THE THEORY OF QUANTUM MECHANICS:
SEVEN LECTURES FOR STUDENTS OF INTRODUCTORY UNIVERSITY PHYSICS

by

Daniel T. Gillespie

Research Department, Naval Weapons Center
China Lake, California 93555

Prepared for the Final Report of the

Carleton College Conference (August 1988)

of the

Introductory University Physics Project

PREFACE

I came to write these lectures in the course of participating in the Introductory University Physics Project (IUPP), which was organized in 1987 by the American Physical Society and the American Association of Physics Teachers under the sponsorship of the National Science Foundation. The ultimate goal of IUPP is to develop some new models for the introductory, calculus-based, university physics course. As envisioned by IUPP chairman John Rigden, the new course models should have a "leaner story line," and should if at all possible contain something meaningful about quantum mechanics. My own involvement in IUPP is as a member of the working group on quantum mechanics, which is chaired by Eugen Merzbacher and co-chaired by Thomas Moore.

I must emphasize that the set of lectures presented here is not patterned after the syllabus recommended by the IUPP quantum mechanics working group. The approach taken in these lectures is frankly too abstract and too formal, too long on theory and too short on applications, for *most* of the students who take the introductory calculus-based physics course. But *some* of those students, those with a more mathematical and philosophical bent, might possibly profit from these lectures. And perhaps also some physics teachers might find one or more features of these lectures worth incorporating into their own lectures and writings.

Why should a university student bother to learn quantum mechanics? For students intent on becoming professional physicists, the motive has always been clear and compelling: If you don't learn quantum mechanics, then you can't join the club. Obviously, a different motive must impel students on a broader spectrum, and perhaps should impel physics majors as well. The motive that I have tried to implicitly establish in these lectures is roughly as follows. Nature has been found to behave, on the microscale, in ways that seem to utterly defy common sense. Mankind has thus been presented with an immense challenge: Make sense of Nature on the microscale! Quantum mechanics, within carefully prescribed limits, has successfully met that challenge. So the theory of quantum mechanics represents an intellectual achievement of truly heroic proportions, and may rightly be regarded as "a principal jewel in the cultural crown of civilization." To appreciate the beauty and mystery of that cultural jewel is alone sufficient reason for any student of any university to undertake a serious study of quantum mechanics.

It seems that there are two approaches to teaching physics in general and quantum mechanics in particular: In the *axiomatic* approach, one first sets forth a minimal number of assumptions or axioms, and one then rigorously deduces their various consequences. By contrast, in the *organic* approach one pays little attention to logical structure, and simply allows the various facts of the theory to arrange themselves in whatever way seems convenient. While I have respect for the organic approach, and in fact suspect that physicists of that camp are more likely than their opposites to make significant advances in physics, I have deliberately written these lectures in the axiomatic vein. I believe that students with an aptitude for mathematics and logical reasoning strongly prefer to *learn* by the axiomatic approach. The "axioms" of quantum mechanics are presented here in the form of *five rules*; the first four rules are stated in Lecture 3, and the fifth rule is stated in Lecture 6. The *language* of those rules, which provides the allowed mechanisms of inference, is the mathematical language of linear algebra, which I have preferred here to call "generalized vector theory." That language is developed, in I think a rather novel way, in Lecture 2. Lecture 2 is therefore a make-or-break lecture for the whole series. Although each of the seven lectures can be *read* in one hour, I think most will take *two* 50-minute lecture sessions to successfully *deliver*. Lectures 5 and 7 are the most challenging of the series, and could be omitted for a less demanding five-lecture sequence; however, a glance at the table of contents will show that omission of Lectures 5 and 7 will cut out several major topics.

I feel obliged to forewarn teachers who are contemplating using these lectures that I have taken what some might consider to be an unnecessarily doctrinaire position, namely, that *observables for microscale systems do not always have values*. In fact, I make that premise to be a principal motive for devising quantum mechanics. A teacher who feels that this premise is unwarranted, or even false, will not like these lectures. But before such a teacher lays these pages aside for that reason, I would respectfully entreat a reading of the entertaining article by N. David Mermin in the April 1985 issue of *Physics Today*. In that article, Mermin discusses the implications of Bell's theorem in terms of an idealized experiment that is equivalent to an electron-pair spin-correlation experiment, but which is happily devoid of the esoteric jargon of quantum physics. Mermin shows with remarkable clarity that there is simply no way of explaining the results of his experiment without doing great violence to one's sense of "reasonableness." I think that a defensible interpretation of the specific conclusions of the Mermin experiment is this: There is no way to assign fixed values (red or green) to each observable (1, 2 and 3) of both particles in such a way that the experimental data can be quantitatively acounted for; hence, we cannot generally ascribe simultaneous values, *even unknown values*, to non-commuting observables. This is one of several non-common-sensical features of quantum mechanics that have been exposed by John Bell's work in 1964, and more recently confirmed experimentally by Alain Aspect and coworkers. I believe it is wrong not to be up-front with students about such issues. I believe it is wrong to pretend to students, on the pretext of "doing physics rather than philosophy," that although quantum mechanics may seem strange and unusual there is nothing about it that should disturb our view of Reality. Indeed, I feel that the strong claim which quantum mechanics has to profound cultural significance stems largely from the fact that it *has* raised serious and as yet unanswered questions about Reality. We need not necessarily make a philosophical study of such issues in a physics course, and we do *not* do that in these lectures, but we should not deny the existence of these issues. As evidence in these lectures for the "fact" that observables for microscale systems do not always have values, I have invoked the double-slit experiment. The evidence provided by that experiment is strong, but not unassailable. Evidence of a more compelling nature would have been provided by an electron-pair spin-correlation experiment, but unfortunately a satisfactory quantum analysis of such an experiment lies beyond the reach of these lectures. The double-slit experiment on the other hand *can* be analyzed within the framework of these lectures, and in fact it provides us with a thematic bridge from Lecture 1 to Lecture 7.

As to the treatment given in these lectures of the quantum theory itself, I have indeed organized and presented the ingredients of the theory in a very different way than is usually done. But the basic view of quantum mechanics taken in these lectures is quite orthodox, and is fully in line with what one will find in such standard textbooks as Dirac, Messiah and Merzbacher. I have tried to present in these lectures a highly simplified and hence unconventional rendering of conventional quantum theory.

I would like to thank John Rigden and Eugen Merzbacher for their encouragement in preparing these lectures. I am happy to acknowledge the benefits of conversations with other physicists who participated in the 1988 IUPP conferences at Harvey Mudd College and Carleton College, especially those in the quantum mechanics working group. And I want to thank Ron Derr, Head of the Research Department of the Naval Weapons Center, for allowing me to participate in IUPP and develop these lectures within the framework of the Center's Independent Research Program.

China Lake, California
May, 1989

D. T. Gillespie

CONTENTS

† These two lectures may be omitted for a less demanding five-lecture series.

LECTURE 1. INTRODUCTION AND MOTIVATION

▸1.1 The Fundamental Problem of Mechanics

The subject of "mechanics" deals with a *system* and its observable properties or *observables*. If A is an observable of a given system, then we can at any time "make a measurement of A on the system," and thereby obtain a *result* A, an ordinary number that we call the *value* of A. For example, the *position* (an observable A) of a *ball* (the system being considered) is measured, with the numerical *result* 4.7 ($=A$, the measured *value* of the observable A).

In the "classical" mechanics of Isaac Newton, there is no need to distinguish between an observable A and its value A: The values of the observables of a system are represented by ordinary mathematical variables, and the central program of mechanics is to find (e.g., by using $F=ma$) the time dependence of those "observable variables." This program is very successful on the *macroscale*, the scale of systems that we are familiar with from our everyday experience (such as the moon, airplanes, tennis balls and dust particles). But the classical program runs into serious difficulties on the *microscale*, the scale of individual atoms and molecules. As a prelude to describing the microscale difficulties of classical mechanics, and how "quantum" mechanics ultimately avoids those difficulties, let's begin by rephrasing the goal of mechanics in a more cautious "operational" way:

THE FUNDAMENTAL PROBLEM OF MECHANICS (FPOM):

- At time zero we measure some observable A on the system, and obtain the result A.
- Then we let the system evolve on its own until some time t ($t \geq 0$).
- At t, we measure some observable B on the system.
- Knowing A, A, t and B, *what can we predict about the result B of the second measurement?*

▸1.2 The Failure of Classical Mechanics on the Microscale

In the classical approach to the FPOM, we know that in some circumstances we can predict B *exactly*, while in other circumstances we can predict B only *probabilistically*. For example, consider a simple harmonic oscillator of mass m and spring constant k. Classically, it executes sinusoidal motion with frequency $2\pi(k/m)^{1/2}$ and some amplitude a. Suppose the first-measured observable A is "total energy," and the second measured observable B is "position." Since the energy value A will be related to the oscillation amplitude a by $A=ka^2/2$, then after the A-measurement we will know that the amplitude of the oscillator motion is $a=[2A/k]^{1/2}$. But that's *all* we will know. So we can't make a unique prediction for the result B of any subsequent position measurement. We can, however, make some *probabilistic* predictions: We know that B must lie somewhere between $\pm[2A/k]^{1/2}$; moreover, since the particle spends more time near the turning points $\pm[2A/k]^{1/2}$, where it moves slowly, than near 0, where it moves quickly, then we should expect B-values near $\pm[2A/k]^{1/2}$ to be more likely than B-values near 0. In fact, if we reason carefully from the classical equations of motion for a harmonic oscillator (we won't go through the details here), we can derive the "probability density" curve for the B-values shown in Fig. 1-1. The shaded area in that figure is equal to the *probability* that the result B will lie between B_1 and B_2. (The curve is the reciprocal of the oscillator's speed at position B, multiplied by a constant that makes the total area under the curve unity.)

Well, what do we find *experimentally*? We get excellent quantitative agreement with this classical prediction for particles of "tangible" mass. But if we could do the experiment with a particle of *very small* mass, such as an atom, we would see marked deviations from this prediction: Depending upon the energy result A, we would find that it is *impossible* to get *some* position values B between $\pm[2A/k]^{1/2}$, and *possible* to get position values B with $|B|>[2A/k]^{1/2}$. [When m is very small we would also notice that the energy values A obtained in the first measurement are always integer multiples of a number proportional to $(k/m)^{1/2}$, another feature that is not predicted by classical mechanics.]

We conclude from this and similar experiments that, at the very least, the values of the observables of a *micro*scale system are not interrelated by the familiar classical formulas that work so well for *macro*scale systems. But in fact, the failure of classical mechanics on the microscale runs

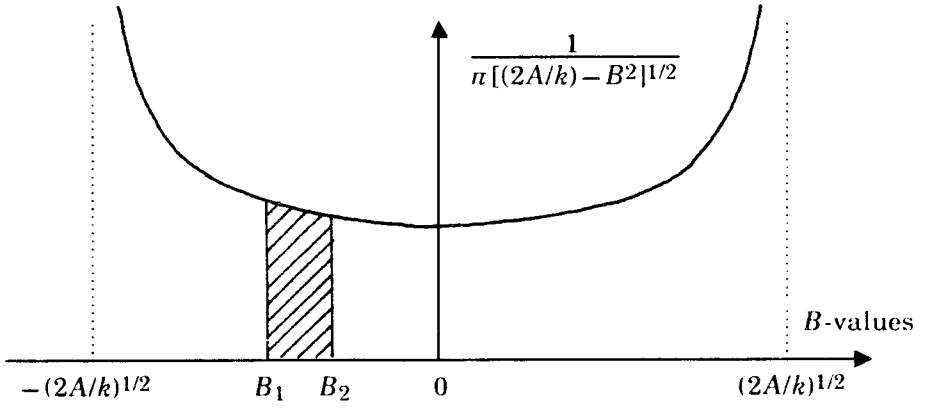


FIG. 1-1. The classical prediction for the result B of a position measurement that is performed on a harmonic oscillator with total energy A and spring constant k. The shaded area is equal to the *probability* that the result of such a measurement will fall between the position values B_1 and B_2.

much deeper than this: Other experiments indicate that it is *pointless*, even *self-contradictory*, to say that microscale system observables *always* "have values." Of those experiments, perhaps the easiest to both describe and analyze is the so-called *double-slit experiment*.

In the double-slit experiment, schematized in Fig. 1-2, a very small particle — let's say an *electron* for definiteness — is fired with horizontal momentum p_0 at a vertical screen S_1 (assume that no gravity forces operate). Screen S_1 is opaque to the electron except for two horizontal slits at $y=y_1$ and $y=y_2$. To actually *see* the effects that we shall describe, these two slits must be *very narrow* and *very close together*, so this is indeed a "microscale" experiment. Beyond S_1 is a second vertical screen S_2, this one coated like a television screen with a substance that emits a spot of light wherever it is struck

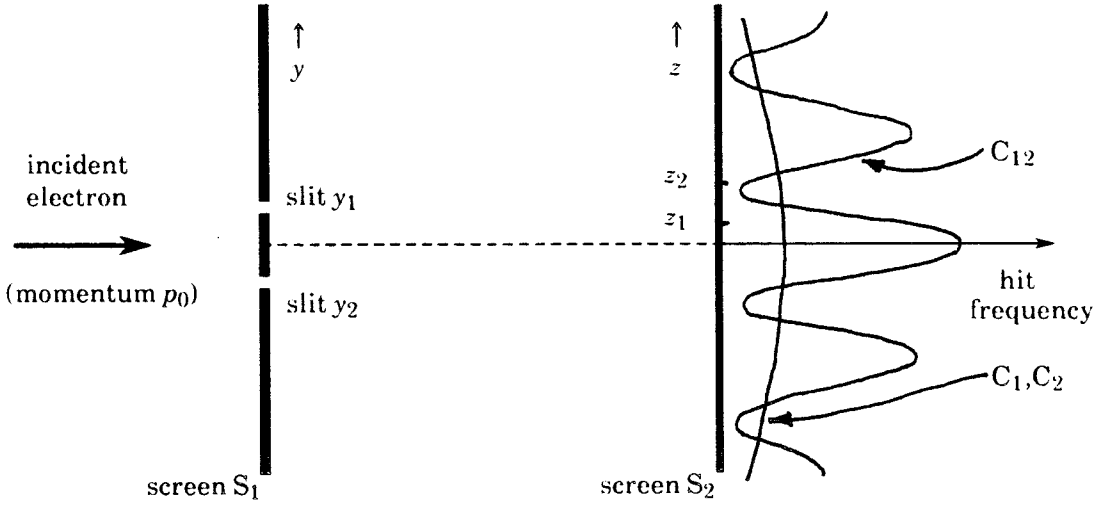


FIG. 1-2. Schematic diagram of the double-slit experiment, showing the hit probability patterns with both slits open (curve C_{12}) and with only one slit open (curves C_1 and C_2). The length scale on screen S_1 has been greatly magnified in this drawing; compared to the "macroscopic" scale of screen S_2, the two slits in screen S_1 are "microscopic" in both their size and their separation.

by an electron. The coating allows us to record the point z where the electron hits S_2. (We measure vertical position along S_1 by y, and vertical position along S_2 by z. Note that in Fig. 1-2 the y-scale is greatly magnified relative to the z-scale.) If we repeat this experiment many times, we can empirically determine the *probability* that the impact point of the electron on screen S_2 will be at any point z. Curve C_{12} shows schematically the hit probability (or hit frequency) pattern that is actually observed when slits y_1 and y_2 are both open. Now it so happens that this curve looks exactly like the interference intensity pattern we would get if we were transmitting through the two slits, instead of particles, *waves* of sound or light with wavelength

$$\lambda = 2\pi\hbar/p_0. \tag{1-1}$$

Here, p_0 is the measured initial momentum of the electrons, and $\hbar$, called "h-bar", is a universal physical constant introduced by Max Planck with the numerical value

$$\hbar = 1.054\ldots\times 10^{-34}\ \text{joule·sec}. \tag{1-2}$$

So the shape of curve C_{12} strongly suggests that an electron is a *wave*; yet the fact that the curve C_{12} is inferred from the statistics of one-at-a-time, point-like scintillations of equal intensity suggests equally strongly that an electron is a *particle*. This curious behavior is not peculiar to electrons alone, but is observed for virtually *all* elementary constituents of matter and energy (such as protons, photons, etc.). Nature seems to exhibit on the microscale a *wave-particle duality* that is simply not understandable in purely classical terms.

Curve C_1 in Fig. 1-2 shows schematically the hit probability pattern that is observed when only slit y_1 is open. The curve C_2 obtained when only slit y_2 is open is virtually indistinguishable from curve C_1, because the slit separation distance on screen S_1 is miniscule on the length scale of screen S_2. Notice that there are some points on screen S_2, such as z_1, that are more likely to be hit when both slits are open than when only one slit is open; that's not surprising. But what *is* surprising is the fact that there are other points on screen S_2, such as z_2, that are *less* likely to be hit when both slits are open than when only one slit is open; i.e., by *closing* one slit we make it *more likely* that point z_2 will be hit on any *one* electron firing! Now, it is obvious that any electron arriving at screen S_2 can only have come by way of the two slits y_1 and y_2 in screen S_1. And it is equally obvious that when *only* slit y_1 is open, any electron arriving at S_2 can only have come through that slit, and hence must have had at screen S_1 the y-position value y_1. When *both* slits are open, common sense would seem to require that any electron arriving at S_2 must have come *either* wholly through slit y_1 *or* wholly through slit y_2. But if that were true, then closing one slit could not possibly *increase* the likelihood that an electron will reach point z_2, as we observe that it does. So we must conclude that "common sense" is *wrong*: any electron that reaches screen S_2 with both slits open did *not* go through *either* slit y_1 *or* slit y_2, but instead somehow made use of *both* slits! It follows that, for such an electron, the observable "y-position at S_1" *cannot* meaningfully be said to have "had a value."

This finding is just one example of the astonishing conclusion that physicists have been forced to by many carefully done experiments: *Observables for microscale systems do not always "have values."* Therefore, we cannot address the Fundamental Problem of Mechanics on the microscale with the usual "common sense" classical approach, because that approach always starts from the assumption that system observables *do* always have values. (For example, we can't analyze the double-slit experiment by trying to figure out how the position value of the electron changes with time, because a "position value of the electron" does not always *exist* during that experiment.) Although classical mechanics works quite well for macroscale systems, it seems that a radically different approach to microscale mechanics must be devised.

▸1.3 The Aim and Plan of These Lectures

The "replacement theory" for classical mechanics on the microscale is called *quantum mechanics*. Quantum mechanics was developed during the first quarter of the Twentieth Century by Werner Heisenberg, Erwin Schrödinger, Niels Bohr, Max Born, Paul Dirac, John von Neumann, and a number of others. We shall not delve into the history of quantum mechanics in these lectures, but we should note that the decision of early Twentieth Century physicists to abandon classical mechanics on

the microscale was at the time both daring and traumatic. The theory that ultimately emerged from the heroic efforts of those physicists, the *standard theory of quantum mechanics* . . .

- takes a radically different approach than classical mechanics;
- is "non-intuitive" from the viewpoint of ordinary experience;
- gives correct results on the microscale;
- reduces to classical mechanics on the macroscale;
- is controversial, even among physicists;
- is one of the most ingenious and significant intellectual achievements in the history of mankind.

There is no way that we can cover quantum mechanics completely in seven lectures. We will have to leave a lot out, and greatly simplify the rest. The limited goal of these lectures will *not* be to show you how to solve lots of quantum mechanics problems, but rather to give you an accurate appreciation of the "essence" of quantum theory. With some effort on your part, you should acquire from these lectures an honest sense of the strange kind of reasoning we have to use in order to successfully describe physical phenomena on the microscale.

We shall first have to learn some new mathematics — calculus alone is not enough for quantum mechanics. We'll get a brief start on that task in just a moment by developing some facts about *complex numbers*. We'll need those facts in our second lecture, where we will develop the mathematical theory of *generalized vectors*. The theory of generalized vectors provides the basic "mathematical language" of quantum mechanics, and we will have to achieve some fluency in that language if we are to gain any real insight into quantum theory. The bad news is that the theory of generalized vectors is very abstract; the good news is that it is in many ways less difficult to learn than calculus, and we won't have to use it here to make any terribly complicated calculations.

In our third lecture we shall use the language of generalized vectors to lay out the first four *rules* (or axioms or laws) of quantum theory. Those four rules will enable us to see how, at least in principle, quantum theory frames its answer to the FPOM for $t \approx 0$ (an A-measurement followed *immediately* by a B-measurement). Those rules thus form the foundation of "quantum statics," where there is no consideration given to the passage of time. Remarkably, *all* the non-intuitive "weirdness" of quantum theory can be exposed in the context of quantum statics. In our fourth and fifth lectures we shall continue our discussion of quantum statics by pursuing some other important results.

In our sixth lecture we shall state the fifth rule of quantum theory, and we shall see how it allows us to finally frame an answer to the FPOM for *any* $t \geq 0$. At that stage, we'll have all the basic ingredients for "quantum dynamics," and an essentially complete quantum theory. Our seventh and final lecture will consider the specific problem of a free particle in one dimension, and will provide us at last with a view of how quantum mechanics accounts for the seemingly unaccountable results of the double-slit experiment.

Note: For an abbreviated five-lecture sequence that is less mathematically demanding of the student, the *fifth* and *seventh* lectures may be omitted.

