics is bound to break down.* Despite this limitation

*It turns out that the concept of a localized state in relativistic quantum mechanics is far more intricate because of the possibility of "negative energy states," or pair creation (Sakurai 1967, 118–19).

our formalism, based on the existence of the position eigenket $|x'\rangle$, has a wide domain of applicability.

Generalization to Three Dimensions

So far in this section we have worked exclusively in one-space for simplicity, but everything we have done can be generalized to three-space, if the necessary changes are made. The base kets to be used can be taken as either the position eigenkets satisfying

$$\mathbf{x}|\mathbf{x}'\rangle = \mathbf{x}'|\mathbf{x}'\rangle \tag{1.7.43}$$

or the momentum eigenkets satisfying

$$\mathbf{p}|\mathbf{p}'\rangle = \mathbf{p}'|\mathbf{p}'\rangle. \tag{1.7.44}$$

They obey the normalization conditions

$$\langle \mathbf{x}'|\mathbf{x}''\rangle = \delta^3(\mathbf{x}'-\mathbf{x}'') \tag{1.7.45a}$$

and

$$\langle \mathbf{p}'|\mathbf{p}''\rangle = \delta^3(\mathbf{p}'-\mathbf{p}''), \tag{1.7.45b}$$

where δ^3 stands for the three-dimensional δ-function

$$\delta^3(\mathbf{x}'-\mathbf{x}'') = \delta(x'-x'')\delta(y'-y'')\delta(z'-z''). \tag{1.7.46}$$

The completeness relations read

$$\int d^3x'|\mathbf{x}'\rangle\langle\mathbf{x}'| = 1 \tag{1.7.47a}$$

and

$$\int d^3p'|\mathbf{p}'\rangle\langle\mathbf{p}'| = 1, \tag{1.7.47b}$$

which can be used to expand an arbitrary state ket:

$$|\alpha\rangle = \int d^3x'|\mathbf{x}'\rangle\langle\mathbf{x}'|\alpha\rangle, \tag{1.7.48a}$$

$$|\alpha\rangle = \int d^3p'|\mathbf{p}'\rangle\langle\mathbf{p}'|\alpha\rangle. \tag{1.7.48b}$$

The expansion coefficients $\langle\mathbf{x}'|\alpha\rangle$ and $\langle\mathbf{p}'|\alpha\rangle$ are identified with the wave functions $\psi_\alpha(\mathbf{x}')$ and $\phi_\alpha(\mathbf{p}')$ in position and momentum space, respectively.

The momentum operator, when taken between $|\beta\rangle$ and $|\alpha\rangle$, becomes

$$\langle\beta|\mathbf{p}|\alpha\rangle = \int d^3x'\psi_\beta^*(\mathbf{x}')(-i\hbar\nabla)\psi_\alpha(\mathbf{x}'). \tag{1.7.49}$$

The transformation function analogous to (1.7.32) is

$$\langle \mathbf{x}'|\mathbf{p}'\rangle = \left[\frac{1}{(2\pi\hbar)^{3/2}}\right]\exp\left(\frac{i\mathbf{p}'\cdot\mathbf{x}'}{\hbar}\right), \tag{1.7.50}$$

so that

$$\psi_\alpha(\mathbf{x}') = \left[\frac{1}{(2\pi\hbar)^{3/2}}\right]\int d^3p' \exp\left(\frac{i\mathbf{p}'\cdot\mathbf{x}'}{\hbar}\right)\phi_\alpha(\mathbf{p}') \tag{1.7.51a}$$