▸1.4 Complex Numbers

You have probably already encountered complex numbers. Here we are going to review a few basic facts about them that we shall need for our development of quantum mechanics.

A *complex number* is a number of the form

$$c = a + ib, \tag{1-3}$$

where $i \equiv \sqrt{-1}$, and a and b are ordinary real numbers. We call a the "real part" of c, and b the "imaginary part" of c, and we write

$$a \equiv \mathrm{Re}\{c\}, \quad b \equiv \mathrm{Im}\{c\}.$$

If $b = 0$ we say that c is pure real, while if $a = 0$ we say that c is pure imaginary. And if a and b are *both* zero, we write $c = 0$. It is important to understand that the terms "real" and "imaginary" are used

here in a strictly technical sense; you should *not* grant them their usual connotations of "genuine" and "artificial."

The key feature of complex numbers is that they can be added and multiplied by using the ordinary rules of algebra, but *supplemented* by the rule that $i^2 = -1$.

●*Exercise 1-1.* If $c_1 = a_1 + ib_1$ and $c_2 = a_2 + ib_2$, show that

$$c_1 + c_2 = (a_1 + a_2) + i(b_1 + b_2), \tag{1-4a}$$

$$c_1 c_2 = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2). \tag{1-4b}$$

The *complex conjugate* of c in (1-3) is defined to be the complex number

$$c^* = a - ib. \tag{1-5}$$

●*Exercise 1-2.* Show that c is pure real if and only if $c^* = c$. Also prove that

$$(c^*)^* = c, \tag{1-6a}$$

$$c + c^* = 2\mathrm{Re}\{c\}, \tag{1-6b}$$

$$(c_1 + c_2)^* = c_1^* + c_2^*, \tag{1-6c}$$

$$(c_1 c_2)^* = c_1^* c_2^*. \tag{1-6d}$$

The *square* of c, namely

$$c^2 \equiv c\,c = (a^2 - b^2) + i(2ab), \tag{1-7a}$$

is obviously a complex number. However, the *square modulus* of c, which we define to be

$$|c|^2 \equiv c\,c^* = c^* c = a^2 + b^2, \tag{1-7b}$$

is evidently a *non-negative real number*, which vanishes if and only if $c = 0$.

●*Exercise 1-3.* Carry out the algebra leading to equations (1-7). Also prove, using the results of Exercise 1-2, that

$$|c_1 c_2| = |c_1||c_2|, \tag{1-8a}$$

$$|c_1 + c_2|^2 = |c_1|^2 + |c_2|^2 + 2\mathrm{Re}\{c_1 c_2^*\}. \tag{1-8b}$$

Finally, for any real number u we define

$$e^{iu} \equiv \cos u + i \sin u. \tag{1-9}$$

One rationale for denoting the complex number on the right side of (1-9) by the *exponential symbol* on the left is this: If a is any real constant and x any real variable, then

$$(d/dx)e^{iax} = (d/dx)[\cos ax + i \sin ax] = -a \sin ax + ia \cos ax = ia[i \sin ax + \cos ax],$$

$$(d/dx)e^{iax} = ia\, e^{iax}, \tag{1-10}$$

which is precisely what we would get if i were an ordinary real number.

●*Exercise 1-4.* Using the definition (1-9), prove that

$$e^{i0} = 1, \tag{1-11a}$$

$$(e^{iu})^* = \cos u - i \sin u = e^{-iu}, \tag{1-11b}$$

$$|e^{iu}|^2 = 1, \tag{1-11c}$$

$$(e^{iu})(e^{iv}) = e^{i(u+v)}. \tag{1-11d}$$

▸ *Before* the next lecture, you should do Exercises 1-1 through 1-4. You should become fairly familiar with these facts about complex numbers before tackling the mathematics of generalized vectors, which is what we will do in the next lecture.

LECTURE 2. THE MATHEMATICS OF GENERALIZED VECTORS

▸2.1 Vectors, Scalars, Components, Basis Expansions

You have learned that there are some quantities in physics, such as mass, temperature and density, that have only "magnitude," and are called *scalars*; other quantities, such as position, velocity and force, have both "magnitude" and "direction," and are called *vectors*. Let's review some things you presumably already know about vectors.

[See Fig. 2-1] We can picture a *vector* **v** as an *arrow*, whose length is the magnitude of the vector, and whose sense (from its tail to its head) is the direction of the vector. If we translate the arrow, i.e., move it without changing its length or direction, we don't change the vector. We can multiply any vector **v** by any real number r, called a *scalar*, to get a new vector, written $r\mathbf{v}$ or $\mathbf{v}r$, which has a magnitude that is $|r|$ times the magnitude of **v** and a direction that is the same or opposite the direction of **v** accordingly as r is positive or negative. Two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ can be added to form a new vector, $\mathbf{v}_1 + \mathbf{v}_2$, by translating them so that the tail of $\mathbf{v}_2$ lies on the head of $\mathbf{v}_1$ and then drawing the sum vector from the tail of $\mathbf{v}_1$ to the head of $\mathbf{v}_2$. And $\mathbf{v}_2 + \mathbf{v}_1$ is the same vector as $\mathbf{v}_1 + \mathbf{v}_2$.

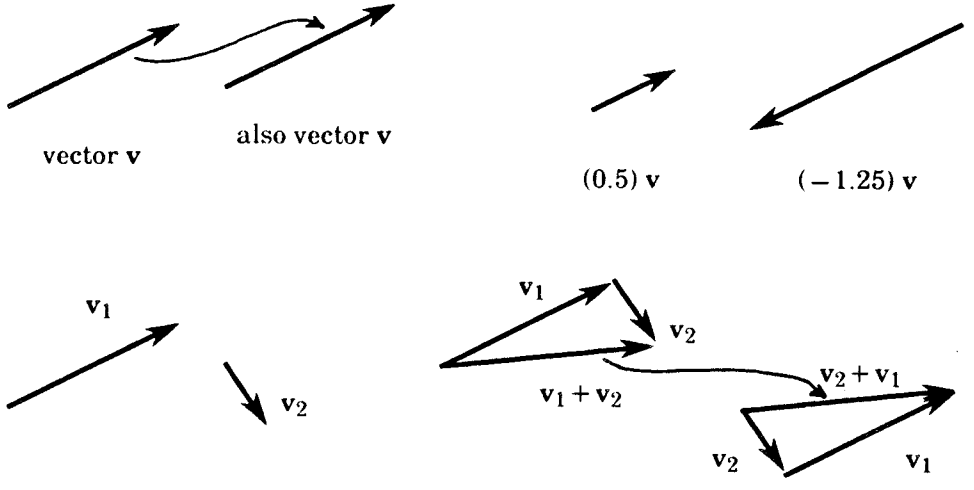


FIG. 2-1. Vectors, scalar multiplication and vector addition.

[See Fig. 2-2] Any vector with length 1 is called a *unit vector*. The *component* of any vector **v** relative to a unit vector **e** is the perpendicularly projected length of **v** onto **e**, a scalar that we denote by **e**·**v**. If the head-to-head angle between **v** and **e** is θ, then **e**·**v** is equal to the magnitude of **v** times $\cos\theta$; thus, **e**·**v** is positive or negative accordingly as θ is less than or greater than $\pi/2$. If $\theta = \pi/2$, then **e**·**v** = 0, and we say that **e** and **v** are *orthogonal*. The **e**-component of $r\mathbf{v}$ is r times the **e**-component of **v**; the **e**-component of $\mathbf{v}_1 + \mathbf{v}_2$ is the sum of the **e**-components of $\mathbf{v}_1$ and $\mathbf{v}_2$.

[See Fig. 2-3] In two dimensions, any two orthogonal unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ form a *basis*, and any vector **v** can be "expanded" in that basis according to the rule

$$\mathbf{v} = \mathbf{e}_1(\mathbf{e}_1\cdot\mathbf{v}) + \mathbf{e}_2(\mathbf{e}_2\cdot\mathbf{v}),$$

wherein the scalar multiplying the unit vector $\mathbf{e}_j$ is just the $\mathbf{e}_j$-component of **v**. Similarly in three dimensions, any three mutually orthogonal unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ constitute a basis, and any vector **v** can be expanded in that basis according to

$$\mathbf{v} = \mathbf{e}_1(\mathbf{e}_1\cdot\mathbf{v}) + \mathbf{e}_2(\mathbf{e}_2\cdot\mathbf{v}) + \mathbf{e}_3(\mathbf{e}_3\cdot\mathbf{v}).$$

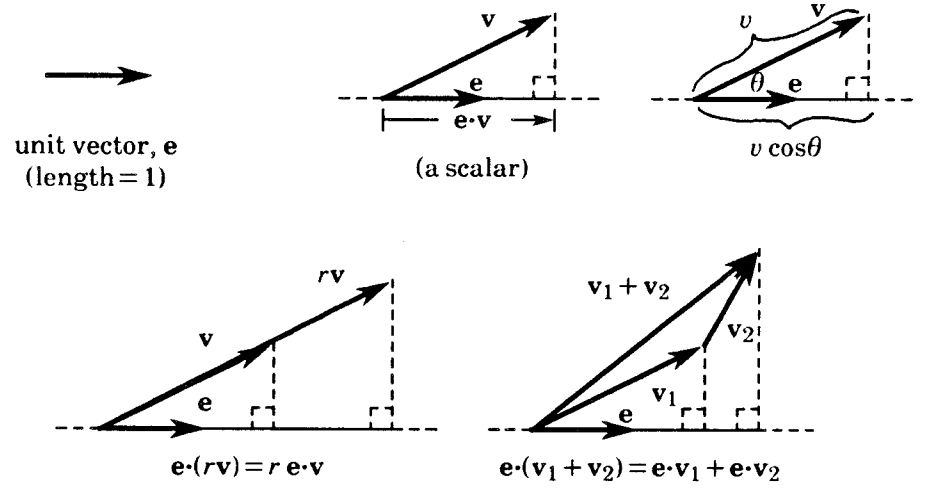


FIG. 2-2. Unit vectors and components relative thereto.

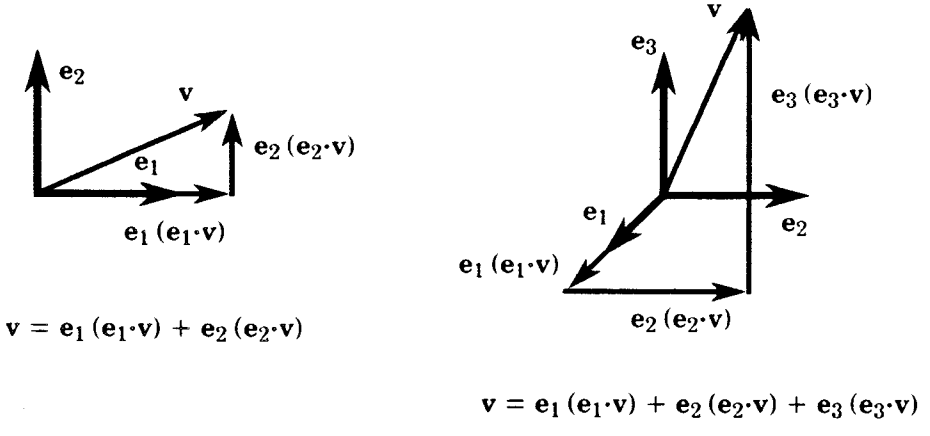


FIG. 2-3. Bases and expansions therein.

The aforementioned facts about vectors should not really be new to you. Now, besides *vectors*, you have no doubt also encountered in your studies *mathematicians*; they too are both amusing and useful. One of the amusing and useful things about mathematicians is their penchant for doing mathematics without relying on pictures, as we just did in our discussion of vectors. Let's see how a mathematician might try to "formalize" the foregoing vector notions, and lay out a mathematical theory of vectors without ever refering to directed line segments. [*Note to students of linear algebra*: The definitions and postulates that follow will be a bit different from what you've learned. We're going to take as direct and elementary an approach as our later needs will allow.]

We define a *vector space* to be a collection of objects **v**, called *vectors*, such that the following things [(a) through (d)] are true:

(a) "Scalar multiplication" is defined: If **v** is any vector in the space and r is any real number, also called a *scalar*, then $r\mathbf{v}$ ($\equiv \mathbf{v}r$), called the product of r and **v**, is also a vector in the space.

(b) "Vector addition" is defined: If $\mathbf{v}_1$ and $\mathbf{v}_2$ are any two vectors in the space, then $\mathbf{v}_1 + \mathbf{v}_2$, called the sum of $\mathbf{v}_1$ and $\mathbf{v}_2$, is also a vector in the space. And $\mathbf{v}_2 + \mathbf{v}_1$ is the same vector as $\mathbf{v}_1 + \mathbf{v}_2$.

Comment: From (a) and (b) we see that, for any two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in the space and any two scalars r_1 and r_2, the *linear combination* $r_1\mathbf{v}_1 + r_2\mathbf{v}_2$ is a well-defined vector in the space.

(c) There exists in the space *unit vectors*, relative to which all vectors in the space have *components*. The component of the vector $\mathbf{v}$ relative to the unit vector $\mathbf{e}$, called the $\mathbf{e}$-*component of* $\mathbf{v}$, is written $\mathbf{e}\cdot\mathbf{v}$, and has the following properties:

(c1) $\mathbf{e}\cdot\mathbf{v}$ is a scalar.

(c2) $\mathbf{e}\cdot(r_1\mathbf{v}_1 + r_2\mathbf{v}_2) = r_1(\mathbf{e}\cdot\mathbf{v}_1) + r_2(\mathbf{e}\cdot\mathbf{v}_2)$.

(c3) $\mathbf{e}\cdot\mathbf{e} = 1$.

(c4) $\mathbf{e}\cdot\mathbf{e}' = \mathbf{e}'\cdot\mathbf{e}$, for any two unit vectors $\mathbf{e}$ and $\mathbf{e}'$.

Comment: (c1) says that the $\mathbf{e}$-component of any vector $\mathbf{v}$ is a real number. (c2) says that the $\mathbf{e}$-component of any linear combination vector is the same linear combination of the $\mathbf{e}$-components. (c3) says that the $\mathbf{e}$-component of itself is unity, and thus serves to *define* a unit vector. And (c4) says that for any two unit vectors $\mathbf{e}$ and $\mathbf{e}'$, the $\mathbf{e}$-component of $\mathbf{e}'$ is equal to the $\mathbf{e}'$-component of $\mathbf{e}$.

(d) If the vector space is N-dimensional, then there exists at least one set of N unit vectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\}$ called a *basis* such that

(d1) $\mathbf{e}_j\cdot\mathbf{e}_k = 0$ whenever $j \neq k$.

(d2) For any vector $\mathbf{v}$ in the space,

$$\mathbf{v} = \mathbf{e}_1(\mathbf{e}_1\cdot\mathbf{v}) + \mathbf{e}_2(\mathbf{e}_2\cdot\mathbf{v}) + \ldots + \mathbf{e}_N(\mathbf{e}_N\cdot\mathbf{v}). \tag{2-1}$$

Comment: (d1) stipulates that the unit vectors composing a basis be *mutually orthogonal*. (d2) says that any vector $\mathbf{v}$ can be written as a linear combination of the basis vectors, and that the scalar coefficient of $\mathbf{e}_j$ in this $\mathbf{e}$-*basis expansion of* $\mathbf{v}$ is just $\mathbf{e}_j\cdot\mathbf{v}$, the $\mathbf{e}_j$-component of $\mathbf{v}$. Eq. (2-1) also suggests something that is quite valid about how we actually specify vectors in practice: We generally specify a vector $\mathbf{v}$ by giving its components relative to some "known" basis, which in turn need *not* be further specified. The point is that we can't really specify a length and direction for $\mathbf{v}$ unless we say "length relative to what" and "direction relative to what," and a *basis* is the "what." So the actual specification of a vector in N dimensions ultimately comes down to specifying N ordinary numbers — the components of $\mathbf{v}$ relative to some particular basis.

Now at this point, if you're not a mathematician, you're probably thinking that all this formalism may be okay, but pictures are better: One picture is worth a thousand mathematicians! In fact, you probably suspect that even though the mathematician didn't *draw* pictures, he/she was *thinking* pictures! But notice something: The mathematician's formal vectors are a little more general than our picture vectors: the dimensionality N of the mathematician's vector space can be *any positive integer*, whereas we can draw picture vectors only for $N \leq 3$. Thus, the mathematician has come up with a significant generalization of the vector concept, one that we will in fact make use of in our subsequent discussions. However, when we speak of "generalized" vectors in what follows, we shall mean something *more* than arbitrary dimensionality. *For generalized vectors, the scalars are complex numbers instead of real numbers.* And then it's "goodbye pictures," even in two dimensions. For generalized vectors, we have no choice but to use the abstract, formal approach of the mathematician. Fortunately, though, *we can lay out the theory of generalized vectors by repeating almost verbatim the formal theory of ordinary vectors.* Here is how it goes: [From now on we shall use the word "vector" to mean "generalized vector," and not a directed line segment.]

We define a *vector space* to be a collection of objects ψ, called *vectors*, such that the following things [(A) through (D)] are true:

(A) "Scalar multiplication" is defined: If ψ is any vector in the space and c is any *complex number*, also called a *scalar*, then $c\psi$ ($\equiv \psi c$), called the product of c and ψ, is also a vector in the space.

(B) "Vector addition" is defined: If ψ_1 and ψ_2 are any two vectors in the space, then $\psi_1+\psi_2$, called the sum of ψ_1 and ψ_2, is also a vector in the space. And $\psi_2+\psi_1$ is the same vector as $\psi_1+\psi_2$.

Comment: From (A) and (B) we see that, for any two vectors ψ_1 and ψ_2 in the space and any two scalars c_1 and c_2, the *linear combination* $c_1\psi_1+c_2\psi_2$ is a well-defined vector in the space.

(C) There exists in the space *unit vectors*, relative to which all vectors in the space have *components*. The component of the vector ψ relative to the unit vector ε, called the *ε-component of* ψ, is written (ε,ψ), and has the following properties:

(C1) (ε,ψ) is a scalar.

(C2) $(\varepsilon, c_1\psi_1+c_2\psi_2) = c_1(\varepsilon,\psi_1) + c_2(\varepsilon,\psi_2)$.

(C3) $(\varepsilon,\varepsilon) = 1$.

(C4) $(\varepsilon,\varepsilon') = (\varepsilon',\varepsilon)^*$, for any two unit vectors ε and ε'.

Comment: (C1) says that the ε-component of any vector ψ is a *complex number*. (C2) says that the ε-component of any linear combination vector is the same linear combination of the ε-components. (C3) says that the ε-component of itself is unity, and thus serves to *define* a unit vector. And (C4) says that for any two unit vectors ε and ε', the ε-component of ε' is equal to the *complex conjugate* of the ε'-component of ε. The conjugation operation here marks an interesting and curious departure from the theory of ordinary vectors. [But notice that we *could* have written (c4) in the ordinary vector theory as $\mathbf{e}\cdot\mathbf{e}' = (\mathbf{e}'\cdot\mathbf{e})^*$ without altering anything, since complex conjugation doesn't change real numbers.]

(D) If the vector space is N-dimensional, then there exists at least one set of N unit vectors $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N\}$ called a *basis* such that

(D1) $(\varepsilon_j,\varepsilon_k) = 0$ whenever $j \neq k$.

(D2) For any vector ψ in the space,

$$\psi = \varepsilon_1(\varepsilon_1,\psi) + \varepsilon_2(\varepsilon_2,\psi) + \ldots + \varepsilon_N(\varepsilon_N,\psi) \equiv \Sigma_j \varepsilon_j(\varepsilon_j,\psi). \tag{2-2}$$

Comment: (D1) stipulates that the unit vectors composing a basis be *mutually orthogonal*. (D2) says that any vector ψ can be written as a linear combination of the basis vectors, and that the scalar coefficient of ε_j in this *ε-basis expansion of* ψ is just (ε_j,ψ), the ε_j-component of ψ.

In analogy with ordinary vectors, we can schematically represent the expansion (2-2) in two dimensions by the diagram in Fig. 2-4; however, we must bear in mind that this is *not* an actual

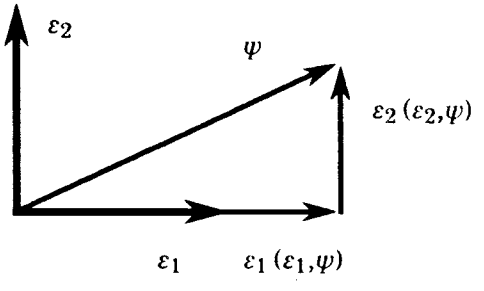


FIG. 2-4. Schematic representation of the expansion (2-2) for $N=2$.

representation of the situation. The reason, of course, is that the components (ε_j,ψ) are generally *complex* numbers. But, if we can't draw an accurate picture of ψ, how can we specify that vector? Just as with an ordinary vector $\mathbf{v}$ in (2-1), we can specify a generalized vector ψ by giving its components, (ε_1,ψ), (ε_2,ψ), ... , (ε_N,ψ), relative to some particular basis; the only difference is that these N components are now complex numbers instead of real numbers.