and

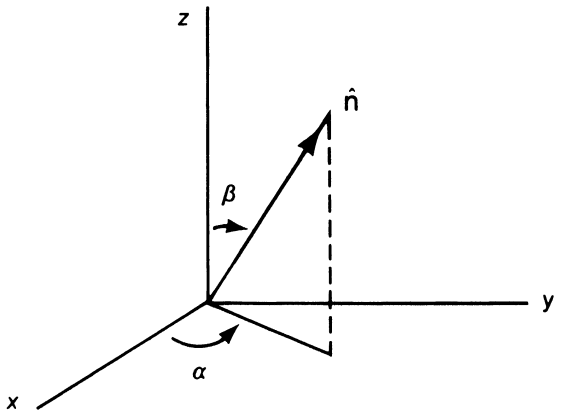
$$\phi_\alpha(\mathbf{p}') = \left[\frac{1}{(2\pi\hbar)^{3/2}}\right]\int d^3x' \exp\left(\frac{-i\mathbf{p}'\cdot\mathbf{x}'}{\hbar}\right)\psi_\alpha(\mathbf{x}'). \tag{1.7.51b}$$

It is interesting to check the dimension of the wave functions. In one-dimensional problems the normalization requirement (1.6.8) implies that $|\langle x'|\alpha\rangle|^2$ has the dimension of inverse length, so the wave function itself must have the dimension of $(\text{length})^{-1/2}$. In contrast, the wave function in three-dimensional problems must have the dimension of $(\text{length})^{-3/2}$ because $|\langle \mathbf{x}'|\alpha\rangle|^2$ integrated over all spatial volume must be unity (dimensionless).

Problems

1. Prove
$$[AB, CD] = -AC\{D, B\} + A\{C, B\}D - C\{D, A\}B + \{C, A\}DB.$$
2. Suppose a 2×2 matrix X (not necessarily Hermitian, nor unitary) is written as
$$X = a_0 + \boldsymbol{\sigma}\cdot\mathbf{a},$$
where a_0 and $a_{1,2,3}$ are numbers.
   a. How are a_0 and a_k $(k=1,2,3)$ related to $\text{tr}(X)$ and $\text{tr}(\sigma_k X)$?
   b. Obtain a_0 and a_k in terms of the matrix elements X_{ij}.
3. Show that the determinant of a 2×2 matrix $\boldsymbol{\sigma}\cdot\mathbf{a}$ is invariant under
$$\boldsymbol{\sigma}\cdot\mathbf{a} \to \boldsymbol{\sigma}\cdot\mathbf{a}' \equiv \exp\left(\frac{i\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\phi}{2}\right)\boldsymbol{\sigma}\cdot\mathbf{a}\exp\left(\frac{-i\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\phi}{2}\right).$$
Find a_k' in terms of a_k when $\hat{\mathbf{n}}$ is in the positive z-direction and interpret your result.
4. Using the rules of bra-ket algebra, prove or evaluate the following:
   a. $\text{tr}(XY) = \text{tr}(YX)$, where X and Y are operators;
   b. $(XY)^\dagger = Y^\dagger X^\dagger$, where X and Y are operators;
   c. $\exp[if(A)] = ?$ in ket-bra form, where A is a Hermitian operator whose eigenvalues are known;
   d. $\Sigma_{a'}\psi^*_{a'}(\mathbf{x}')\psi_{a'}(\mathbf{x}'')$, where $\psi_{a'}(\mathbf{x}') = \langle \mathbf{x}'|a'\rangle$.