In our later work with generalized vectors we shall have occasion to use *two theorems*. One is the

Component Expansion Theorem: If ψ is any vector, ε' is any unit vector, and $\{\varepsilon_j\}$ is any basis, then the ε'-component of ψ can be "expanded in the basis $\{\varepsilon_j\}$" according to the formula

$$(\varepsilon',\psi) = \Sigma_j\,(\varepsilon',\varepsilon_j)\,(\varepsilon_j,\psi), \tag{2-3}$$

where the sum runs over the dimensionality of the space. [*Mnemonic*: Sum over the "inside" ε_j's.]

Proof:
$$\begin{aligned}(\varepsilon',\psi) &= (\varepsilon', \Sigma_j\,\varepsilon_j\,(\varepsilon_j,\psi)) && \text{[expanding } \psi \text{ in the } \{\varepsilon_j\}\text{-basis according to (2-2)]}\\ &\equiv (\varepsilon', \Sigma_j\,c_j\,\varepsilon_j) && \text{[}(\varepsilon_j,\psi)\equiv c_j \text{ is some scalar]}\\ &= \Sigma_j\,c_j\,(\varepsilon',\varepsilon_j) && \text{[invoking the fundamental component property (C2)]}\\ &\equiv \Sigma_j\,(\varepsilon_j,\psi)\,(\varepsilon',\varepsilon_j) \equiv \Sigma_j\,(\varepsilon',\varepsilon_j)\,(\varepsilon_j,\psi).\end{aligned}$$

It's easier, in fact *recommended*, that you remember this important proof in the *two-step* form:

$$(\varepsilon',\psi) = (\varepsilon', \Sigma_j\,\varepsilon_j\,(\varepsilon_j,\psi)) = \Sigma_j\,(\varepsilon',\varepsilon_j)\,(\varepsilon_j,\psi).$$

The other theorem we shall need later on is the

Generalized Pythagorean Theorem: If ε' is any unit vector and $\{\varepsilon_j\}$ is any basis, then the sum of the square moduli of all the $\{\varepsilon_j\}$-components of ε' is one; i.e.,

$$\Sigma_j\,|(\varepsilon_j,\varepsilon')|^2 = 1. \tag{2-4}$$

●*Exercise 2-1*. Prove the generalized Pythagorean theorem. [*Hint*: First apply the component expansion theorem with ψ replaced by ε'. Then make use of properties (C3) and (C4).]

The name of this second theorem comes from the fact that it generalizes the following familiar result: For any two-dimensional real unit vector $\mathbf{e}'$, as shown in Fig. 2-5,

$$(\mathbf{e}_1\cdot\mathbf{e}')^2 + (\mathbf{e}_2\cdot\mathbf{e}')^2 = 1^2.$$

But notice in (2-4) that we sum, not the squares of the components of ε', but their *square moduli*.

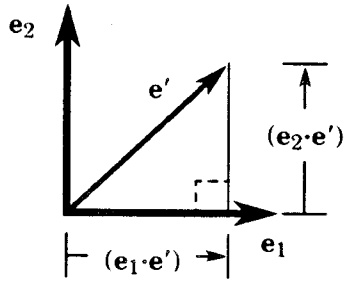


FIG. 2-5. A two-dimensional unit vector $\mathbf{e}'$ and a basis $\{\mathbf{e}_1,\mathbf{e}_2\}$.

▸2.2 Operators, Linearity, Eigenvectors and Eigenvalues, Eigenbases

In calculus you learned that a *function* f transforms a number x into a new number, $f(x)$. Similarly, an *operator* $\mathbf{O}$ transforms a vector ψ into a new vector, written $\mathbf{O}\psi$. In the picture terminology of ordinary vectors, we can loosely think of $\mathbf{O}$ as generally "stretching" and "rotating" the vector ψ into a new vector $\mathbf{O}\psi$.

If the action of $\mathbf{O}$ on a particular vector ϕ is a pure "stretch" by a scalar factor o,

$$\mathbf{O}\phi = o\phi, \tag{2-5}$$

then we say that ϕ is an *eigenvector* of $\mathbf{O}$, and o is the corresponding *eigenvalue*.

●*Exercise 2-2*. If $\mathbf{O}$ is the rule "multiply the given vector by the scalar $(3+i4)$," find all the eigenvectors and eigenvalues of $\mathbf{O}$.

•*Exercise 2-3.* If **O** is the rule "add to the given vector the vector (i2)β," where β is some vector, show that β is an eigenvector of **O**, and find the corresponding eigenvalue.

In quantum mechanics, we will *only* have to deal with operators **O** that have the following two special properties:

(a) **O** is *linear*. This means that, for any two vectors ψ_1 and ψ_2 and any two scalars c_1 and c_2,

$$\mathbf{O}(c_1\psi_1 + c_2\psi_2) = c_1\mathbf{O}\psi_1 + c_2\mathbf{O}\psi_2. \tag{2-6}$$

(b) **O** has an *eigenbasis*. This means that there is some basis $\{\phi_1, \phi_2, \ldots\}$, each vector of which is an eigenvector of **O**. We shall denote the eigenvalue corresponding to eigenvector ϕ_j by o_j. We call the basis $\{\phi_1, \phi_2, \ldots\}$ the *eigenbasis* of **O**, and the scalars $\{o_1, o_2, \ldots\}$ the *eigenvalue set* of **O**. Thus we have

$$(\phi_j,\phi_j) = 1 \text{ for all } j, \tag{2-7a}$$

$$(\phi_j,\phi_k) = 0 \text{ for all } j \neq k, \tag{2-7b}$$

$$\psi = \Sigma_j \phi_j (\phi_j,\psi) \text{ for any vector } \psi, \tag{2-7c}$$

as with any basis, but additionally,

$$\mathbf{O}\phi_j = o_j\phi_j \text{ for all } j. \tag{2-7d}$$

To *define* an operator **O**, we must specify how it produces, for any given vector ψ, the vector $\mathbf{O}\psi$. Of the many ways in which we might do that, one especially simple way invokes the eigenbasis and eigenvalue set of **O**. To show that a knowledge of the eigenbasis and eigenvalue set of **O** suffices to determine $\mathbf{O}\psi$ for any given ψ — and hence suffices to *define* **O** — we reason as follows:

$$\begin{aligned}
\mathbf{O}\psi &= \mathbf{O}\,\Sigma_j \phi_j(\phi_j,\psi) && \text{[expand } \psi \text{ in the eigenbasis of } \mathbf{O}\text{]} \\
&= \Sigma_j \mathbf{O}\phi_j(\phi_j,\psi) && \text{[invoke the linearity of } \mathbf{O}\text{]} \\
&= \Sigma_j o_j\phi_j(\phi_j,\psi) && \text{[invoke the eigenvector condition]} \\
\mathbf{O}\psi &= \Sigma_j \phi_j[o_j(\phi_j,\psi)]. && (2\text{-}8a)
\end{aligned}$$

Equation (2-8a) tells us that the ϕ_j-component of vector $\mathbf{O}\psi$ is $o_j(\phi_j,\psi)$; i.e.,

$$(\phi_j,\mathbf{O}\psi) = o_j(\phi_j,\psi). \tag{2-8b}$$

Thus, given the ϕ_j-component of any vector ψ, we can obtain the ϕ_j-component of vector $\mathbf{O}\psi$ by simply multiplying the former by the corresponding eigenvalue o_j, as illustrated in Fig. 2-6.

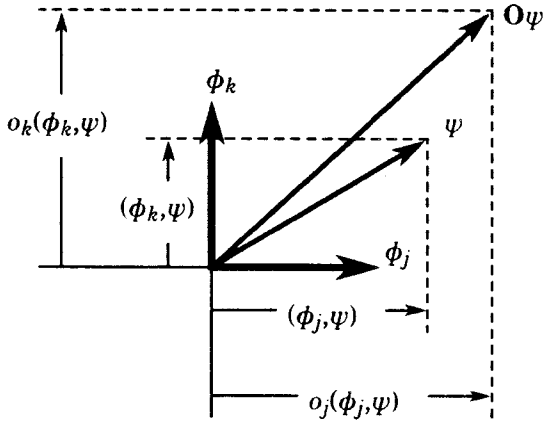


FIG. 2-6. Illustrating schematically how to construct vector $\mathbf{O}\psi$ for any given vector ψ by using the eigenbasis $\{\phi_1, \phi_2, \ldots\}$ and eigenvalue set $\{o_1, o_2, \ldots\}$ of **O**.

•*Exercise 2-4.* Derive equation (2-8b) directly by first substituting (2-7c) on the left (first change the summation index to k), and then applying (2-6), (2-7d), (C2), (2-7b) and (2-7a), in that order.

▸2.3 Summary

The table below summarizes generalized vector theory as we shall require it for our development of quantum mechanics. *Before* the next lecture, you should do three things: First, make sure you understand everything in the summary table. Second, work Exercises 2-1 through 2-4; if you can master them, then you'll be able to handle all the mathematics in the remaining lectures. Finally, read again Sec. 1.1 to remind yourself of the FPOM; in the next lecture, we'll use the theory of generalized vectors to show how quantum mechanics proposes to answer the FPOM for $t \approx 0$.

TABLE SUMMARIZING GENERALIZED VECTOR THEORY

name	item	comment
A *vector*:	ψ	A "very abstract" arrow.
A *scalar*:	c	Any *complex* number.
A linear combination:	$c_1\psi_1 + c_2\psi_2$	A vector in the space.
The *component* of ψ relative to the *unit vector* ε:	(ε,ψ)	A scalar, i.e., a *complex* number. The "ε-component of ψ."
Key properties thereof:	$(\varepsilon, c_1\psi_1+c_2\psi_2) = c_1(\varepsilon,\psi_1) + c_2(\varepsilon,\psi_2)$	Used again and again.
	$(\varepsilon,\varepsilon) = 1$	*Defines* a unit vector ε.
	$(\varepsilon,\varepsilon') = (\varepsilon',\varepsilon)^*$	The curious conjugation.
[1] A *basis* $\{\varepsilon_1, \varepsilon_2, \dots\}$:	$(\varepsilon_j,\varepsilon_j) = 1$ for all j	Unit vectors,
	$(\varepsilon_j,\varepsilon_k) = 0$ for all $j \neq k$	mutually orthogonal.
	$\psi = \Sigma_j \varepsilon_j (\varepsilon_j,\psi)$ for all ψ	The "$\{\varepsilon_j\}$-expansion of ψ."
[3] Component expansion thm.:	$(\varepsilon',\psi) = \Sigma_j (\varepsilon',\varepsilon_j)(\varepsilon_j,\psi)$	Sum over "inside" ε_j's.
[3] Pythagorean theorem:	$\Sigma_j \lvert(\varepsilon_j,\varepsilon')\rvert^2 = 1$	For any unit vector ε'.
An *operator*:	$\mathbf{O}$	Transforms ψ into $\mathbf{O}\psi$.
Linearity of $\mathbf{O}$:	$\mathbf{O}(c_1\psi_1 + c_2\psi_2) = c_1\mathbf{O}\psi_1 + c_2\mathbf{O}\psi_2$	Used again and again.
[2] The *eigenbasis* $\{\phi_j\}$	$\{\phi_1, \phi_2, \dots\}$ is a basis	See above.
and *eigenvalue set* $\{o_j\}$	$o_1, o_2, \dots$ are scalars	I.e., complex numbers.
of $\mathbf{O}$:	$\mathbf{O}\phi_j = o_j\phi_j$ for all j	The eigenvector condition.
[3] Effect of $\mathbf{O}$ on any ψ:	$(\phi_j,\mathbf{O}\psi) = o_j(\phi_j,\psi)$	See Fig. 2-6.

Notes:

[1] Any vector ψ in N dimensional space can be specified by N complex numbers, namely, the N components of ψ, (ε_1,ψ), (ε_2,ψ), ..., (ε_N,ψ), relative to some basis $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N\}$.

[2] Any linear operator $\mathbf{O}$ defined on an N dimensional vector space can be specified by specifying its eigenbasis $\{\phi_1, \phi_2, \dots, \phi_N\}$ and eigenvalue set $\{o_1, o_2, \dots, o_N\}$.

[3] A "theorem," derivable from the preceding definitions and properties.

LECTURE 3. QUANTUM MECHANICS: STATICS I

▸3.1 States and Observables, Vectors and Operators

In our first lecture, we said that "mechanics" deals with systems and their observables. If A is an *observable* (say, position) of a certain *system* (say, a ball), then we can at any time measure A on the system and get a *result* A (say, 4.7), which we call the *value* of A. (So, the measured value of the ball's position is 4.7.) We also noted that the chief concern of mechanics is to answer the following general kind of question (the FPOM): If we measure an observable A on a system with result A, and then a time t later measure observable B on the system, what can we predict about the result B of the second measurement?

The answer to this question proffered by *classical mechanics* is generally framed in terms of the system's "classical state." For example, suppose our system is a simple particle of mass m that moves on the x-axis in a potential energy field $V(x)$ [or equivalently, in a force field $F(x) = -V'(x)$]. Classical mechanics defines the *state* of this system at time t to be the pair of variables $[x(t),p(t)]$, the particle's position and momentum at time t. This definition of state is motivated by two considerations: First, a knowledge of the particle's instantaneous state $[x(t),p(t)]$ allows us to calculate the instantaneous values of all other meaningful observables, such as the particle's velocity $p(t)/m$, or its energy $p^2(t)/2m + V(x(t))$. And second, a knowledge of the particle's *initial* state $[x(0),p(0)]$ provides precisely the information needed (two integration constants) to uniquely solve Newton's equation $F = ma$,

$$m\mathrm{d}^2x/\mathrm{d}t^2 = -V'(x), \tag{3-1}$$

for the particle's state $[x(t),p(t) \equiv m\mathrm{d}x(t)/\mathrm{d}t]$ at any time $t > 0$. In general, then, the answer to the FPOM from the classical viewpoint hinges on how well we can know the system's *state*, which classical mechanics *defines* to be the values of some fixed set of key observables.

We noted in our first lecture that this classical approach to the FPOM works very well for *macro*scale systems, but not for *micro*scale systems. We explained that the failure of classical mechanics on the microscale is quite profound, and arises basically from the fact that observables of microscale systems (such as the "position" of an electron) cannot always be said to "have values." And we don't mean by this that sometimes we are just ignorant of an observable's value; we mean that sometimes we will involve ourselves in a *contradiction* if we even assume the *existence* of such a value. (Recall our discussion of the double-slit experiment.) The replacement for classical mechanics on the microscale is called *quantum mechanics*, and is the subject of this and the following four lectures.

In essence, the "standard theory" of quantum mechanics takes the following tack. It proposes to *retain* the classical idea that a physical system always has a definite state, but to *reject* the classical *definition* of "state" as the instantaneous values of some fixed set of key observables. For example, quantum mechanics would go along with classical mechanics in saying that a microscale particle on the x-axis is always in a well-defined state, but quantum mechanics would not define that state to be the instantaneous values of the particle's position and momentum; indeed, it *couldn't*, because those values won't always exist. And how does quantum mechanics propose to *separate* the notions of "state" and "observable"? By doing something very un-intuitive, something so very outlandish that the only rationale we can give for it is simply the fact that *it ultimately works*! Quantum mechanics turns to the abstract, mathematical theory of generalized vector spaces and says, let's represent *states* by *vectors*, and *observables* by *operators*. More precisely, quantum mechanics starts by laying down the following two *rules* (or axioms or postulates):

▸ **Rule 1**: Corresponding to any isolated physical system there is a generalized vector space. The *unit vectors* in this space "represent" the possible *physical states* of the system. The particular unit vector representing the system state at time t is written Ψ_t, and is called the *state vector* of the system at time t; we say that the system is "in the state Ψ_t" at time t.

▸ **Rule 2**: Each system *observable* A is "represented" by a *linear operator* **A** that has an *eigenbasis* $\{a_1, a_2, \ldots\}$ in the system's generalized vector space, and a *real* eigenvalue set $\{A_1, A_2, \ldots\}$; that is, the linear operator **A** representing the observable A is such that

$$\mathbf{A}a_j = A_j a_j \quad (j = 1, 2, \ldots) \tag{3-2}$$

where the eigenvectors $\{a_1, a_2, \ldots\}$ form a basis in the system's generalized vector space, and where the eigenvalues $\{A_1, A_2, \ldots\}$ are all real numbers.

Now at this point, you must be *very patient*, because these rules don't just look abstract, they look pointless! What does it mean to say that the vector Ψ_t "represents" the system's state, or the operator **A** "represents" the observable A? How should we go about *finding* either that vector or that operator? Why should we even *want* to? In truth, Rules 1 and 2 are just getting things set up for Rules 3 and 4, which we'll write down in a minute, and which will answer most of those very reasonable questions. But before we go on to those next rules, let's clarify some implications of Rules 1 and 2 that will turn out later to be important.

First, Rule 1 implies that every unit vector in the system's generalized vector space corresponds to a possible physical state of the system, and conversely, every possible physical state of the system corresponds to some unit vector in the system's generalized vector space. (The correspondence between physical states and unit vectors is *not quite* one-to one, but we're not going to worry about that technicality here.) In writing the system's state vector at time t as Ψ_t, Rule 1 obviously suggests that *the state vector evolves with time*; i.e., for two different times t and t', Ψ_t and $\Psi_{t'}$ will usually be two different unit vectors in the space. The time-dependence of the state vector Ψ_t is described by Rule 5, but we won't get to Rule 5 until our Lecture 6. In this and the next two lectures, we'll be concerned solely with what can be said about the system at a single instant t — i.e., with the "static" features of quantum mechanics.

Rule 2 similarly implies that there is an essentially one-to-one correspondence between physical observables on the one hand, and linear operators with an eigenbasis and real eigenvalue set on the other. Of course, we may not know, or even care, how to *measure* all those observables. In fact, it is the high calling of the physicist as a "diviner of Nature" to identify the physically relevant observables of a system, and to mathematically construct their associated operators. Notice that, in the process of constructing operator **A**, the physicist will also construct, either explicitly or implicitly, the associated basis $\{a_1, a_2, \ldots\}$, and hence the entire generalized vector space of the system. Clearly, this is not a trivial task! In Lecture 5 we'll see how to construct the operators **X** and **P** that correspond to the position and momentum of a particle on the x-axis. But mostly in these lectures, we'll just be talking about what you can do with observable operators once you have them.

Notice that neither the general observable operator **A**, nor its eigenbasis or eigenvalues, depend upon time. In the formulation of quantum mechanics that we are considering here, only the system's state vector evolves with time.

Since the **A**-eigenvectors $\{a_1, a_2, \ldots\}$ are a basis in the system's state space, it follows that the system's instantaneous state vector Ψ_t can always be "expanded in the **A**-eigenbasis" as

$$\Psi_t = \Sigma_j a_j (a_j, \Psi_t), \tag{3-3}$$

where (a_j, Ψ_t), the a_j-component of Ψ_t, is just some complex number. We'll see later that such expansions of the state vector in the eigenbasis of observable operators play an important role in quantum theory. Also important will be the fact that, since Ψ_t is a *unit* vector, then by the (generalized) Pythagorean theorem the sum of the square moduli of its $\{a_j\}$-components must equal unity:

$$\Sigma_j |(a_j, \Psi_t)|^2 = 1. \tag{3-4}$$

We're not specifying the summation limits in (3-3) and (3-4), because the dimensionality of the generalized vector space is not the same for all systems. Some systems require only a two dimensional vector space, but most systems (even a simple particle on the x-axis) call for an infinite dimensional vector space.

If the observable A is the first-measured observable in our statement of the FPOM, then by Rule 2 the second-measured observable B will be similarly represented by some linear operator **B** with eigenbasis $\{\beta_1, \beta_2, \ldots\}$ and real eigenvalue set $\{B_1, B_2, \ldots\}$. And just as with Eqs. (3-3) and (3-4), we can always expand the system's state vector Ψ_t in the **B**-eigenbasis according to

$$\Psi_t = \Sigma_k \beta_k (\beta_k, \Psi_t),$$

where the components of Ψ_t relative to the **B**-eigenbasis are complex numbers satisfying

$$\Sigma_k |(\beta_k, \Psi_t)|^2 = 1.$$

If the **A**-eigenbasis components of Ψ_t happen to be known, then we can calculate from them the **B**-eigenbasis components of Ψ_t by using the component expansion theorem [see Eq. (2-3)]:

$$(\beta_k, \Psi_t) = \Sigma_j (\beta_k, \alpha_j)(\alpha_j, \Psi_t). \tag{3-5}$$

Here, the coefficient (β_k, α_j) is of course the β_k-component of the vector α_j in the expansion

$$\alpha_j = \Sigma_k \beta_k (\beta_k, \alpha_j). \tag{3-6}$$

●*Exercise 3-1.* Eq. (3-6) is the expansion of the **A**-eigenvector α_j in the **B**-eigenbasis. Write down the expansion of the **B**-eigenvector β_k in the **A**-eigenbasis.

●*Exercise 3-2.* Prove the following two relations between the β_k-component of vector α_j and the α_j-component of vector β_k:

$$(\beta_k, \alpha_j) = (\alpha_j, \beta_k)^*, \tag{3-7a}$$

$$|(\beta_k, \alpha_j)|^2 = |(\alpha_j, \beta_k)|^2. \tag{3-7b}$$

●*Exercise 3-3.* What is the formula for calculating the **A**-eigenbasis components of Ψ_t from the **B**-eigenbasis components of Ψ_t? How are the coefficients in this formula related to the coefficients in formula (3-5)?