5. a. Consider two kets $|\alpha\rangle$ and $|\beta\rangle$. Suppose $\langle a'|\alpha\rangle$, $\langle a''|\alpha\rangle$,... and $\langle a'|\beta\rangle$, $\langle a''|\beta\rangle$,... are all known, where $|a'\rangle$, $|a''\rangle$,... form a complete set of base kets. Find the matrix representation of the operator $|\alpha\rangle\langle\beta|$ in that basis.
   b. We now consider a spin $\frac{1}{2}$ system and let $|\alpha\rangle$ and $|\beta\rangle$ be $|s_z = \hbar/2\rangle$ and $|s_x = \hbar/2\rangle$, respectively. Write down explicitly the square matrix that corresponds to $|\alpha\rangle\langle\beta|$ in the usual (s_z diagonal) basis.
6. Suppose $|i\rangle$ and $|j\rangle$ are eigenkets of some Hermitian operator A. Under what condition can we conclude that $|i\rangle + |j\rangle$ is also an eigenket of A? Justify your answer.
7. Consider a ket space spanned by the eigenkets $\{|a'\rangle\}$ of a Hermitian operator A. There is no degeneracy.
   a. Prove that
$$\prod_{a'}(A - a')$$
   is the null operator.
   b. What is the significance of
$$\prod_{a'' \neq a'} \frac{(A - a'')}{(a' - a'')}?$$
   c. Illustrate (a) and (b) using A set equal to S_z of a spin $\frac{1}{2}$ system.
8. Using the orthonormality of $|+\rangle$ and $|-\rangle$, prove
$$[S_i, S_j] = i\varepsilon_{ijk}\hbar S_k, \qquad \{S_i, S_j\} = \left(\frac{\hbar^2}{2}\right)\delta_{ij},$$
   where
$$S_x = \frac{\hbar}{2}(|+\rangle\langle -| + |-\rangle\langle +|), \qquad S_y = \frac{i\hbar}{2}(-|+\rangle\langle -| + |-\rangle\langle +|),$$
$$S_z = \frac{\hbar}{2}(|+\rangle\langle +| - |-\rangle\langle -|).$$
9. Construct $|\mathbf{S}\cdot\hat{\mathbf{n}}; +\rangle$ such that
$$\mathbf{S}\cdot\hat{\mathbf{n}}|\mathbf{S}\cdot\hat{\mathbf{n}}; +\rangle = \left(\frac{\hbar}{2}\right)|\mathbf{S}\cdot\hat{\mathbf{n}}; +\rangle$$
   where $\hat{\mathbf{n}}$ is characterized by the angles shown in the figure. Express your answer as a linear combination of $|+\rangle$ and $|-\rangle$. [*Note*: The answer is
$$\cos\left(\frac{\beta}{2}\right)|+\rangle + \sin\left(\frac{\beta}{2}\right)e^{i\alpha}|-\rangle.$$
   But do not just verify that this answer satisfies the above eigenvalue equation. Rather, treat the problem as a straightforward eigenvalue



problem. Also do not use rotation operators, which we will introduce later in this book.]

10. The Hamiltonian operator for a two-state system is given by

$$H = a(|1\rangle\langle 1| - |2\rangle\langle 2| + |1\rangle\langle 2| + |2\rangle\langle 1|),$$

where a is a number with the dimension of energy. Find the energy eigenvalues and the corresponding energy eigenkets (as linear combinations of $|1\rangle$ and $|2\rangle$).

11. A two-state system is characterized by the Hamiltonian

$$H = H_{11}|1\rangle\langle 1| + H_{22}|2\rangle\langle 2| + H_{12}[|1\rangle\langle 2| + |2\rangle\langle 1|]$$

where H_{11}, H_{22}, and H_{12} are real numbers with the dimension of energy, and $|1\rangle$ and $|2\rangle$ are eigenkets of some observable ($\neq H$). Find the energy eigenkets and corresponding energy eigenvalues. Make sure that your answer makes good sense for $H_{12} = 0$. (You need not solve this problem from scratch. The following fact may be used without proof:

$$(\mathbf{S}\cdot\hat{\mathbf{n}})|\hat{\mathbf{n}};+\rangle = \frac{\hbar}{2}|\hat{\mathbf{n}};+\rangle,$$

with $|\hat{\mathbf{n}};+\rangle$ given by

$$|\hat{\mathbf{n}};+\rangle = \cos\frac{\beta}{2}|+\rangle + e^{i\alpha}\sin\frac{\beta}{2}|-\rangle,$$

where β and α are the polar and azimuthal angles, respectively, that characterize $\hat{\mathbf{n}}$.)

12. A spin $\frac{1}{2}$ system is known to be in an eigenstate of $\mathbf{S}\cdot\hat{\mathbf{n}}$ with eigenvalue $\hbar/2$, where $\hat{\mathbf{n}}$ is a unit vector lying in the xz-plane that makes an angle γ with the positive z-axis.

a. Suppose S_x is measured. What is the probability of getting $+\hbar/2$?
b. Evaluate the dispersion in S_x, that is,

$$\langle (S_x - \langle S_x \rangle)^2 \rangle.$$

(For your own peace of mind check your answers for the special cases $\gamma = 0$, $\pi/2$, and π.)

13. A beam of spin $\frac{1}{2}$ atoms goes through a series of Stern-Gerlach–type measurements as follows:
a. The first measurement accepts $s_z = \hbar/2$ atoms and rejects $s_z = -\hbar/2$ atoms.
b. The second measurement accepts $s_n = \hbar/2$ atoms and rejects $s_n = -\hbar/2$ atoms, where s_n is the eigenvalue of the operator $\mathbf{S}\cdot\hat{\mathbf{n}}$, with $\hat{\mathbf{n}}$ making an angle β in the xz-plane with respect to the z-axis.
c. The third measurement accepts $s_z = -\hbar/2$ atoms and rejects $s_z = \hbar/2$ atoms.

What is the intensity of the final $s_z = -\hbar/2$ beam when the $s_z = \hbar/2$ beam surviving the first measurement is normalized to unity? How must we orient the second measuring apparatus if we are to maximize the intensity of the final $s_z = -\hbar/2$ beam?

14. A certain observable in quantum mechanics has a 3×3 matrix representation as follows:

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

a. Find the normalized eigenvectors of this observable and the corresponding eigenvalues. Is there any degeneracy?
b. Give a physical example where all this is relevant.

15. Let A and B be observables. Suppose the simultaneous eigenkets of A and B $\{|a', b'\rangle\}$ form a *complete* orthonormal set of base kets. Can we always conclude that

$$[A, B] = 0?$$

If your answer is yes, prove the assertion. If your answer is no, give a counterexample.

16. Two Hermitian operators anticommute:

$$\{A, B\} = AB + BA = 0.$$

Is it possible to have a simultaneous (that is, common) eigenket of A and B? Prove or illustrate your assertion.

17. Two observables A_1 and A_2, which do not involve time explicitly, are known not to commute,

$$[A_1, A_2] \neq 0,$$

yet we also know that A_1 and A_2 both commute with the Hamiltonian:

$$[A_1, H] = 0, \qquad [A_2, H] = 0.$$

Prove that the energy eigenstates are, in general, degenerate. Are there exceptions? As an example, you may think of the central-force problem $H = \mathbf{p}^2/2m + V(r)$, with $A_1 \to L_z$, $A_2 \to L_x$.

18. a. The simplest way to derive the Schwarz inequality goes as follows. First, observe

$$\left(\langle\alpha| + \lambda^*\langle\beta|\right)\cdot\left(|\alpha\rangle + \lambda|\beta\rangle\right) \geq 0$$

for any complex number λ; then choose λ in such a way that the preceding inequality reduces to the Schwarz inequality.

b. Show that the equality sign in the generalized uncertainty relation holds if the state in question satisfies

$$\Delta A|\alpha\rangle = \lambda\,\Delta B|\alpha\rangle$$

with λ purely *imaginary*.

c. Explicit calculations using the usual rules of wave mechanics show that the wave function for a Gaussian wave packet given by

$$\langle x'|\alpha\rangle = (2\pi d^2)^{-1/4}\exp\left[\frac{i\langle p\rangle x'}{\hbar} - \frac{(x'-\langle x\rangle)^2}{4d^2}\right]$$

satisfies the minimum uncertainty relation

$$\sqrt{\langle(\Delta x)^2\rangle}\,\sqrt{\langle(\Delta p)^2\rangle} = \frac{\hbar}{2}.$$

Prove that the requirement

$$\langle x'|\Delta x|\alpha\rangle = (\text{imaginary number})\langle x'|\Delta p|\alpha\rangle$$

is indeed satisfied for such a Gaussian wave packet, in agreement with (b).