▸ 3.2 Results and Effects of Measurements

Now that we have "set the stage" with Rules 1 and 2, let's "raise the curtain and start the play:" Let's address the crucial question of what happens when we measure a given observable **A** on a system in a known state Ψ_t. The answer to that question comes in two parts. The first part, which we'll call Rule 3, deals with *predicting the result of a measurement.*

▸ **Rule 3**: If observable **A** is measured on a system in state Ψ_t, then the *most* that can be predicted about the result of that measurement is this: The *probability* that the result will be the eigenvalue A_j is equal to the square modulus of the α_j-component of Ψ_t, namely $|(\alpha_j, \Psi_t)|^2$.

Rule 3 implies that even though the state of the system is completely specified, in that the state vector Ψ_t is known, we can nevertheless make only *probabilistic* predictions about measurement results. This is in sharp contrast to the situation in classical mechanics, where the (classical) state uniquely determines the result of any observable measurement. But in quantum mechanics, all we can say prospectively about a measurement of **A** on the system in state Ψ_t is that the probability that the result will be A_1 is $|(\alpha_1, \Psi_t)|^2$, the probability that the result will be A_2 is $|(\alpha_2, \Psi_t)|^2$, etc.

The addition law for probabilities implies that we can calculate the probability that an **A**-measurement on state Ψ_t will yield *any* of the eigenvalues of **A** — i.e., either A_1 or A_2 or A_3 ... — by simply adding up all the individual probabilities. Thus,

$$\text{Prob}\{A_1 \text{ or } A_2 \text{ or } \ldots\} = |(\alpha_1, \Psi_t)|^2 + |(\alpha_2, \Psi_t)|^2 + \ldots = \Sigma_j |(\alpha_j, \Psi_t)|^2.$$

But Eq. (3-4) tells us that the sum on the right is one, which in probability theory means "absolute certainty." Thus, a measurement of **A** is *certain* to yield *some* one of the eigenvalues of **A**. In other words, the eigenvalues $\{A_1, A_2, \ldots\}$ introduced in Rule 2 are the *only* values that any measurement of **A** can ever yield. Now, depending upon the observable, eigenvalue sets can be either *discretely* distributed (like the integer numbers) or *continuously* distributed (like the real numbers). Historically, the fact that many observables on the microscale (such as the energy of an electron inside an atom) turn out to have discrete or "quantized" values, is what led to the name "quantum" mechanics. (Of course, you can see already that the difference between classical and quantum mechanics runs much deeper that quantized measurement results.) In our work here we are going to

pretend, insofar as possible, that *all* observable eigenvalue sets are discretely distributed, because that's the easiest case to treat mathematically.

We have seen how Rule 3 illuminates the significance of the *eigenvalues* of the operator **A** corresponding to the observable A: those eigenvalues are just the allowed results of an A-measurement. Rule 3 also sheds some light on the significance of the *eigenvectors* of **A** as well: It says that if we expand the system's state vector Ψ_t in the **A**-eigenbasis, as in Eq. (3-3), then the square modulus of the coefficient of eigenvector a_j is the probability of measuring the corresponding eigenvalue A_j. And this in turn implies another interesting fact:

> If the system's state vector Ψ_t coincides with the eigenvector a_k, then an A-measurement on the system is *certain* to yield the eigenvalue A_k.

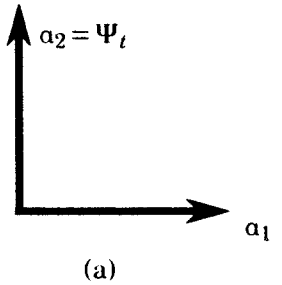
To prove this statement, we simply observe from Rule 3 that an A-measurement on the system in the state $\Psi_t = a_k$ will yield eigenvalue A_j with probability $|(a_j, \Psi_t)|^2 = |(a_j, a_k)|^2$. This probability is *zero* if $j \neq k$, owing to the orthogonality of different eigenbasis vectors, and *one* if $j = k$, owing to the fact that a_k is a unit vector. So we see that Rule 3 implies that the eigenbasis vectors of **A** define the possible states of the system for which an A-measurement will have a *uniquely predictable result* (namely the corresponding eigenvalue of **A**).

Further insight into the significance of the eigenbasis vectors of **A** is provided by Rule 4. This rule deals with *the effect of a measurement on the state.*

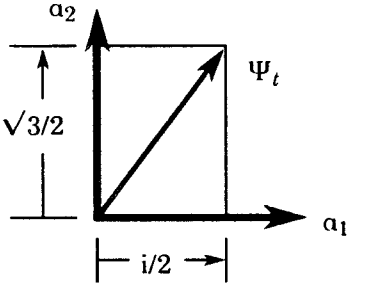
▸ **Rule 4**: If a measurement of A *does* yield eigenvalue A_j, then *immediately after* that measurement the system's state vector will coincide with the corresponding eigenvector a_j, regardless of what the state vector was just before the measurement.

Rule 4 implies that when we measure A on a system, the system's state vector immediately "jumps" to one of the eigenbasis vectors of **A**, namely to that eigenbasis vector corresponding to the measured eigenvalue. This is another dramatic departure from classical mechanics, where a measurement (of the ideal kind we're considering here) has no effect on the state. Notice that in quantum mechanics a measurement result tells us much more about the state of the system immediately *after* the measurement than immediately before: If the measurement result is A_j, then immediately *after* the measurement we know that the state vector of the system is a_j; however, all we can infer about the state immediately *before* the measurement is that its a_j-component was not zero (otherwise, we couldn't have got the result A_j).

●*Exercise 3-4.* Suppose our system's generalized vector space is two dimensional, so that the eigenbasis of **A** consists of the two orthogonal unit vectors a_1 and a_2.



(a)



(b)

(a) If $\Psi_t = a_2$, as indicated schematically in (a) above, calculate the probability that an A-measurement will yield the value A_1. Ditto for the value A_2. What will the state vector of the system be immediately *after* an A-measurement?

(b) If $\Psi_t = (i/2)a_1 + (\sqrt{3}/2)a_2$, as indicated schematically in (b) above, calculate the probability that an A-measurement will yield the value A_1. Ditto for the value A_2. For each outcome, describe the state vector immediately after the measurement.

▸ 3.3 Answer to the FPOM for $t \approx 0$

With Rules 1 through 4, you have now seen the most bizarre features of quantum mechanics; you have arrived at the "radical core" of the theory! Also with these four rules, you are now in a position to see how quantum mechanics proposes to answer the FPOM for $t \approx 0$, i.e., for no appreciable time lapse between the A-measurement and the subsequent B-measurement. We reason as follows:

- When A is measured on the system, the result must (by Rules 2 and 3) be some A-eigenvalue; let's call the measured eigenvalue A_j.

- So immediately after the A-measurement, the system will (by Rule 4) be in the state a_j. And since $t \approx 0$, the system will *still* be in state a_j for the B-measurement.

- Measuring B on the state a_j will (by Rule 3) yield any B-eigenvalue B_k with probability $|(\beta_k, a_j)|^2$.

▸ Therefore, quantum theory's answer to the FPOM in the $t \approx 0$ case is:

$$\text{Prob}\{\mathsf{B} = B_k, \text{ given that } \mathsf{A} = A_j \text{ and } t \approx 0\} = |(\beta_k, a_j)|^2. \quad (k = 1, 2, \ldots) \tag{3-8}$$

Notice that all we need to know to answer the FPOM for $t \approx 0$ are:

- the eigenvalues of **A** and **B**, and
- the components of the **A** eigenvectors relative to the **B** eigenbasis.

Notice also, from Eq. (3-7b), that β_k and a_j in (3-8) can be interchanged.

As a simple application of the result (3-8), suppose that A and B are the *same observable* (call it A). Then (3-8) implies that the probability that the second A-measurement will give the result A_k when the first A-measurement gave the result A_j is

$$\begin{aligned} |(a_k, a_j)|^2 &= 0 \quad \text{if } k \neq j \quad (\text{since } a_k \text{ and } a_j \text{ are orthogonal for } k \neq j) \\ &= 1 \quad \text{if } k = j \quad (\text{since } a_j \text{ is a unit vector}) \end{aligned}$$

Thus, the second A-measurement is *certain* to give the *same* result A_j as the first A-measurement. More generally, *the results obtained in two rapid, successive measurements of any observable will always agree with each other*. So quantum mechanics is not *totally* crazy! But notice that this consistency of immediate remeasurement results depends strongly on the "measurement jump" of the state vector postulated by Rule 4: After the first A-measurement, with whatever result A_j, the state vector *must* coincide with the corresponding eigenvector a_j in order for Rule 3 to guarantee the result A_j on the second A-measurement.

●*Exercise 3-5.* Suppose A and B are observables of a system with a two-dimensional state space, and suppose the B-eigenbasis vectors are given in terms of the A-eigenbasis vectors by

$$\beta_1 = i(1/3)^{1/2} a_1 + (2/3)^{1/2} a_2 \quad \text{and} \quad \beta_2 = (2/3)^{1/2} a_1 + i(1/3)^{1/2} a_2.$$

(a) By inspecting the above formulas, identify the components (a_j, β_k) for all j and k. Then use Eq. (3-7a) to deduce (β_k, a_j) for all j and k. Using the latter numbers and Eq. (3-6), write down a_1 and a_2 in terms of β_1 and β_2. Check your answer by solving the above two equations simultaneously for a_1 and a_2.

(b) In the FPOM, suppose the A-measurement yields the result A_2. What is the probability that the immediately subsequent B-measurement will yield the result B_1? The result B_2?

(c) Suppose the B-measurement gives the result B_2, and then an immediate remeasurement of A is made. What is the probability that the result of this second A-measurement will give the same value A_2 as obtained in the first measurement?

LECTURE 4. QUANTUM MECHANICS: STATICS II

▸4.1 Recapitulation

Let's review our story so far.

Classical mechanics founders on the microscale, essentially because it tries to identify the enduring "state" of a physical system with the not-so-enduring values of certain observables of the system. Quantum mechanics attempts to salvage things by *separating* the notions of "state" and "observable" using the mathematical theory of generalized vectors. Specifically, quantum mechanics postulates, in its Rules 1 and 2, that to every physical system there corresponds an abstract generalized vector space, in which the system's *states* are represented by certain *vectors* and the system's *observables* are represented by certain *operators*. The vectors representing system's states are all *unit* vectors. The operators representing system's observables are all *linear* with *eigenbases* and *real* eigenvalue sets. And just how are we supposed to *find* the operators that represent specific system observables, and that in fact define through their eigenbases the system's abstract vector space? That is something that quantum theory leaves to the wit and wiles of the physicist, who thus is given the challenging task of hooking up the rules of quantum mechanics to the real world.

But we can't expect to form a meaningful physical theory by doing nothing more than associating states with vectors and observables with operators. What ties all these concepts together is the notion of *measurement*. Rules 3 and 4 describe what happens when we make a measurement of an observable A on a system whose state vector is Ψ_t. Those rules say, firstly, that the outcome of such a measurement must be one of the real eigenvalues $\{A_1, A_2, \dots\}$ of the operator **A** associated with observable A. But which eigenvalue? Quantum theory says that usually we can't be sure, and that the *best* we can do is this: Expand the state vector Ψ_t in the eigenbasis $\{a_1, a_2, \dots\}$ of **A**,

$$\Psi_t = \Sigma_j a_j (a_j, \Psi_t), \tag{4-1}$$

and observe the complex component (a_j, Ψ_t) of Ψ_t along the eigenbasis vector a_j; the square modulus of that component, namely $|(a_j, \Psi_t)|^2$, is numerically equal to the *probability* that an A-measurement on the system in the state Ψ_t will give the eigenvalue A_j corresponding to eigenvector a_j. And while we can't always be sure of the result of an A-measurement, we can be sure of this: If the result *is* the particular eigenvector A_j, then the state vector of the system immediately *after* the measurement will coincide with the corresponding eigenbasis vector a_j. In other words, a measurement of A causes the system's state vector to "jump" to one of the eigenvectors of the observable's operator **A**, namely to that eigenvector corresponding to the eigenvalue found in the measurement. This measurement jump, like the measurement result itself, is inherently random and uncontrollable, and does not seem to be explainable in terms of any underlying deterministic mechanism.

That, in brief, is the gist of quantum mechanics as embodied by our Rules 1 through 4. At the end of our last lecture, we showed how these four rules allow us to frame an answer to the FPOM for the "static" case $t \approx 0$. Let's review that answer in the context of a simple hypothetical system whose generalized vector space is two-dimensional. Let's assume that we have defined the operators **A** and **B** for observables A and B by specifying their respective eigenbases, $\{a_1, a_2\}$ and $\{\beta_1, \beta_2\}$, and their corresponding real eigenvalue sets, $\{A_1, A_2\}$ and $\{B_1, B_2\}$. The eigenbases define in turn the generalized vector space of the system. To help us visualize that abstract space, and in particular the relation between the two eigenbases, we'll use the pictorial representation in Fig. 4-1a. We can think of the components of the various vectors relative to each other as"projections," just as we do for ordinary vectors, *provided* we keep in mind the caveat illustrated in Fig. 4-1b: Although the drawing implies that the component of, say, vector β_1 relative to vector a_1 is *equal* to the component of a_1 relative to β_1, those two components are in fact *complex conjugates* of each other; i.e., $(a_1, \beta_1) = (\beta_1, a_1)^*$. Now, the $t \approx 0$ version of the FPOM contemplates a measurement of observable A followed immediately by a measurement of observable B. We know that the A-measurement must yield one of the two eigenvalues A_1 or A_2. If the result is A_1, then regardless of what the state vector of the system was just *before* the A-measurement, it will coincide with a_1 immediately *after* that measurement; hence, the system will be in the state a_1 for the B-measurement. So to predict the

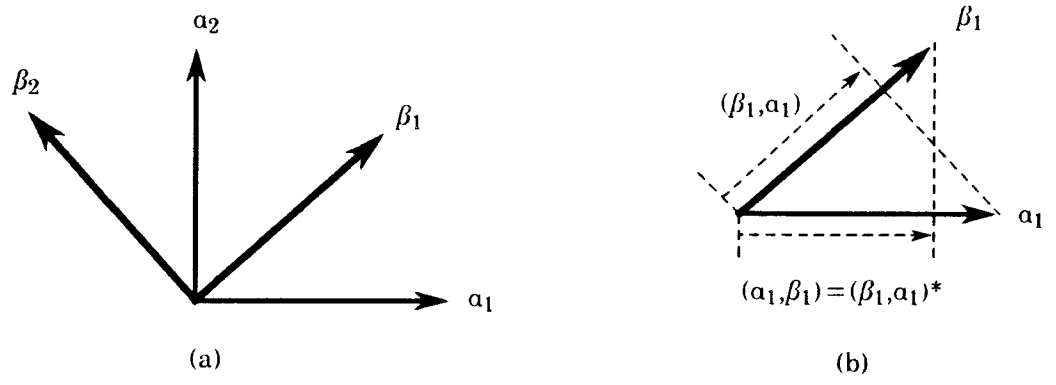


FIG. 4-1. Generalized two-dimensional vector space for a hypothetical system, showing in (a) the eigenbases $\{\alpha_1,\alpha_2\}$ and $\{\beta_1,\beta_2\}$ of two observable operators **A** and **B**, and reminding us in (b) of the inherent limitation of this kind of pictorial representation of generalized vectors.

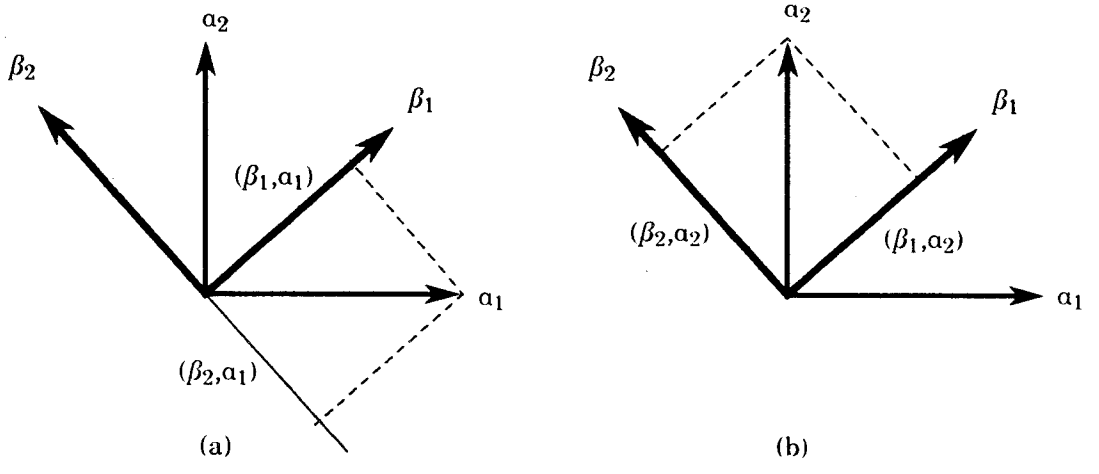


FIG. 4-2. Showing the complex components needed to answer the FPOM for $t \approx 0$, (a) when the result of the A-measurement is A_1, and (b) when the result of the A-measurement is A_2.

result of the B-measurement, we first calculate the complex components (β_1,α_1) and (β_2,α_1) of α_1 relative to the **B**-eigenbasis [see Fig. 4-2a]. We then assert that the B-measurement will yield the result B_1 with probability $|(\beta_1,\alpha_1)|^2$, and the result B_2 with probability $|(\beta_2,\alpha_1)|^2$. On the other hand, if the A-measurement had given the result A_2, then we would know that the system's state vector just before the B-measurement would be α_2. In that case, we would calculate the complex components (β_1,α_2) and (β_2,α_2) of α_2 relative to the **B**-eigenbasis [see Fig. 4-2b], and then assert that the B-measurement will yield the result B_1 with probability $|(\beta_1,\alpha_2)|^2$, and the result B_2 with probability $|(\beta_2,\alpha_2)|^2$. That, in brief, is quantum theory's answer to the FPOM in the "static" case.

Now let's pursue some other implications of the first four rules of quantum mechanics.

▸4.2 Compatible and Incompatible Observables

First we're going to discuss a notion, or if you will a *definition*, that is very revealing of the non-classical character of quantum mechanics. Two observables A and B are said to be *compatible* if and only if, in a rapid sequence of three measurements — of A then of B then of A again — the results of the first and third measurements will *always* agree with each other. But if, in those three measurements, it *might* happen that the first and third measurements will *not* agree, then A and B are said to be *incompatible*.

In classical mechanics this definition is rather useless, because there *all* observables are compatible: The intervening B-measurement, being "ideal," has no effect on the state of the system, and hence no effect on the value of observable A. But in quantum mechanics, it is easy to see how two observables might very well be *incompatible*. For example, consider a system with a two-dimensional state space, and suppose that the eigenbases $\{a_1,a_2\}$ and $\{\beta_1,\beta_2\}$ of the operators for observables A and B do *not* coincide with each other, as is indicated schematically in Fig. 4-3a. If the first

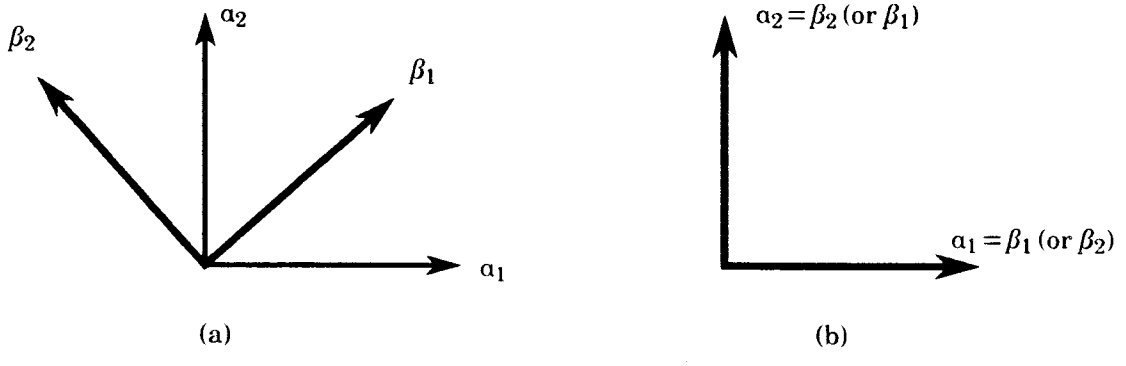


FIG. 4-3. Illustrating, for two observables A and B with repective eigenbases $\{a_1,a_2\}$ and $\{\beta_1,\beta_2\}$, the conditions of (a) incompatibility, and (b) compatibility.

A-measurement yields eigenvalue A_1, then the system's state vector will jump to state a_1. The subsequent B-measurement will make the system's state vector jump to either state β_1 or state β_2. But in neither of those two states is it guaranteed that an A-measurement will give the result A_1 that was obtained in the first measurement. Of course, we *might* get the result A_1 on the third measurement, but that is not a certainty; so by our definition, A and B are *incompatible* observables.

On the other hand, suppose that the eigenbases $\{a_1,a_2\}$ and $\{\beta_1,\beta_2\}$ of the A and B operators *coincide* with each other, as is indicated schematically in Fig. 4-3b (whether the eigenbasis indices match or not is immaterial for our arguments). Now what happens in our A-B-A measurement sequence? If the first A-measurement gives the result A_1, then the system will jump to state a_1. Since a_1 is also an eigenbasis vector of B's operator, then the B-measurement will simply *leave* the system in that state; thus, the remeasurement of A will necessarily yield the eigenvalue A_1 again. Obviously, a similar argument will show that if the first measurement had given the result A_2, then the third measurement would also have to give the result A_2. So A and B in this case are *compatible* observables.

Our conclusions here represent a general result in quantum mechanics: *A necessary and sufficient condition for two observables to be compatible is that their operators have a common eigenbasis.* If the operators for A and B do *not* have a common eigenbasis, then in a rapid A-B-A measurement sequence the B-measurement always has the potential of "spoiling" the remeasurement of A. But it is very important to understand that this spoilage, when it occurs, is *not* the result of sloppy measuring technique, but rather is intrinsic to the nature of the measured observables.

▸ 4.3 Interference. The "Value" of an Observable

A fascinating property of incompatible observables is their apparent tendency to "interfere" with each other. The most famous manifestation of quantum interference occurs in the double-slit experiment, which we discussed in our first lecture. However, the double-slit experiment is an example of "dynamic" interference, since it involves in a fundamental way the passage of time; we shall consider the double-slit experiment in some detail in Lecture 7. The general phenomenon of interference can be demonstrated in a static context rather easily by appealing once again to a hypothetical system with a two-dimensional state space. Suppose A and B are two incompatible observables for such a system, so that the eigenbases $\{a_1,a_2\}$ and $\{\beta_1,\beta_2\}$ of their associated operators **A** and **B** do not coincide, as schematized in Fig. 4-4. Suppose further that the system's instantaneous

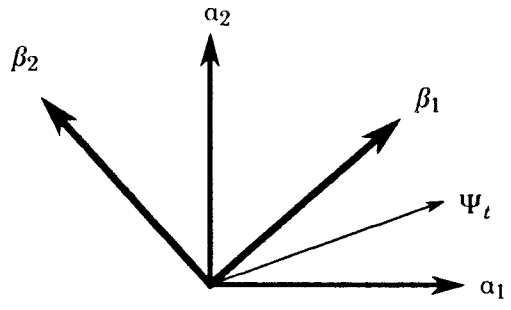


FIG. 4-4. A situation leading to "interference" between the incompatible observables A and B.

state vector Ψ_t does not coincide with *any* of those four eigenbasis vectors. As we know quite well by now, the measurement probabilities for observables A and B on state Ψ_t are as follows:

$$\text{Prob}\{\text{Meas}(\mathsf{A})=A_j\} = |(a_j,\Psi_t)|^2 \ (j=1,2); \quad \text{Prob}\{\text{Meas}(\mathsf{B})=B_k\} = |(\beta_k,\Psi_t)|^2 \ (k=1,2).$$

Now, if we expand Ψ_t in the **A**-eigenbasis,

$$\Psi_t = a_1\,(a_1,\Psi_t) + a_2\,(a_2,\Psi_t),$$

then we can write the β_1-component of Ψ_t in the form

$$(\beta_1,\Psi_t) = (\beta_1, a_1\,(a_1,\Psi_t) + a_2\,(a_2,\Psi_t)) = (\beta_1,a_1)(a_1,\Psi_t) + (\beta_1,a_2)(a_2,\Psi_t),$$

which of course is just the change-of-basis formula (3-5). Therefore, the probability that a B-measurement on state Ψ_t will yield the result B_1 can be computed as

$$\begin{aligned}\text{Prob}\{\text{Meas}(\mathsf{B})=B_1\} &= |(\beta_1,\Psi_t)|^2 \\ &= |(\beta_1,a_1)(a_1,\Psi_t) + (\beta_1,a_2)(a_2,\Psi_t)|^2.\end{aligned}$$

Observing that the right side is just the square modulus of the sum of two complex numbers, we need only invoke the identities (1-8) to get

$$\begin{aligned}\text{Prob}\{\text{Meas}(\mathsf{B})=B_1\} = |(\beta_1,a_1)|^2\,|(a_1,\Psi_t)|^2 + |(\beta_1,a_2)|^2\,|(a_2,\Psi_t)|^2 \\ + 2\text{Re}\{(\beta_1,a_1)(a_1,\Psi_t)(\beta_1,a_2)^*(a_2,\Psi_t)^*\}\end{aligned} \tag{4-2}$$

●*Exercise 4-1.* Carry out the algebra leading to Eq. (4-2).

Notice in Eq. (4-2) that the first two terms on the right are strictly positive (none of the paired unit vectors are orthogonal), while the third term may be either positive or negative but *not* zero.

The third term on the right side of Eq. (4-2) is called the "interference" term. That name is partly a holdover from a way of trying to interpret Eq. (4-2) that was popular in the early days of quantum mechanics. In this (unrecommended) view, one regards A and B as "joint random variables" — i.e., as variables that can *simultaneously* take values which are predictable only in a *probabilistic* sense. For

such variables the event $\mathsf{B}=B_1$ always occurs in conjunction with one of the two mutually exclusive events $\mathsf{A}=A_1$ or $\mathsf{A}=A_2$, so the laws of probability theory imply that

$$\text{Prob}\{\mathsf{B}=B_1\} = \text{Prob}\{\mathsf{B}=B_1 \text{ and } \mathsf{A}=A_1\} + \text{Prob}\{\mathsf{B}=B_1 \text{ and } \mathsf{A}=A_2\}. \tag{4-3}$$

One now tries to view Eq. (4-2) as an instance of Eq. (4-3). The difficulty is that *there is no consistent way of doing that.* In particular, the most plausible association, of the first two terms on the right side of Eq. (4-2) with the corresponding two terms on the right side of Eq. (4-3), leaves one quite unable to account for the third term on the right side of Eq. (4-2). What some people did was to still regard those associations as meaningful, and then regard the "interference" term as a manifestation of the overall quantum mystery.

Without denying the presence of a "quantum mystery," let's recommend another way to resolve the discrepancy posed by Eqs. (4-2) and (4-3). In the circumstance illustrated in Fig. 4-4, where the state vector is such that it is absolutely impossible to predict with certainty what value will be obtained in an A-measurement, does it really make any sense to speak of A as "having a value"? Let's answer this question in the negative; in other words, let's adopt the following position:

> *An observable* A *can be said to "have a value" if and only if the system's state vector guarantees a unique result of measuring* A.

Therefore, if Ψ_t is a linear combination of two or more eigenbasis vectors of operator **A**, then observable A *cannot* be said to have a value. In that circumstance, it is the *act of measuring* A that *develops* an A-value; the measurement does this by causing the system's state vector to jump to one of the **A**-eigenbasis vectors, so that *then* A will have a value. But an A-value does not exist *prior* to the A-measurement. Now, in our interference problem, since it is obviously not possible for Ψ_t to simultaneously coincide with an eigenvector of *both* **A** and **B**, then it is not possible for A and B to *simultaneously* have values; hence, it is not possible to regard A and B as *joint* random variables, *and Eq. (4-3) does not apply.* With Eq. (4-3) thus eliminated, there is no discrepancy. Or, if you prefer, we have replaced the "interference" mystery with the mystery that there are perfectly legitimate system observables that sometimes *have* values and sometimes do *not* have values.

But then, the latter mystery was precisely the conclusion that we seemed to be *forced* to by the results of the double-slit experiment. Referring to Fig. 1-2, we concluded that the results of that experiment implied that any electron reaching screen S_2 with *both* slits y_1 and y_2 open could not plausibly be said to have come through one slit *exclusive* of the other; therefore, such an electron could *not* be said to have had a y-position value when it passed screen S_1. In light of such experimental evidence, it does not seem terribly unreasonable to *disallow* the underlying premise of Eq. (4-3) that incompatible observables will always simultaneously have values.

Before the next lecture you should review this and the preceding lecture, and work all the exercises therein (the five exercises in Lecture 3 and one in the present Lecture 4). We will use the remaining time here to answer any questions you might have about our development so far.

LECTURE 5. QUANTUM MECHANICS: STATICS III

▸5.1 The Position and Momentum Operators. Wave Functions

In this lecture we're going to discuss two *specific* physical observables, the *position* X and the *momentum* P of a particle that is constrained to move in one physical dimension — say along the x-axis. Our first task is to produce the operators **X** and **P** that physicists, in their role as "diviners of Nature," have decreed shall represent those two observables. There are several equivalent procedures for defining the operators **X** and **P**. Our procedure here will be to simply specify their eigenbases and eigenvalue sets; that this will completely define those operators follows from the general result of Exercise 2-5 [see also Fig. 2-6]. In our definition, you will notice that the eigenvalue sets of **X** and **P** are taken to be *continuously* distributed. While that seems very reasonable, it will make for some mathematical complications further on — which is why we have thus far been "pretending" that all observable eigenvalue sets were discretely distributed.

• *DEFINITION OF* **X** *AND* **P**:

- The eigenvalue set of **X** is the set $\{x\}$ of all real numbers x, and the eigenvalue set of **P** is the set $\{p\}$ of all real numbers p.

- The **X**-eigenbasis vector δ_x that corresponds to the eigenvalue x ($\mathbf{X}\delta_x = x\delta_x$), and the **P**-eigenbasis vector ϕ_p that corresponds to the eigenvalue p ($\mathbf{P}\phi_p = p\phi_p$), are such that the component of ϕ_p relative to δ_x is the complex number

$$(\delta_x,\phi_p) = r\,e^{-ixp/\hbar} \equiv r\,[\cos(xp/\hbar) - i\sin(xp/\hbar)], \tag{5-1a}$$

where $\hbar$ is Planck's constant in Eq. (1-2), and r is a real constant whose value will not concern us here. [In a more advanced quantum mechanics course, you'd learn that $r = (2\pi\hbar)^{-1/2}$.]

Notice that we have cleverly specified the two eigenbases $\{\delta_x\}$ and $\{\phi_p\}$ by giving their components *relative to each other*; indeed, it follows from Eqs. (5-1a) and (1-11b) that the component of δ_x relative to ϕ_p must be given by

$$(\phi_p,\delta_x) = (\delta_x,\phi_p)^* = r\,e^{+ixp/\hbar} \equiv r\,[\cos(xp/\hbar) + i\sin(xp/\hbar)]. \tag{5-1b}$$

Eqs. (5-1) make it rather plain that the eigenbases $\{\delta_x\}$ and $\{\phi_p\}$ do *not* coincide with each other; thus, we can expect X and P to be *incompatible* observables. We will explore that issue in more detail a little later.

In allowing the eigenvalue sets of **X** and **P** to be continuously distributed, we have evidently created a generalized vector space that has as many dimensions as there are real numbers; because, for each real number x there is a distinct basis vector δ_x, and similarly, for each real number p there is a distinct basis vector ϕ_p. This "heavy" kind of infinite-dimensionality gives rise to some thorny mathematical complications, the most bizarre of which is this: Although the position and momentum eigenbasis vectors are mutually orthogonal in the usual sense that

$$(\delta_x,\delta_{x'}) = 0 \text{ if } x \neq x' \quad \text{and} \quad (\phi_p,\phi_{p'}) = 0 \text{ if } p \neq p', \tag{5-2a}$$

it turns out that the "self-components" of those eigenbasis vectors, instead of being unity, are *infinite*:

$$(\delta_x,\delta_x) = \infty \quad \text{and} \quad (\phi_p,\phi_p) = \infty. \tag{5-2b}$$

But we hasten to add that *all other* unit vectors in the space are assumed to have self-components of one — i.e., $(\varepsilon,\varepsilon) = 1$.

Somewhat more reasonably, the expansion formulas for the state vector Ψ_t in the position and momentum eigenbases, instead of being *discrete* sums as in Eq. (3-3), are taken to be *continuous* sums — i.e., integrals:

$$\Psi_t = \sum_{j=1}^{\infty} a_j (a_j,\Psi_t) \quad \rightarrow \quad \Psi_t = \int_{-\infty}^{\infty} \delta_x (\delta_x,\Psi_t)\,dx, \quad \Psi_t = \int_{-\infty}^{\infty} \phi_p (\phi_p,\Psi_t)\,dp. \tag{5-3}$$

Similarly, the Pythagorean formula (4-3) takes the "continuous-sum" forms

$$\sum_{j=1}^{\infty} |(a_j,\Psi_t)|^2 = 1 \quad \rightarrow \quad \int_{-\infty}^{\infty} |(\delta_x,\Psi_t)|^2\,dx = 1, \quad \int_{-\infty}^{\infty} |(\phi_p,\Psi_t)|^2\,dp = 1. \tag{5-4}$$

And the change-of-basis formula (3-5) takes the continuous-sum forms

$$(\beta_k,\Psi_t) = \sum_{j=1}^{\infty} (\beta_k,a_j)(a_j,\Psi_t) \rightarrow (\phi_p,\Psi_t) = \int_{-\infty}^{\infty} (\phi_p,\delta_x)(\delta_x,\Psi_t)dx, \; (\delta_x,\Psi_t) = \int_{-\infty}^{\infty} (\delta_x,\phi_p)(\phi_p,\Psi_t)\,dp. \tag{5-5}$$

Apart from these mathematical technicalities, a noteworthy modification is made in our Rule 3, for predicting the result of a measurement: *For position and momentum measurements, Rule 3 is to be interpreted as follows*:

$|(\delta_x,\Psi_t)|^2\,dx$ = probability that an X-measurement, made on the particle in the state Ψ_t, will yield some value between x and $x+dx$, (5-6a)

$|(\phi_p,\Psi_t)|^2\,dp$ = probability that a P-measurement, made on the particle in the state Ψ_t, will yield some value between p and $p+dp$. (5-6b)

With these rules, Eqs. (5-4) tell us that the result of an X- or P-measurement on the particle in any state Ψ_t must be *some* real number between $-\infty$ and $+\infty$, just as we should expect.

Of all the foregoing statements, only parts (a) and (b) of our definition of the operators **X** and **P** contain really new information; all the other formulas are just restatements in "continuous" eigenvalue language of things we've said before in "discrete" eigenvalue language.

The δ_x-component of the state vector Ψ_t, namely (δ_x,Ψ_t), is evidently a complex number that depends on the two parameters x and t; hence, it is a *complex function* of x and t. As such, it is often written in the alternate functional form

$$(\delta_x,\Psi_t) \equiv \Psi_X(x,t), \tag{5-7a}$$

and called the *position wave function* of the system. If we know the position wave function $\Psi_X(x,t)$ for all values of its argument x, then we obviously know all the components of the state vector relative to a basis, namely the **X**-eigenbasis $\{\delta_x\}$; thus, knowing the position wave function is tantamount to knowing the "state" of the particle. Similarly, the ϕ_p-component of Ψ_t, (ϕ_p,Ψ_t), is a complex function of p and t that is often written

$$(\phi_p,\Psi_t) \equiv \Psi_P(p,t), \tag{5-7b}$$

and called the *momentum wave function* of the system. If we know the momentum wave function $\Psi_P(p,t)$ for all values of its argument p, then we obviously know all the components of the state vector relative to the **P**-eigenbasis $\{\phi_p\}$, so knowing the momentum wave function is *also* tantamount to knowing the "state" of the particle. Wave functions are a convenient and frequently used way of representing the state of a particle.

Since Eqs. (5-4) through (5-6) all involve the δ_x- and ϕ_p-components of the state vector, we can *rewrite* all of those equations in terms of the wave functions. Let's do that now. Beginning with Eq. (5-6), we see that the square moduli of the position and momentum wave functions have the special significance that

$|\Psi_X(x,t)|^2\,dx$ = probability that an X-measurement, made on the particle in the state Ψ_t, will yield some value between x and $x+dx$, (5-8a)

$|\Psi_P(p,t)|^2\,dp$ = probability that a P-measurement, made on the particle in the state Ψ_t, will yield some value between p and $p+dp$. (5-8b)

And Eqs. (5-4) tell us that these square moduli also satisfy

$$\int_{-\infty}^{\infty} |\Psi_X(x,t)|^2 \, dx = 1 \qquad \text{and} \qquad \int_{-\infty}^{\infty} |\Psi_P(p,t)|^2 \, dp = 1. \tag{5-9}$$

Fig. 5-1 shows schematic plots of the square moduli of the position and momentum wave functions. Eqs. (5-8) imply that the shaded areas in those figures are respectively equal to the probabilities for X- and P-measurements on the particle in the state Ψ_t to yield results in the indicated intervals. And Eqs. (5-9) imply that the total area under each curve is one.

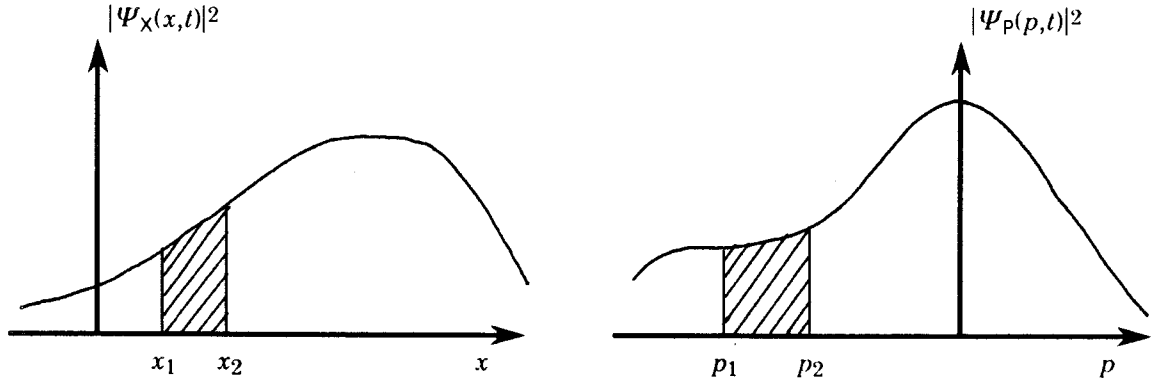


FIG. 5-1. Schematic plots of the *square moduli* of the position and momentum wave functions. The shaded areas equal the probabilities for X- and P-measurements on the particle in the state Ψ_t to give results in the indicated intervals. The total area under each curve is unity.

And finally, it follows from the change-of-basis formulas (5-5) and our fundamental Eqs. (5-1) that either of the two wave functions can be calculated from the other through the formulas

$$\Psi_P(p,t) = r \int_{-\infty}^{\infty} e^{ixp/\hbar} \Psi_X(x,t) \, dx, \tag{5-10a}$$

$$\Psi_X(x,t) = r \int_{-\infty}^{\infty} e^{-ixp/\hbar} \Psi_P(p,t) \, dp. \tag{5-10b}$$

●*Exercise 5-1.* Derive Eqs. (5-10).

Now let us deduce some of the interesting physical consequences of this rather heavy formalism.

▸ 5.2 The Wave-Particle Duality

In developing the consequences of our definition of the position and momentum operators, we don't want to confuse the particle that is our system with the particle *attribute* of being "localized at some point." So let's suppose that our system particle in an *electron* on the x-axis. Then the foregoing considerations enable us to make the following deductions:

Deduction 1. If the electron has a position x', then its position wave function $\Psi_X(x,t)$ is zero for all x except $x = x'$, where it has an infinite spike.

Proof: If the electron has a position x', then by definition its state vector Ψ_t coincides with the X-eigenbasis vector $\delta_{x'}$. The position wave function is therefore

$$\Psi_X(x,t) \equiv (\delta_x,\Psi_t) = (\delta_x,\delta_{x'}),$$

which by Eqs. (5-2) is zero if $x \neq x'$ and infinite if $x=x'$, precisely as claimed. Notice that this result is consistent with the probability interpretation of the square modulus of $\Psi_X(x,t)$ in (5-8a), since it implies that a position measurement on an electron with position x' cannot yield any value other than x'.

Deduction 2. If the electron has a momentum p', then its momentum wave function $\Psi_P(p,t)$ is zero for all p except $p=p'$, where it has an infinite spike.

•*Exercise 5-2.* Prove Deduction 2. [Hint: Repeat the proof of Deduction 1, except interchange the roles of position and momentum.]

Deduction 3. If the electron has a momentum p, then in some sense it also has a *position wavelength*

$$\lambda_X = 2\pi\hbar/p. \tag{5-11}$$

Proof: If the electron has a momentum p, then its state vector Ψ_t coincides with the **P**-eigenbasis vector ϕ_p. Its position wave function is then

$$\Psi_X(x,t) \equiv (\delta_x,\Psi_t) = (\delta_x,\phi_p) = r\,e^{-ixp/\hbar} = r[\cos(xp/\hbar) - i\sin(xp/\hbar)].$$

Now since

$$\cos[(x+2\pi\hbar/p)p/\hbar] = \cos[xp/\hbar + 2\pi] = \cos[xp/\hbar],$$

and similarly for the sine function, it follows that

$$\Psi_X(x+2\pi\hbar/p,t) = \Psi_X(x,t).$$

Thus, the position wave function of an electron with momentum p is periodic in x, with period or wavelength $2\pi\hbar/p$. Since this function completely defines the state of the electron, then the electron too must, in some obscure sense, "have a wavelength" $2\pi\hbar/p$.

Notice that Deduction 3 provides some glimmer of insight into the double-slit experiment discussed in Lecture 1, wherein an electron with momentum p_0 exhibited an interference pattern characteristic of a wave with wavelength $2\pi\hbar/p_0$. We shall give a fuller account of the double slit experiment in Lecture 7.