19. a. Compute

$$\langle(\Delta S_x)^2\rangle \equiv \langle S_x^2\rangle - \langle S_x\rangle^2,$$

where the expectation value is taken for the S_z+ state. Using your result, check the generalized uncertainty relation

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geq \frac{1}{4}|\langle[A, B]\rangle|^2,$$

with $A \to S_x$, $B \to S_y$.

b. Check the uncertainty relation with $A \to S_x$, $B \to S_y$ for the S_x+ state.

20. Find the linear combination of $|+\rangle$ and $|-\rangle$ kets that maximizes the

uncertainty product

$$\left\langle (\Delta S_x)^2 \right\rangle \left\langle (\Delta S_y)^2 \right\rangle .$$

Verify explicitly that for the linear combination you found, the uncertainty relation for S_x and S_y is not violated.

21. Evaluate the x-p uncertainty product $\langle (\Delta x)^2 \rangle \langle (\Delta p)^2 \rangle$ for a one-dimensional particle confined between two rigid walls

$$V = \begin{cases} 0 & \text{for } 0 < x < a, \\ \infty & \text{otherwise.} \end{cases}$$

Do this for both the ground and excited states.

22. Estimate the rough order of magnitude of the length of time that an ice pick can be balanced on its point if the only limitation is that set by the Heisenberg uncertainty principle. Assume that the point is sharp and that the point and the surface on which it rests are hard. You may make approximations which do not alter the general order of magnitude of the result. Assume reasonable values for the dimensions and weight of the ice pick. Obtain an approximate numerical result and express it *in seconds*.

23. Consider a three-dimensional ket space. If a certain set of orthonormal kets—say, $|1\rangle$, $|2\rangle$, and $|3\rangle$—are used as the base kets, the operators A and B are represented by

$$A \doteq \begin{pmatrix} a & 0 & 0 \\ 0 & -a & 0 \\ 0 & 0 & -a \end{pmatrix}, \qquad B \doteq \begin{pmatrix} b & 0 & 0 \\ 0 & 0 & -ib \\ 0 & ib & 0 \end{pmatrix}$$

with a and b both real.

   a. Obviously A exhibits a degenerate spectrum. Does B also exhibit a degenerate spectrum?
   b. Show that A and B commute.
   c. Find a new set of orthonormal kets which are simultaneous eigenkets of both A and B. Specify the eigenvalues of A and B for each of the three eigenkets. Does your specification of eigenvalues completely characterize each eigenket?

24. a. Prove that $(1/\sqrt{2})(1 + i\sigma_x)$ acting on a two-component spinor can be regarded as the matrix representation of the rotation operator about the x-axis by angle $-\pi/2$. (The minus sign signifies that the rotation is clockwise.)
   b. Construct the matrix representation of S_z when the eigenkets of S_y are used as base vectors.

25. Some authors define an *operator* to be real when every member of its matrix elements $\langle b'|A|b''\rangle$ is real in some representation ($\{|b'\rangle\}$ basis in this case). Is this concept representation independent, that is, do the

matrix elements remain real even if some basis other than $\{|b'\rangle\}$ is used? Check your assertion using familiar operators such as S_y and S_z (see Problem 24) or x and p_x.