•*Exercise 5-3.* Show that if the electron has a position x, then in some sense it also has a *momentum wavelength* $\lambda_P = 2\pi\hbar/x$. [Hint: Repeat the proof of Deduction 3, except interchange the roles of position and momentum.]

Deduction 4. If the electron has a momentum p, then a position measurement will yield *any value with equal probability*; consequently, it makes absolutely no sense to ascribe a position value to an electron that "has a momentum."

Proof: If the electron has a momentum p, then its state vector Ψ_t coincides with the **P**-eigenbasis vector ϕ_p, so its position wave function is

$$\Psi_X(x,t) = (\delta_x,\Psi_t) = (\delta_x,\phi_p) = r\,e^{-ixp/\hbar}.$$

Then according to Eq. (5-8a), the probability that a position measurement will yield a result between x and $x+dx$ is

$$|\Psi_X(x,t)|^2\,dx = r^2|e^{-ixp/\hbar}|^2\,dx = r^2\,dx.$$

where we have invoked Eq. (1-11c). Since this probability is independent of x, then we must conclude that all x-values are equally likely.

Deduction 5. If the electron has a position x, then a momentum measurement will yield *any value with equal probability*; consequently, it makes absolutely no sense to ascribe a momentum value to an electron that "has a position."

•*Exercise 5-4.* Prove Deduction 5. [Hint: Modify the proof of Deduction 4.]

We had already inferred from our fundamental premise (5-1) that position and momentum were incompatible observables, and Deductions 4 and 5 evidently confirm that inference in the strongest

possible terms. Apparently, when we measure the position of an electron, we force the electron's state vector into one of the **X**-eigenbasis vectors δ_x; there the electron can be said to have a position value x but *not* a momentum value, and the electron's *spatial localization* causes us to think of it as a "particle." But if we then measure the momentum of that same electron, we force its state vector into one of the **P**-eigenbasis vectors ϕ_p; there the electron *cannot* be said to have a position value. But it *can* be said to have a momentum value p, and hence also a *spatial wavelength* $2\pi\hbar/p$, so we can think of the electron as a "wave." That, in essence, is how quantum mechanics explains, or at least accommodates, the wave-particle duality of Nature.

▸ 5.3 The Heisenberg Uncertainty Principle

We shall conclude our sojourn into quantum statics by saying something about the celebrated Heisenberg Uncertainty Principle, which you have no doubt heard of. The Heisenberg Uncertainty Principle is a purely mathematical consequence of Eqs. (5-10); however, its actual derivation would take more time than we have available here. We shall simply content ourselves with an explanation of what it says. But before we begin, here is a suggestion: Forget anything you *think* you already know about the Heisenberg Uncertainty Principle — just pretend you're hearing about it for the first time right now.

Suppose we have a particle on the x-axis in some state Ψ_t. As we have seen, the components of Ψ_t relative to the **X**-eigenbasis $\{\delta_x\}$ are given by the (complex) values of the position wave function $\Psi_X(x,t)$, while the components of Ψ_t relative to the **P**-eigenbasis $\{\phi_p\}$ are given by the (complex) values of the momentum wave function $\Psi_P(p,t)$. We have also seen that the *square moduli* of those complex functions have special import for predicting the results of position and momentum measurements. Specifically, as illustrated in Fig. 5-1, the area under the $|\Psi_X(x,t)|^2$-versus-x curve between x_1 and x_2 is numerically equal to the probability that a position measurement on the particle in state Ψ_t will yield a value between x_1 and x_2; similarly,the area under the $|\Psi_P(p,t)|^2$-versus-p curve between p_1 and p_2 is numerically equal to the probability that a momentum measurement on the particle in state Ψ_t will yield a value between p_1 and p_2. Now, it is possible to mathematically define for any $|\Psi_X(x,t)|^2$ curve a generic quantity Δ_X that measures the "width" or "spread" or "fatness" of that curve. The magnitude of Δ_X will therefore characterize the uncertainty we would have to contend with in trying to predict the result of a *position* measurement on the particle in state Ψ_t. And the same mathematical definition gives for any $|\Psi_P(p,t)|^2$ curve a quantity Δ_P that measures its width, and hence the amount of uncertainty we would have to contend with in trying to predict the result of a *momentum* measurement on the particle in state Ψ_t. We show in Fig. 5-2 the particle's *position uncertainty* Δ_X and *momentum uncertainty* Δ_P for a hypothetical state Ψ_t.

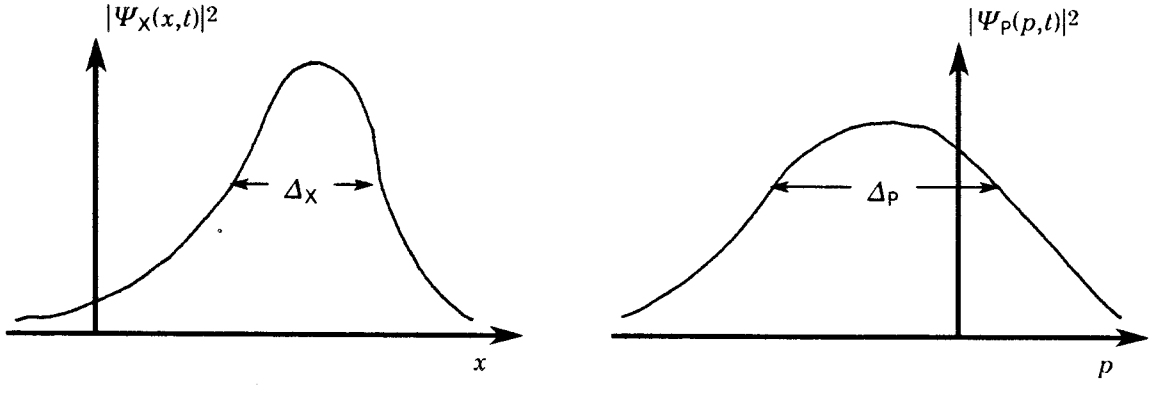


FIG. 5-2. Schematic plots of the square moduli of the position and momentum wave functions of a particle in some state Ψ_t. The Heisenberg Uncertainty Principle says that the products of the (suitably defined) widths Δ_X and Δ_P can never be smaller than $\hbar/2$.

Now, since the two wave functions $\Psi_X(x,t)$ and $\Psi_P(p,t)$ are related according to Eqs. (5-10), it should come as no surprise that the quantities Δ_X and Δ_P are related to each other. In fact, if we start with Eqs. (5-10) and go through a lengthy but purely mathematical argument, it is possible to prove the following inequality:

$$\Delta_X \cdot \Delta_P \geq \hbar/2. \tag{5-12}$$

This is the Heisenberg Uncertainty Principle, at least as it applies to the position and momentum of a particle. It implies that, for *any* state Ψ_t, there is a *fundamental limit* on the *simultaneous smallness* of the two uncertainties Δ_X and Δ_P: For a given position uncertainty Δ_X, then Δ_P can be no smaller than $\hbar/2\Delta_X$, a lower bound that approaches infinity as $\Delta_X \rightarrow 0$. Or, for a given momentum uncertainty Δ_P, then Δ_X can be no smaller than $\hbar/2\Delta_P$, a lower bound that approaches infinity as $\Delta_P \rightarrow 0$.

The most difficult thing to understand about the Heisenberg Uncertainty Principle is the precise physical meanings of the "uncertainties" Δ_X and Δ_P. Perhaps the best way to clarify those meanings is to imagine that we have 100 identical electrons, all in the same state Ψ_t. Suppose we make identical *position* measurements on 50 of those electrons. If the square modulus of the position wave function corresponding to state Ψ_t looks anything like that shown in Fig. 5-2, then those electrons obviously do *not* have a position value *prior* to the position measurement, and the results obtained in those 50 measurements will generally all be different. The quantity Δ_X characterizes the probable "scatter" in those 50 measurement results — i.e., the approximate amount by which any two of those measurement results will be found to differ from each other. Similarly, if on the other 50 electrons we make identical *momentum* measurements, then Δ_P will characterize the scatter in the resulting 50 momentum values. Notice that Δ_X is *not* a measure of our ignorance of the "true value" of X prior to an X-measurement; rather, it is a measure of the intrinsic uncertainty as to what X-value will be "developed" in an X-measurement. The "uncertainties" in the Heisenberg Uncertainty Principle obviously have a very technical definition that cannot be fully appreciated outside of the logic established by Rules 1 through 4. That's why the Heisenberg Uncertainty Principle, which is an important and often useful fact of modern science, is so often misrepresented outside of physics.

We have already seen the implications of the Heisenberg Uncertainty Principle in two limiting cases: If $\Psi_t = \delta_x$, so that the electron "has a position x," then our Deduction 1 implies that $\Delta_X = 0$ while Deduction 5 implies that $\Delta_P = \infty$, just as required by (5-12). And at the other extreme, if $\Psi_t = \phi_p$, so that the electron "has a momentum p," then our Deduction 2 implies that $\Delta_P = 0$ while Deduction 4 implies that $\Delta_X = \infty$, again in consonance with (5-12).

But if (5-12) is true, then how is it that we can *ever* use classical mechanics, which assumes that Δ_X and Δ_P are both always zero? The answer is that $\hbar$ is such a fantastically small number from a macroscopic point of view that Δ_X and Δ_P can appear to be *macroscopically* zero and *still* satisfy (5-12). It is only on the microscale that the Heisenberg Uncertainty Principle makes itself felt. The following exercise is intended to demonstrate this point.

●*Exercise 5-5.* Since momentum and velocity are related by $p = mv$, then $\Delta_P = m\Delta_V$, and (5-12) can be written in the form

$$\Delta_X \cdot \Delta_V \geq \hbar/2m.$$

(a) Show that a 1-gram particle can have its position certain to within 10^{-8} cm and its velocity certain to within 10^{-10} cm/sec, and yet still be a very long way from violating (5-12).

(b) Suppose an electron (mass $\approx 10^{-27}$ grams) is confined to an atom (diameter of 10^{-8} cm). What would be the *minimum* value for the velocity uncertainty Δ_V? Would it be at all meaningful to ascribe a "velocity value" to an electron that is inside an atom?

This concludes our discussion of quantum statics. In the next lecture we will take a look at quantum dynamics.

Lecture 6. QUANTUM MECHANICS: DYNAMICS I

▸ 6.1 The Hamiltonian Operator and Time-evolution

In Rules 1 and 2, we attached a time variable t to the state vector Ψ_t but not to the general observable operator **A**. The obvious implication is that the state vector changes with time, while all observable operators (along with their eigenbasis vectors and eigenvalues) are constant in time. In order to answer the FPOM for any $t>0$, it is necessary to specify precisely how Ψ_t evolves with time. This is the subject of Rule 5, and leads us into the "dynamics" part of quantum mechanics.

Rule 5 comes in two parts. The first part introduces an operator called the "Hamiltonian operator," and the second part introduces a closely related operator called the "time-evolution operator." We'll first state Rule 5 in full, and then elaborate its two parts separately.

▸ **Rule 5**:

(a) Every physical system has an observable H called the "total energy." The operator **H** that represents this observable is called the system's *Hamiltonian operator*, and we denote its eigenbasis by $\{\eta_n\}$ and its associated eigenvalue set by $\{E_n\}$:

$$\mathbf{H}\eta_n = E_n\eta_n. \quad (n = 0, 1, 2, ...) \tag{6-1}$$

(b) If the system is in some state Ψ_0 at time 0, then provided the system is not disturbed (such as by being measured), its state vector at time t will be

$$\Psi_t = \mathbf{U}_t\Psi_0. \quad (t\geq 0) \tag{6-2}$$

Here, $\mathbf{U}_t$, called the *time evolution operator*, is the linear operator with eigenbasis $\{\eta_n\}$ and associated eigenvalue set $\{e^{-iE_nt/\hbar}\}$,

$$\mathbf{U}_t\eta_n = e^{-iE_nt/\hbar}\,\eta_n \equiv [\cos(E_nt/\hbar) - i\sin(E_nt/\hbar)]\,\eta_n, \quad (n = 0, 1, 2, ...) \tag{6-3}$$

where $\{\eta_n\}$ and $\{E_n\}$ are as defined in part (a), and $\hbar$ is Planck's constant [see Eq. (1-2)].

Part (a) of Rule 5 simply asserts that every physical system has a *Hamiltonian* operator **H**, which is the operator that represents the system observable "total energy." But Rule 5 does *not* tell us how to *find* that operator **H**. As with any observable, it's up to the physicist to "discover" the Hamiltonian operator **H** for each physical system of interest. In discovering **H** the physicist defines, either directly or indirectly, its eigenbasis $\{\eta_n\}$ and eigenvalue set $\{E_n\}$; the former characterizes the generalized vector space of the system, while the latter characterizes the allowed energy values of the system. The task of finding **H** is obviously non-trivial, but not so impossible as it might seem. This is because various rules-of-thumb have been discovered over the years that give demonstrably correct Hamiltonians for many physical systems. Those rules-of-thumb have in fact become an important part of quantum mechanics as an "applied science." But we're not going to discuss those rules here; we're just going to describe what can be done once the Hamiltonian operator **H** is in hand.

Part (b) of Rule 5 says that once we know the system's Hamiltonian operator **H**, then we can proceed to determine the time-evolution of the state vector. We do that by first defining a *new* linear operator $\mathbf{U}_t$, called the *time-evolution* operator, as follows: $\mathbf{U}_t$ is to have the same eigenbasis $\{\eta_n\}$ as **H**, but the eigenvalue corresponding to η_n, instead of being E_n, is to be $e^{-iE_nt/\hbar}$, where $\hbar$ is Planck's constant. That this specification procedure completely defines the operator $\mathbf{U}_t$ follows from our discussion of Fig. 2-6. Notice that $\mathbf{U}_t$ *cannot* be regarded as an "observable operator," because its eigenvalue spectrum is not pure real. Indeed, the role of $\mathbf{U}_t$ in our theory is quite different from the role of **H** or any other observable operator: $\mathbf{U}_t$ *acts on* the time-0 state vector Ψ_0 and *transforms* it into the time-t state vector Ψ_t. Since Rule 1 requires the system's state vector to always be a *unit* vector, then we may expect that the action of $\mathbf{U}_t$ on Ψ_0 will be a "pure rotation" with no "stretching;" we'll verify shortly that that is indeed the case.

Now we're going to use Rule 5 to derive an explicit formula for the time-varying state vector. We proceed as follows:

$$\begin{aligned}
\Psi_t &= \mathbf{U}_t\Psi_0 && \text{[by (6-2)]}\\
&= \mathbf{U}_t[\Sigma_n \eta_n(\eta_n,\Psi_0)] && \text{[by expanding } \Psi_0 \text{ in the } \mathbf{H}\text{-eigenbasis]}\\
&= \Sigma_n \mathbf{U}_t\eta_n(\eta_n,\Psi_0) && \text{[since } \mathbf{U}_t \text{ is a linear operator]}\\
&= \Sigma_n e^{-iE_nt/\hbar}\eta_n(\eta_n,\Psi_0) && \text{[by (6-3)]}\\
\Psi_t &= \Sigma_n \eta_n[e^{-iE_nt/\hbar}(\eta_n,\Psi_0)]. && (6\text{-}4)
\end{aligned}$$

Eq. (6-4) shows how, in principle, we can calculate Ψ_t from a knowledge of Ψ_0, $\{\eta_n\}$ and $\{E_n\}$. Notice that the coefficient of the vector η_n in this formula is just the product of two complex numbers, namely $e^{-iE_nt/\hbar}$ and (η_n,Ψ_0), and hence is itself a complex number. In fact, since (6-4) is just an expansion of Ψ_t in the eigenbasis $\{\eta_n\}$, it follows that the coefficient of η_n in that formula is none other than the η_n-component of Ψ_t; i.e., Eq. (6-4) tells us that

$$(\eta_n,\Psi_t) = e^{-iE_nt/\hbar}(\eta_n,\Psi_0). \qquad (6\text{-}5)$$

So to get the η_n-component of Ψ_t, we merely multiply the η_n-component of Ψ_0 by the complex number $e^{-iE_nt/\hbar}$. Fig. 6-1 illustrates this important relationship pictorially.

•*Exercise 6-1.* Show that Eqs. (6-4) and (6-5) are identities for $t=0$. [Hint: Recall (1-11a).]

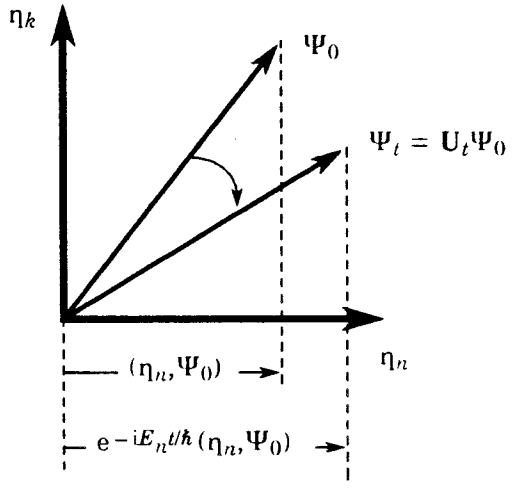


FIG. 6-1. Showing schematically how the state vector evolves with time from the perspective of the eigenbasis $\{\eta_n\}$ of the Hamiltonian operator $\mathbf{H}$.

Notice that the state vector of the system can *change* in *two different ways*: One way is the smooth, orderly evolutionary process described by Rule 5 that occurs when the system is left undisturbed. The other way is the sudden, random change described by Rule 4 that occurs when the system is measured. Much effort has gone into trying to explain the measurement change of Rule 4 *in terms of* the evolutionary change of Rule 5, but with no real success. Perhaps that's not surprising in view of the fact that the measurement change mandated by Rule 4 is *non-deterministic*, while the evolutionary change described by Rule 5 is quite deterministic. So we apparently must accord Rules 4 and 5 equal status in the theory. Thoughtful physicists still worry about whether these five rules are logically consistent; however, the fact remains that these rules seem to work very well together in the laboratory.

▸ 6.2 Conserved Quantities. Stationary States

We're now in a position to answer the FPOM for $t>0$. But first, let's savor the result (6-4) a bit by drawing out some of its other interesting consequences.

If we take the square modulus of Eq. (6-5), we get

$$|(\eta_n,\Psi_t)|^2 = |e^{-iE_n t/\hbar}(\eta_n,\Psi_0)|^2 = |e^{-iE_n t/\hbar}|^2\,|(\eta_n,\Psi_0)|^2 = |(\eta_n,\Psi_0)|^2, \tag{6-6}$$

where the two final steps have invoked the identities (1-8a) and (1-11c). Now this result has two interesting implications.

First, since Rule 3 tells us that $|(\eta_n,\Psi_t)|^2$ is the probability that an energy measurement on the system in the state Ψ_t will yield the **H**-eigenvalue E_n, then Eq. (6-6) implies that

$$\text{Prob}\{\text{Meas}(H)=E_n \text{ at time t}\} \text{ is independent of } t. \tag{6-7}$$

So *energy measurement probability is always conserved in time*, provided of course that the system is not measured or otherwise disturbed. Notice that this does *not* say that the *value* of the system's energy is always conserved in time — indeed, such a value will not exist unless Ψ_t happens to coincide with one of the **H**-eigenbasis vectors. All that (6-7) says is that the *probability* for an energy measurement to yield any particular result is conserved in time.

A second implication of (6-6) follows by summing both sides over all n:

$$\Sigma_n |(\eta_n,\Psi_t)|^2 = \Sigma_n |(\eta_n,\Psi_0)|^2 = 1,$$

where the last equality is due to the fact that Ψ_0 is a *unit* vector. In other words, the fact that Ψ_0 is a unit vector guarantees that $\Psi_t = \mathbf{U}_t\Psi_0$ will also be a unit vector, as is required by Rule 1. This confirms our earlier conjecture that the effect of the time-evolution operator $\mathbf{U}_t$ on Ψ_0 is a "pure rotation" in the generalized vector space, with no "stretching."

Another interesting consequence of Eq. (6-4) arises when the system's initial state Ψ_0 coincides with some **H**-eigenvector η_j, so that the system at time 0 has energy E_j. According to Eq. (6-4), the state vector of the system at time t will then be given by

$$\Psi_t = \Sigma_n \eta_n [e^{-iE_n t/\hbar}(\eta_n,\eta_j)].$$

Since (η_n,η_j) is zero for $n\neq j$ and one for $n=j$, then the sum here collapses, and we obtain the theorem

$$\Psi_0=\eta_j \;\Rightarrow\; \Psi_t = \eta_j e^{-iE_j t/\hbar} \quad \text{for all } t>0. \tag{6-8}$$

This theorem has two important implications: If $\Psi_0=\eta_j$, and the system remains undisturbed, then: (a) the system will have energy E_j for *all* $t>0$; and (b) the measurement probabilities for *all* system observables will be conserved. We'll leave the proofs of these two facts as an exercise.

●*Exercise 6-2.*

(a) Prove the first assertion above by using theorem (6-8) to show that

$$\text{Prob}\{\text{Meas}(H)=E_j \text{ at any time } t>0, \text{ given } \Psi_0=\eta_j\} = 1. \tag{6-9a}$$

[*Hint*: Start with Rule 3.]

(b) Prove the second assertion above by using theorem (6-8) to show that, for any observable A,

$$\text{Prob}\{\text{Meas}(A)=A_k \text{ at time } t, \text{ given } \Psi_0=\eta_j\} \text{ is independent of } t. \tag{6-9b}$$

[*Hint*: Start with Rule 3, and remember that $(\varepsilon,c\psi)=c(\varepsilon,\psi)$.]

The time-dependent vector $\eta_j e^{-iE_j t/\hbar}$ appearing on the right side of theorem (6-8) is called a *stationary state* of the system. Obviously, there is one stationary state for each eigenbasis vector of the system's Hamiltonian operator. What we have proved here is that if the system starts out in a stationary state, then the system will remain in that stationary state with a constant energy value and constant measurement probabilities for all its observables.

▸ 6.3 Answer to the FPOM for Any $t \geq 0$

Now let us see how quantum mechanics proposes to answer the Fundamental Problem of Mechanics for any $t \geq 0$. We reason as follows

- When A is measured on the system, the result must (by Rules 2 and 3) be some A-eigenvalue; let's call the measured eigenvalue A_j.
- So immediately after the A-measurement, the system will (by Rule 4) be in the state a_j.
- Starting from the state $\Psi_0 = a_j$ at time 0, the state vector will then evolve smoothly until time t, when it will be [by (6-4)] the vector

  $\Psi_t = \Sigma_n \eta_n [e^{-iE_n t/\hbar} (\eta_n, a_j)]$.
- Then measuring B at time t will (by Rule 3) yield any B-eigenvalue B_k with probability

  $|(\beta_k, \Psi_t)|^2 = |(\beta_k, \Sigma_n \eta_n [e^{-iE_n t/\hbar} (\eta_n, a_j)])|^2 = |\Sigma_n (\beta_k, \eta_n)[e^{-iE_n t/\hbar} (\eta_n, a_j)]|^2$.

▸ Therefore, quantum theory's answer to the FPOM in the general t case is:

$$\text{Prob}\{\mathsf{B} = B_k \text{ at time } t, \text{ given } \mathsf{A} = A_j \text{ at time } 0\} = |\Sigma_n (\beta_k, \eta_n)\, e^{-iE_n t/\hbar} (\eta_n, a_j)|^2. \quad (k = 1, 2, \ldots) \qquad (6\text{-}10)$$

Notice that all we need to know to answer the FPOM for any $t \geq 0$ are:
- the eigenvalues of **A**, **B**,and **H**, and
- the components of the **A** and **B** eigenvectors relative to the **H**-eigenbasis.

Let's verify that the general result (6-10) indeed reduces to the static result (3-8) when $t \approx 0$. Since $e^{i0} = 1$ by (1-11a), then the right side of (6-10) becomes

$$|\Sigma_n (\beta_k, \eta_n)\, e^{i0} (\eta_n, a_j)|^2 = |\Sigma_n (\beta_k, \eta_n) (\eta_n, a_j)|^2 = |(\beta_k, a_j)|^2,$$

where in the last step we have invoked the component expansion theorem (2-3). This is precisely the result (3-8) for the $t \approx 0$ case.

●*Exercise 6-3.* In the FPOM, suppose the second-measured observable B is the same as the first-measured observable A. Prove that

$$\text{Prob}\{\mathsf{A} = A_j \text{ at time } t, \text{ given } \mathsf{A} = A_j \text{ at time } 0\} = |\Sigma_n |(\eta_n, a_j)|^2\, e^{-iE_n t/\hbar}|^2. \qquad (6\text{-}11)$$

Show that this probability is unity if $t = 0$, as required by Rules 3 and 4, but need not be unity for any $t > 0$.

●*Exercise 6-4.* Use (6-10) to show that an energy measurement at time 0 followed by an energy measurement at any time $t > 0$ will always give equal results, in agreement with theorem (6-9a).

Rules 1 through 5 form the logical foundations of the "standard" theory of quantum mechanics, and (6-10) is the answer that those rules give to the Fundamental Problem of Mechanics. All the rest of quantum mechanics consists of: generalizing those five rules somewhat to accommodate more complicated situations; finding specific operators to represent specific system observables, especially specific Hamiltonian operators for specific systems; and, of course, evaluating (6-10) for specific problems.

We shall conclude this lecture by sketching a philosophically interesting shift of perspective on quantum theory, along lines originally suggested by Richard Feynman.

▸6.4 Transition Amplitudes and Virtual Processes

It is possible to look upon our fundamental result (6-10) as giving the probability that the two consecutive processes "t-passage" and "B-measurement" will *jointly* carry the system *from* state α_j at time zero *to* state β_k at time t. In this view, (6-10) gives the probability that the system will "make a *transition*" from state α_j to state β_k in time t, with the tacit understanding that B is actually measured at time t. In the spirit of this view, here is an *alternate way* of arriving at the result (6-10):

▸ *Rule A.* Imagine that the α_j-to-β_k transition occurs, in some vaguely defined way, *via* the energy eigenstates $\{\eta_n\}$, and assign to the particular *virtual process* $\{\alpha_j \to \eta_n \to \beta_k$ in time $t\}$ the *transition amplitude,*

$$\Pi_n(\alpha_j \to \eta_n \to \beta_k \text{ in } t) \equiv (\beta_k, \eta_n)\, e^{-iE_n t/\hbar}\, (\eta_n, \alpha_j). \tag{6-12a}$$

[The quantity on the right side is most suggestively read from right-to-left. Notice that this transition amplitude is simply the product of three complex numbers, and hence is itself just a complex number; indeed, it is just the summand in (6-10).]

▸ *Rule B.* Similarly, associate with the *net* process $\{\alpha_j \to \beta_k$ in time $t\}$ a transition amplitude $\Pi(\alpha_j \to \beta_k$ in $t)$, and postulate that this net transition amplitude is equal to the *sum* of all the virtual process amplitudes:

$$\Pi(\alpha_j \to \beta_k \text{ in } t) = \Sigma_n\, \Pi_n(\alpha_j \to \eta_n \to \beta_k \text{ in } t). \tag{6-12b}$$

[This net amplitude, being a sum of complex numbers, is then itself a complex number.]

▸ *Rule C.* Finally, postulate that the *probability* of the net process $\{\alpha_j \to \beta_k$ in time $t\}$ is equal to the *square modulus* of its amplitude:

$$\text{Prob}(\alpha_j \to \beta_k \text{ in } t) = |\Pi(\alpha_j \to \beta_k \text{ in } t)|^2. \tag{6-12c}$$

It is easy to see that if (6-12a) is substituted into (6-12b), and then that is substituted into (6-12c), we get precisely the result (6-10). The foregoing describes what might be called a "virtual process" formulation of quantum theory. It's relation to the "conventional" formulation of quantum theory, as embodied by Rules 1 through 5, can be briefly summarized this way: The virtual process formulation *retains* Rules 1, 2 and 5(a), but *replaces* Rules 3, 4 and 5(b) with Rules A, B and C.

We know that, when considering mutually exclusive *real* processes, we must sum their *probabilities* to obtain the total process probability. But for the mutually exclusive *virtual* processes contemplated above, the rules are evidently different: We must sum their *amplitudes*, and then obtain the total process probability as the square modulus of that amplitude sum.

But, you might ask, how does Nature *really* work? Does the system go from state α_j to state β_k by way of one, or all, of the states $\{\eta_n\}$? Or does the system's state vector smoothly rotate from state α_j to state Ψ_t and then suddenly jump to state β_k when the B-measurement is made? In truth, no one really knows. Moreover, no one has been able to paint a detailed, consistent, plausible picture of *either* the mysterious virtual transition process $\alpha_j \to \eta_n \to \beta_k$, *or* the equally mysterious measurement jump process $\Psi_t \to \beta_k$.

So what are we to make of this state of affairs? The answer seems to be this: *All that present day quantum theory can do is provide a computational algorithm — namely (6-10) — for quantitatively answering the Fundamental Problem of Mechanics.* That algorithm has been amazingly successful in laboratory applications, not only in describing phenomena that were already known, but also in *predicting* phenomena that were subsequently discovered. But no one has been able to undergird that algorithm with a common-sense "picture of reality" analogous to the one suggested by classical mechanics for macroscale phenomena. The suspicion among some is that microscopic reality may not be "understandable" by minds whose criteria for understanding have been conditioned so thoroughly by macroscale experiences. Serious contemplation of this prospect leaves most physicists in some linear combination state of dismay, skepticism, awe and fascination.

Lecture 7. QUANTUM MECHANICS: DYNAMICS II

▸ 7.1 Recapitulation

In our last lecture we introduced Rule 5, our final rule of quantum mechanics. Rule 5 comes in two parts. Part (a) asserts that every physical system has a "total energy" — an observable that we represent in the system's generalized vector space by an operator **H** called the system's *Hamiltonian*. We denote the eigenbasis of **H** by $\{\eta_n\}$ and the corresponding eigenvalue set by $\{E_n\}$:

$$\mathbf{H}\eta_n = E_n\eta_n. \quad (n=1,2,\dots) \tag{7-1}$$

Rule 5 does *not* tell us how to *find* the Hamiltonian operator for any given physical system; the "discovery" of **H** is up to us. In discovering **H**, we also discover, either directly or indirectly, its eigenbasis $\{\eta_n\}$ and eigenvalue set $\{E_n\}$. The former effectively defines the system's generalized vector space, while the latter determines the system's allowed energy values. What makes the Hamiltonian operator more important than any other observable operator of the system is the fact that **H** determines the time evolution of the system's state vector. Specifically, part (b) of Rule 5 says that, given that the system is in state Ψ_0 at time 0, then if we leave the system alone until time t its state vector will be

$$\Psi_t = \mathbf{U}_t\Psi_0. \quad (t\geq 0) \tag{7-2}$$

Here, $\mathbf{U}_t$, called the *time-evolution* operator, is the linear operator with eigenbasis $\{\eta_n\}$ and eigenvalue set $\{e^{-iE_nt/\hbar}\}$:

$$\mathbf{U}_t\eta_n = e^{-iE_nt/\hbar}\eta_n. \quad (n=1,2,\dots) \tag{7-3}$$

So specifying **H** determines $\mathbf{U}_t$, and hence the time-evolution of the system's state vector.

If in Eq. (7-2) we expand Ψ_0 in the **H**-eigenbasis, and then use the fact that $\mathbf{U}_t$ is linear and statisfies Eqs. (7-3), we obtain

$$\Psi_t = \mathbf{U}_t\Sigma_n\,\eta_n(\eta_n,\Psi_0) = \Sigma_n\,\mathbf{U}_t\eta_n(\eta_n,\Psi_0) = \Sigma_n\,e^{-iE_nt/\hbar}\eta_n(\eta_n,\Psi_0),$$

or

$$\Psi_t = \Sigma_n\,\eta_n[e^{-iE_nt/\hbar}(\eta_n,\Psi_0)]. \tag{7-4}$$

This is evidently an explicit formula for the time-evolving state vector in terms of the Hamiltonian's eigenbasis and eigenvalue set.

With the result (7-4) it is straightforward to use Rules 1 through 4 to formulate a general answer to the Fundamental Problem of Mechanics. In the FPOM, we measure some observable A at time 0, note the result, and then let the system evolve to some time t when we measure some observable B. To predict the result of the B-measurement, we reason as follows: When we measure A at time zero, we will obtain some **A**-eigenvalue A_j, thereby leaving the system in the corresponding **A**-eigenstate α_j. The state vector then evolves to time t according to Eq. (7-4) with $\Psi_0=\alpha_j$. At time t, a B-measurement will yield any **B**-eigenvalue B_k with probability

$$\text{Prob}\{\mathsf{B}=B_k \text{ at time t, given } \mathsf{A}=A_j \text{ at time } 0\} = |(\beta_k,\Psi_t)|^2, \quad (k=1,2,\dots) \tag{7-5}$$

where β_k is the **B**-eigenbasis vector corresponding to B_k. Substituting Eq. (7-4), with $\Psi_0=\alpha_j$, into Eq. (7-5), we obtain the key result of our theory:

$$\text{Prob}\{\mathsf{B}=B_k \text{ at time t, given } \mathsf{A}=A_j \text{ at time } 0\} = |\Sigma_n\,(\beta_k,\eta_n)\,e^{-iE_nt/\hbar}(\eta_n,\alpha_j)|^2. \quad (k=1,2,\dots) \tag{7-6}$$

▸ 7.2 (Optional) The Schrödinger Equations

It may happen that Eq. (7-6) is computationally inconvenient. This would be the case if the system's Hamiltonian operator **H** is specified in some way *other* than by calling out its eigenbasis and eigenvalues, and those then turn out to be analytically so complicated that the right side of Eq. (7-6) becomes difficult to evaluate. In such cases, it is useful to have an alternate way of finding the time-varying state vector from the system's Hamiltonian operator. To that end, consider the *time-*

derivative of the state vector Ψ_t, which we define by the usual rule

$$\frac{d\Psi_t}{dt} \equiv \lim_{\Delta t \to 0} \frac{\Psi_{t+\Delta t} - \Psi_t}{\Delta t} \equiv \lim_{\Delta t \to 0} \left((\Delta t)^{-1} \Psi_{t+\Delta t} + (-\Delta t)^{-1} \Psi_t \right).$$

(The last expression shows that the quantity we are taking the limit of is simply a linear combination of two vectors in our generalized vector space.) Let's see what we get by evaluating this derivative using the formula (7-4). We have

$$\begin{aligned}
d\Psi_t/dt &= (d/dt)\left(\Sigma_n \eta_n [e^{-iE_n t/\hbar} (\eta_n, \Psi_0)]\right) && \text{[by Eq. (7-4)]} \\
&= \Sigma_n \eta_n (d/dt)[e^{-iE_n t/\hbar}] (\eta_n, \Psi_0) && \text{[differentiation is a linear operation]} \\
&= \Sigma_n \eta_n (-iE_n/\hbar)[e^{-iE_n t/\hbar}] (\eta_n, \Psi_0) && \text{[by Eq. (1-10)]} \\
&= (1/i\hbar) \Sigma_n E_n \eta_n e^{-iE_n t/\hbar} (\eta_n, \Psi_0) && \text{[since } -i = 1/i\text{]} \\
&= (1/i\hbar) \Sigma_n \mathbf{H} \eta_n e^{-iE_n t/\hbar} (\eta_n, \Psi_0) && \text{[by Eq. (7-1)]} \\
&= (1/i\hbar)\mathbf{H} \Sigma_n \eta_n [e^{-iE_n t/\hbar} (\eta_n, \Psi_0)] && \text{[H is a linear operator]} \\
&= (1/i\hbar)\mathbf{H}\Psi_t. && \text{[by Eq. (7-4)]}
\end{aligned}$$

Thus we have derived the differential equation

$$i\hbar \, d\Psi_t/dt = \mathbf{H}\Psi_t. \tag{7-7}$$

which is to be solved for Ψ_t subject to the initial condition $\Psi_{t=0} = \Psi_0$. This prescription for finding the time-evolving state vector is of course fully equivalent to Eqs. (7-2) and (7-4).

Eq. (7-7) is called *the time-dependent Schrödinger equation*, and Eq. (7-1) is called *the time-independent Schrödinger equation*. It is important that the similarity in the names of those two equations not obscure the fundamental difference between their roles: The time-independent Schrödinger equation (7-1) is the eigenvalue-eigenvector equation for the total energy operator **H**. As such, it defines the energy eigenbasis $\{\eta_n\}$ and eigenvalue set $\{E_n\}$ in those circumstances when **H** has been specified in some other way. The time-dependent Schrödinger equation (7-7), on the other hand, is the fundamental time-evolution equation for the system's state vector. It's purpose is to determine Ψ_t from Ψ_0 in those circumstances in which the computations required by Eq. (7-4) happen to be inconvenient to carry out. Notice that if we *do* find Ψ_t by solving Eq. (7-7) instead of by evaluating formula (7-4), then to compute an answer to the FPOM we would want impose the initial condition $\Psi_{t=0} = a_j$ and use Eq. (7-5) instead of Eq. (7-6).

In practical applications, the Schrödinger equations are nearly always expressed in "component form" relative to the eigenbasis of some convenient observable operator. To derive the A-component form of the time-independent Schrödinger equation (7-1), we first observe that

$$\mathbf{H}\eta_n = \mathbf{H} \Sigma_j a_j (a_j, \eta_n) = \Sigma_j \mathbf{H} a_j (a_j, \eta_n);$$

thus, Eq. (7-1) can be written

$$\Sigma_j \mathbf{H} a_j (a_j, \eta_n) = E_n \eta_n.$$

Now taking the a_k-component of this equation and invoking the linearity property, we get

$$\Sigma_j (a_k, \mathbf{H} a_j) (a_j, \eta_n) = E_n (a_k, \eta_n). \qquad (k = 1, 2, \ldots ; n = 1, 2, \ldots) \tag{7-1'}$$

This is the "A-component form" of the time-independent Schrödinger equation (7-1). In it, the set of complex numbers $\{(a_k, \eta_n)$ for $k = 1, 2, \ldots\}$ "represents" the vector η_n relative to the **A**-eigenbasis, and the set of complex numbers $\{(a_k, \mathbf{H}a_j)$ for $j = 1, 2, \ldots$ and $k = 1, 2, \ldots\}$ "represents" the operator **H** relative to the **A**-eigenbasis. If **H** is defined by *specifying* the set of numbers $\{(a_k, \mathbf{H}a_j)\}$, then we can solve Eq. (7-1') for the set of numbers $\{(a_k, \eta_n)\}$ and the number E_n — i.e., for η_n and its associated eigenvalue. Similar manipulations applied to the time-dependent Schrödinger equation (7-7) will show that its A-component form is *(optional exercise!)*

$$i\hbar \, d(a_k, \Psi_t)/dt = \Sigma_j (a_k, \mathbf{H} a_j) (a_j, \Psi_t). \qquad (k = 1, 2, \ldots) \tag{7-7'}$$

This is a set of coupled, first order differential equations for the **A**-components $\{(a_k, \Psi_t)\}$ of the state vector Ψ_t, and these equations can in principle be solved if we know the set of numbers $\{(a_k, \mathbf{H}a_j)\}$.

▸ 7.3 The Free Particle

Perhaps the simplest of all dynamical systems is the *free particle*, a particle of mass m on which no forces act. You know the solution to this problem offered by *classical* mechanics: The particle moves with constant velocity, say along the x-axis. More precisely, if the particle's initial position and momentum are x_0 and p_0, so that its initial velocity is p_0/m, then its position and momentum at any later time t will be

$$x(t) = x_0 + p_0 t/m, \quad p(t) = p_0. \tag{7-8}$$

To see how the free particle is analyzed in quantum mechanics, we first recall from Lecture 5 that in quantum mechanics we represent the observables "position" and "momentum" for any particle, free or otherwise, by two linear operators $\mathbf{X}$ and $\mathbf{P}$, which are defined as follows: The eigenvalue sets of both $\mathbf{X}$ and $\mathbf{P}$ are all the real numbers, and their eigenbases $\{\delta_x\}$ and $\{\phi_p\}$, which satisfy

$$\mathbf{X}\delta_x = x\delta_x \ \ (-\infty<x<\infty), \qquad \mathbf{P}\phi_p = p\phi_p \ \ (-\infty<p<\infty), \tag{7-9}$$

are defined by the prescription

$$(\delta_x,\phi_p) = r\,e^{-ixp/\hbar}. \ \ (-\infty<x,p<\infty) \tag{7-10}$$

But what about the observable "total energy"? What shall we take to be the Hamiltonian operator $\mathbf{H}$ for a free particle? You will recall that in *classical* mechanics the total energy of a free particle of momentum p is given by the formula $\frac{1}{2}m(p/m)^2 = p^2/2m$. Well, the physicist's "Handbook of Tried-and-True Hamiltonian Recipes" declares that the Hamiltonian operator for the free particle is

$$\mathbf{H} = (1/2m)\mathbf{P}^2. \tag{7-11}$$

Fair enough, but what do we mean by the "square" of the operator $\mathbf{P}$? We mean simply this: If ψ is any vector, then to find the vector $\mathbf{P}^2\psi$ we first operate on ψ with $\mathbf{P}$ to get the vector $\mathbf{P}\psi$, and we then operate on *that* vector with $\mathbf{P}$. So for the operator $\mathbf{H}$ in Eq. (7-11), the vector $\mathbf{H}\psi$ is

$$\mathbf{H}\psi = (1/2m)\mathbf{P}^2\psi \equiv (1/2m)(\mathbf{P}(\mathbf{P}\psi)).$$

With this definition, we see that the free-particle Hamiltonian is indeed a well-defined operator; furthermore, this definition implies the following important result:

- For the free particle, the eigenbasis of $\mathbf{H}$ is the momentum eigenbasis $\{\phi_p\}$, and the eigenvalue of $\mathbf{H}$ corresponding to the eigenbasis vector ϕ_p is $p^2/2m$.

The proof of this assertion goes as follows: For any momentum eigenbasis vector ϕ_p we have

$$(1/2m)\mathbf{P}^2\phi_p = (1/2m)(\mathbf{P}(\mathbf{P}\phi_p)) = (1/2m)(\mathbf{P}(p\phi_p)) = (p/2m)(\mathbf{P}\phi_p) = (p/2m)(p\phi_p).$$

Thus, *for the free particle* we indeed have the eigenvalue-eigenvector relation

$$\mathbf{H}\phi_p = (p^2/2m)\phi_p. \ \ (-\infty<p<\infty) \tag{7-12}$$

• *Exercise 7-1.* Of the three free-particle observables "position," "momentum" and "energy," which pairs are *compatible* and which are *incompatible*?