26. Construct the transformation matrix that connects the S_z diagonal basis to the S_x diagonal basis. Show that your result is consistent with the general relation

$$U = \sum_r |b^{(r)}\rangle\langle a^{(r)}|.$$

27. a. Suppose that $f(A)$ is a function of a Hermitian operator A with the property $A|a'\rangle = a'|a'\rangle$. Evaluate $\langle b''|f(A)|b'\rangle$ when the transformation matrix from the a' basis to the b' basis is known.

    b. Using the continuum analogue of the result obtained in (a), evaluate

$$\langle \mathbf{p}''|F(r)|\mathbf{p}'\rangle.$$

    Simplify your expression as far as you can. Note that r is $\sqrt{x^2 + y^2 + z^2}$, where x, y, and z are *operators*.

28. a. Let x and p_x be the coordinate and linear momentum in one dimension. Evaluate the classical Poisson bracket

$$[x, F(p_x)]_{\text{classical}}.$$

    b. Let x and p_x be the corresponding quantum-mechanical operators this time. Evaluate the commutator

$$\left[x, \exp\left(\frac{ip_x a}{\hbar}\right)\right].$$

    c. Using the result obtained in (b), prove that

$$\exp\left(\frac{ip_x a}{\hbar}\right)|x'\rangle, \qquad (x|x'\rangle = x'|x'\rangle)$$

    is an eigenstate of the coordinate operator x. What is the corresponding eigenvalue?

29. a. On page 247, Gottfried (1966) states that

$$[x_i, G(\mathbf{p})] = i\hbar\frac{\partial G}{\partial p_i}, \qquad [p_i, F(\mathbf{x})] = -i\hbar\frac{\partial F}{\partial x_i}$$

    can be "easily derived" from the fundamental commutation relations for all functions of F and G that can be expressed as power series in their arguments. Verify this statement.

    b. Evaluate $[x^2, p^2]$. Compare your result with the classical Poisson bracket $[x^2, p^2]_{\text{classical}}$.

30. The translation operator for a finite (spatial) displacement is given by

$$\mathscr{T}(\mathbf{l}) = \exp\left(\frac{-i\mathbf{p}\cdot\mathbf{l}}{\hbar}\right),$$

where $\mathbf{p}$ is the momentum *operator*.

a. Evaluate

$$[x_i, \mathscr{T}(\mathbf{l})].$$

b. Using (a) (or otherwise), demonstrate how the expectation value $\langle \mathbf{x} \rangle$ changes under translation.

31. In the main text we discussed the effect of $\mathscr{T}(d\mathbf{x}')$ on the position and momentum eigenkets and on a more general state ket $|\alpha\rangle$. We can also study the behavior of expectation values $\langle \mathbf{x} \rangle$ and $\langle \mathbf{p} \rangle$ under infinitesimal translation. Using (1.6.25), (1.6.45), and $|\alpha\rangle \rightarrow \mathscr{T}(d\mathbf{x}')|\alpha\rangle$ only, prove $\langle \mathbf{x} \rangle \rightarrow \langle \mathbf{x} \rangle + d\mathbf{x}', \langle \mathbf{p} \rangle \rightarrow \langle \mathbf{p} \rangle$ under infinitesimal translation.

32. a. Verify (1.7.39a) and (1.7.39b) for the expectation value of p and p^2 from the Gaussian wave packet (1.7.35).

b. Evaluate the expectation value of p and p^2 using the momentum-space wave function (1.7.42).

33. a. Prove the following:

$$\text{(i)} \qquad \langle p'|x|\alpha\rangle = i\hbar \frac{\partial}{\partial p'} \langle p'|\alpha\rangle,$$

$$\text{(ii)} \qquad \langle \beta|x|\alpha\rangle = \int dp' \, \phi_\beta^*(p') i\hbar \frac{\partial}{\partial p'} \phi_\alpha(p'),$$

where $\phi_\alpha(p') = \langle p'|\alpha\rangle$ and $\phi_\beta(p') = \langle p'|\beta\rangle$ are momentum-space wave functions.

b. What is the physical significance of

$$\exp\left(\frac{ix\Xi}{\hbar}\right),$$

where x is the position operator and Ξ is some number with the dimension of momentum? Justify your answer.