Since for the free particle the vector ϕ_p is an eigenbasis vector of $\mathbf{H}$ with eigenvalue $p^2/2m$, then it follows from Eq. (6-8) that

$$\Psi_0 = \phi_p \quad \Rightarrow \quad \Psi_t = \phi_p\, e^{-i(p^2/2m)t/\hbar}, \ \ (t>0) \tag{7-13}$$

the latter being a "stationary state" of the system. Therefore, *if* the free particle is in state ϕ_p at time 0, then at any later time $t>0$ an energy measurement on the particle would yield the certain value $p^2/2m$, and a momentum measurement would yield the certain value p; however, a position measurement would yield *any* real number with *equal* probability.

• *Exercise 7-2.* Prove the three assertions in the last sentence by applying Rule 3 to the state vector Ψ_t in (7-13).

So we see that if we initially measure the momentum of a free particle, then the particle's momentum value *and* energy value will be sharply fixed thereafter; however, the particle *cannot* thereafter be said to have a position value.

To calculate Ψ_t for initial states Ψ_0 other than one of the vectors $\{\phi_p\}$, we must resort to formula (7-4). Making the appropriate substitutions there, namely

$$\eta_n \to \phi_p, \quad E_n \to p^2/2m, \quad \Sigma_n \to \int dp,$$

we see that the general formula for the time-varying state vector of the free particle is

$$\Psi_t = \int_{-\infty}^{\infty} \phi_p [e^{-ip^2 t/2m\hbar} (\phi_p, \Psi_0)] \, dp. \tag{7-14}$$

Since we'll be especially interested in *position* measurements at time t, let's take the δ_x-component of this vector equation. We get

$$(\delta_x, \Psi_t) = \left(\delta_x, \int_{-\infty}^{\infty} \phi_p [e^{-ip^2 t/2m\hbar} (\phi_p, \Psi_0)] \, dp \right)$$

$$= \int_{-\infty}^{\infty} (\delta_x, \phi_p) e^{-ip^2 t/2m\hbar} (\phi_p, \Psi_0) \, dp .$$

Using Eq. (7-10), and recalling that (δ_x, Ψ_t) is just the position wave function $\Psi_X(x,t)$, this is

$$\Psi_X(x,t) = r \int_{-\infty}^{\infty} e^{-ixp/\hbar} e^{-ip^2 t/2m\hbar} (\phi_p, \Psi_0) \, dp . \tag{7-15}$$

This **X**-basis version of Eq. (7-14) is useful for predicting the results of position measurements on the free particle because, as you will recall from Lecture 5 [see Eq. (5-8a)], $|\Psi_X(x,t)|^2 dx$ gives the probability that a position measurement at time t will give a result between x and $x + dx$.

Let's evaluate Eq. (7-15) for the particular initial state $\Psi_0 = \delta_a$, which corresponds to the free particle having position value a at time 0. Eq. (7-15) says that

$$\Psi_0 = \delta_a \quad \Rightarrow \quad \Psi_X(x,t) = G(x,t,m,a), \tag{7-16}$$

where for reasons that will become clear shortly we have introduced the function

$$G(x,t,m,a) \equiv r \int_{-\infty}^{\infty} e^{-ixp/\hbar} e^{-ip^2 t/2m\hbar} (\phi_p, \delta_a) \, dp . \tag{7-17}$$

Surprisingly, we can evaluate $G(x,t,m,a)$ up to an unimportant multiplicative constant *without* actually performing any integration! Here's how: Remembering that

$$(\phi_p, \delta_a) = (\delta_a, \phi_p)^* = r e^{+iap/\hbar},$$

and also that $e^{iu} e^{iv} \equiv e^{i(u+v)}$, we get from Eq. (7-17),

$$G(x,t,m,a) = rr \int_{-\infty}^{\infty} \exp\left\{ -\frac{ixp}{\hbar} - \frac{ip^2 t}{2m\hbar} + \frac{iap}{\hbar} \right\} dp$$

$$= r^2 \int_{-\infty}^{\infty} \exp\left\{ -\frac{it}{2m\hbar} \left(p^2 + \frac{2pm(x-a)}{t} \right) \right\} dp$$

$$= r^2 \int_{-\infty}^{\infty} \exp\left\{ -\frac{it}{2m\hbar} \left(p + \frac{m(x-a)}{t} \right)^2 + \frac{it}{2m\hbar} \left(\frac{m(x-a)}{t} \right)^2 \right\} dp$$

$$= r^2 \exp\left\{ \frac{im(x-a)^2}{2\hbar t} \right\} \int_{-\infty}^{\infty} \exp\left\{ -\frac{it}{2m\hbar} \left(p + \frac{m(x-a)}{t} \right)^2 \right\} dp .$$

Changing the integration variable from p to

$$u = \left(\frac{t}{2m\hbar}\right)^{1/2}\left(p + \frac{m(x-a)}{t}\right), \qquad du = \left(\frac{t}{2m\hbar}\right)^{1/2} dp,$$

this becomes

$$G(x,t,m,a) = r^2 \exp\left\{\frac{im(x-a)^2}{2\hbar t}\right\}\left(\frac{2m\hbar}{t}\right)^{1/2}\int_{-\infty}^{\infty} e^{-iu^2}\,du.$$

The u-integral here is evidently just some complex constant; thus we conclude that

$$G(x,t,m,a) = c\left(\frac{m}{t}\right)^{1/2} \exp\left\{\frac{im(x-a)^2}{2\hbar t}\right\}, \tag{7-18}$$

where c is a complex constant whose precise value will not concern us here.

If we now substitute the result (7-18) into (7-16), and then take the square modulus, we conclude:

$$\Psi_0 = \delta_a \quad \Rightarrow \quad |\Psi_X(x,t)|^2 = |c|^2\,(m/t). \tag{7-19}$$

• *Exercise 7-3.* Show in detail how Eq. (7-19) follows from Eqs. (7-16) and (7-18).

The fact that the right side of Eq. (7-19) is independent of both x and a is the key result here. It means that if the free particle has the precise position a at time 0, then a position measurement at any finitely later time will yield *any* x-value with *equal* probability. So a free particle, if localized to one point, will immediately and completely "de-localize" itself!

Holding this surprising and peculiar result in abeyance for the moment, let's now consider what would happen if the intial state were $\Psi_0 = \delta_a + \delta_{-a}$. That initial state corresponds to a situation in which a time-zero measurement of position would produce, with equal probability, either $x = a$ or $x = -a$. Substituting this initial state into Eq. (7-15) gives

$$\Psi_X(x,t) = r\int_{-\infty}^{\infty} e^{-ixp/\hbar} e^{-ip^2 t/2m\hbar}(\phi_p, \delta_a + \delta_{-a})\,dp$$

$$= G(x,t,m,a) + G(x,t,m,-a),$$

where we first expanded the component in the usual way and then invoked our definition (7-17) of G. Taking the square modulus of this function gives, using the identity (1-8b),

$$|\Psi_X(x,t)|^2 = |\,G(x,t,m,a) + G(x,t,m,-a)\,|^2$$
$$= |G(x,t,m,a)|^2 + |G(x,t,m,-a)|^2 + 2\mathrm{Re}\{G(x,t,m,a)\,G^*(x,t,m,-a)\}.$$

Substituting our result (7-18) for G into this expression and simplifying algebraically, we conclude:

$$\Psi_0 = \delta_a + \delta_{-a} \quad \Rightarrow \quad |\Psi_X(x,t)|^2 = |c|^2\,(m/t)\,2[1 + \cos(2max/\hbar t)]. \tag{7-20}$$

•*Exercise 7-4.*

(a) Show in detail how Eq. (7-20) follows from the preceding equation and Eq. (7-18). [*Hint*: Besides the identities developed at the end of Lecture 1, you'll need to recall the trigonometric identity for $\cos(u-v)$.]

(b) Show that the x-distance between an adjacent maximum and minimum of the function in Eq. (7-20) is

$$\Delta(t) = n\hbar t/2ma. \tag{7-21}$$

Fig. 7-1 shows plots of the free-particle position probability density function $|\Psi_X(x,t)|^2$ for the two initial states $\Psi_0 = \delta_a$ and $\Psi_0 = \delta_a + \delta_{-a}$, as given in Eqs. (7-19) and (7-20). These quantum predictions for the free particle are quite bizarre compared to the classical prediction (7-8). But we

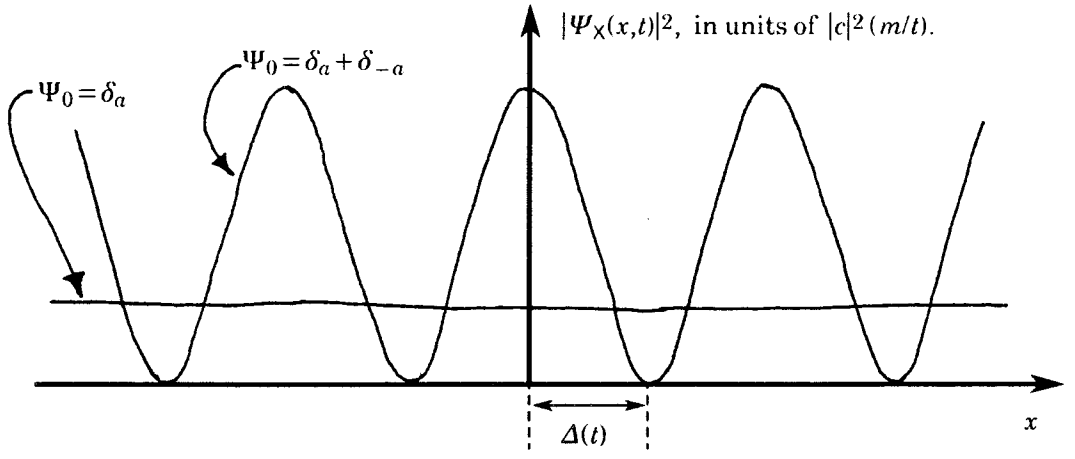


FIG. 7-1. Plot of $|\Psi_X(x,t)|^2$ for the free particle, given the two initial states $\Psi_0=\delta_a$ and $\Psi_0=\delta_a+\delta_{-a}$. The distance $\Delta(t)$ is as given in Eq. (7-21).

are now going to show that these seemingly absurd quantum predictions successfully explain the results of the double-slit experiment that we described back in Lecture 1.

▸ 7.4 Explanation of the Double-Slit Experiment

The set-up of the double-slit experiment is sketched in Fig. 7-2. An electron of mass m and horizontal momentum p_0 is incident on a vertical screen S_1, which is opaque except for two narrow slits. We measure vertical distance on screen S_1 by the variable y, and we'll take slit 1 to be at $y=a$ and slit 2 to be at $y=-a$. A distance L beyond screen S_1 is a second vertical screen S_2, this one coated with a phosphor that enables us to record the point of impact of any electron. We'll measure vertical distance on screen S_2 by the variable z, with $z=0$ on the same level as $y=0$.

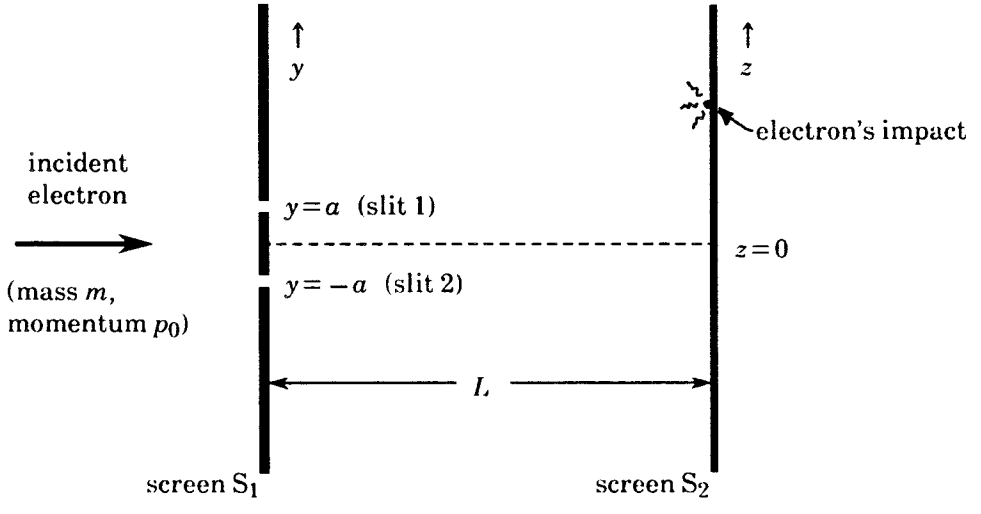


FIG. 7-2. Geometry of the double-slit experiment. (See Fig. 1-2.)

The key to understanding the double-slit experiment in terms of the dynamics of a one-dimensional free particle is this: In quantum mechanics, just as in classical mechanics, *the horizontal*

and vertical components of the motion of a free particle are independent of each other. Thus, any electron that makes it through screen S_1 retains its horizontal momentum component p_0 until it strikes screen S_2. But the vertical momentum component is drastically effected by screen S_1; in fact, any electron passing through screen S_1 with only slit 1 open is forced into the vertical position state δ_a, while any electron passing through screen S_1 with both slits open is forced into the vertical position state $\delta_a + \delta_{-a}$. Since the horizontal velocity in either case is p_0/m, then the electron will in either case reach screen S_2 at a time

$$t = L/(p_0/m) = mL/p_0 \tag{7-22}$$

after it leaves screen S_1. Therefore, screen S_2 is essentially measuring, at time $t = mL/p_0$, the vertical position of an electron that was prepared at time $t = 0$ in the vertical position state δ_a if only slit 1 was open, or in the vertical position state $\delta_a + \delta_{-a}$ if both slits were open. So in terms of our quantum model of a free particle confined to the x-axis, the y-axis on screen S_1 in Fig. 7-2 corresponds to the x-axis at time $t = 0$, while the z-axis on screen S_2 corresponds to the x-axis at time $t = mL/p_0$.

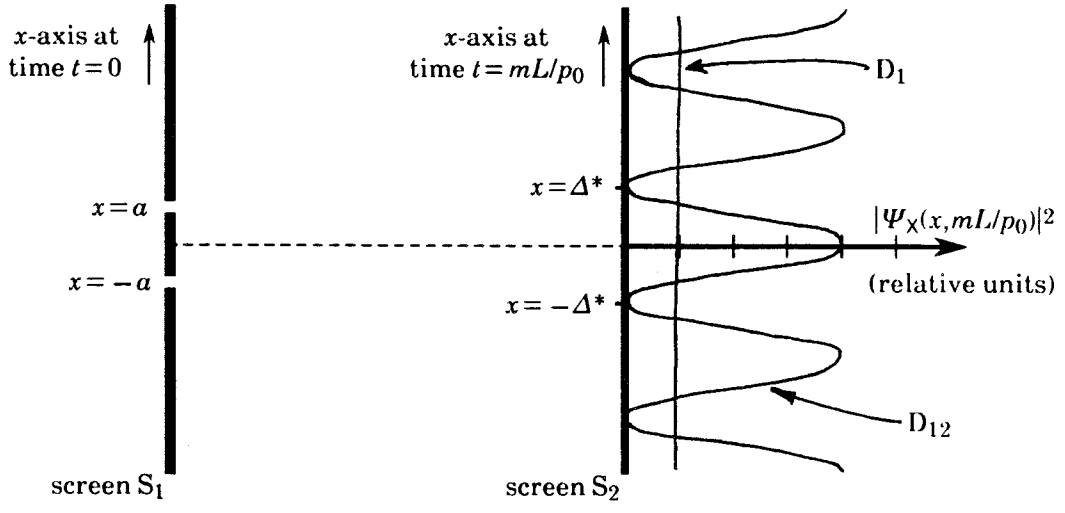


FIG. 7-3. Quantum free-particle predictions for the double-slit experiment, assuming the slit openings are infinitesimally narrow. Curve D_1 is for only slit 1 open, and curve D_{12} is for both slits open, both as deduced from Fig. 7-1.

In Fig. 7-3 we show the consequent quantum predictions of Fig. 7-1, as given by $|\Psi_X(x, t = mL/p_0)|^2$, for the electron hit probability patterns on screen S_2; curve D_1 is the predicted pattern when only slit 1 is open, and curve D_{12} is the predicted pattern when both slits are open. Notice from Eqs. (7-21) and (7-22) that the distance Δ^* between an adjacent maximum and minimum of curve D_{12} is predicted by quantum mechanics to be

$$\Delta^* = \Delta(t = mL/p_0) = \frac{\pi\hbar}{2ma}\left(\frac{mL}{p_0}\right) = \frac{\pi\hbar L}{2ap_0}. \tag{7-23}$$

Now let's compare these quantum predictions with the experimental results sketched in Fig. 1-2. We immediately notice an apparent discrepancy: Whereas the experimental curves C_1 and C_{12} die off in amplitude with increasing distance from the center line, our theoretical curves D_1 and D_{12} have constant amplitudes. But if we were to repeat our experiment with *smaller slit openings*, we would actually find that the curves C_1 and C_{12} would then die off more slowly. And we would infer that in the idealized limit of *zero*-width slits, which we assumed for the sake of mathematical simplicity in our quantum calculations, curves C_1 and C_{12} would not die off at all. So, for comparison purposes, we may ignore the amplitude tail-offs in curves C_1 and C_{12}. When we do that, the two sets of curves seem

to match rather well. In each case, the two-slit curve oscillates smoothly between about 0 and 4, on a scale where 1 is the constant amplitude of the one-slit curve.

A crucial test is provided by comparing the wavelengths of the oscillations of the two-slit curves. Remember in our first lecture, we said that the experimental curve looked like the two-slit intensity interference pattern of a *wave* with wavelength $\lambda = 2\pi\hbar/p_0$. In Fig. 7-4 we show the condition defining

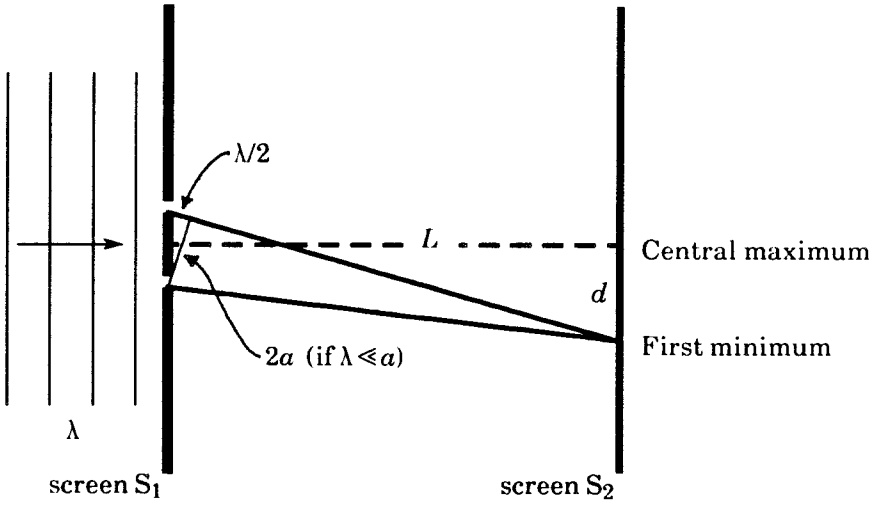


FIG. 7-4. Two-slit diffraction of an ordinary wave, showing the condition for the location of the first minimum in the intensity interference pattern.

the location of the first minimum in such an interference pattern: d is such that the difference in the distances to the two slits is $\lambda/2$, resulting in destructive interference at that point. For the experimental condition $\lambda \ll a$, we have by similar triangles that

$$(\lambda/2)/(2a) = d/L.$$

Solving for d, and then invoking the aforementioned result $\lambda = 2\pi\hbar/p_0$, we get

$$d = \frac{L\lambda}{4a} = \frac{L}{4a}\left(\frac{2\pi\hbar}{p_0}\right) = \frac{\pi\hbar L}{2a\,p_0}.$$

This agrees exactly with our quantum prediction for Δ^* in Eq. (7-23)!

Our conclusion: *Quantum mechanics provides an accurate description of the double-slit experiment — an experiment that defies a plausible description in terms of classical mechanics.* On that optimistic note, which is echoed again and again for the other systems on Nature's amazing microscale, we shall conclude our brief look at the theory of quantum mechanics.

●*Exercise 7-5.*

(a) Suppose the electron in the double-slit experiment acquires its horizontal momentum p_0 by being accelerated through a potential difference V, thus acquiring a kinetic energy eV, where e is the electron's charge. Write the formula for the distance Δ^* in Fig. 7-3 in terms of m, e and V.

(b) If $V = 10$ volts and $L = 10$ meters, what should a be in order to make $\Delta^* = 1$ millimeter? [Answer: About 100 angstroms, which is why this experiment is so hard to do.]

(c) With the values of V, L and a as in (b), suppose the electron were a "macroscopic" particle, say of mass 10^{-5} grams. Calculate Δ^* for that case. Would the oscillations in the two-slit curve be detectable in this macroscopic case?

* * *