

The Hamiltonian Matrix

8-1 Amplitudes and vectors

Before we begin the main topic of this chapter, we would like to describe a number of mathematical ideas that are used a lot in the literature of quantum mechanics. Knowing them will make it easier for you to read other books or papers on the subject. The first idea is the close mathematical resemblance between the equations of quantum mechanics and those of the scalar product of two vectors. You remember that if χ and ϕ are two states, the amplitude to start in ϕ and end up in χ can be written as a sum over a complete set of base states of the amplitude to go from ϕ into one of the base states and then from that base state out again into χ :

$$\langle \chi | \phi \rangle = \sum_{\text{all } i} \langle \chi | i \rangle \langle i | \phi \rangle. \quad (8.1)$$

We explained this in terms of a Stern-Gerlach apparatus, but we remind you that there is no need to have the apparatus. Equation (8.1) is a mathematical law that is just as true whether we put the filtering equipment in or not—it is not always necessary to imagine that the apparatus is there. We can think of it simply as a formula for the amplitude $\langle \chi | \phi \rangle$.

We would like to compare Eq. (8.1) to the formula for the dot product of two vectors \mathbf{B} and \mathbf{A} . If \mathbf{B} and \mathbf{A} are ordinary vectors in three dimensions, we can write the dot product this way:

$$\sum_{\text{all } i} (\mathbf{B} \cdot \mathbf{e}_i)(\mathbf{e}_i \cdot \mathbf{A}), \quad (8.2)$$

with the understanding that the symbol \mathbf{e}_i stands for the three unit vectors in the x , y , and z -directions. Then $\mathbf{B} \cdot \mathbf{e}_1$ is what we ordinarily call B_x ; $\mathbf{B} \cdot \mathbf{e}_2$ is what we ordinarily call B_y ; and so on. So Eq. (8.2) is equivalent to

$$B_x A_x + B_y A_y + B_z A_z,$$

which is the dot product $\mathbf{B} \cdot \mathbf{A}$.

Comparing Eqs. (8.1) and (8.2), we can see the following analogy: The states χ and ϕ correspond to the two vectors \mathbf{A} and \mathbf{B} . The base states i correspond to the special vectors \mathbf{e}_i to which we refer all other vectors. Any vector can be represented as a linear combination of the three “base vectors” \mathbf{e}_i . Furthermore, if you know the coefficients of each “base vector” in this combination—that is, its three components—you know everything about a vector. In a similar way, any quantum mechanical state can be described completely by the amplitude $\langle i | \phi \rangle$ to go into the base states; and if you know these coefficients, you know everything there is to know about the state. Because of this close analogy, what we have called a “state” is often also called a “state vector.”

Since the base vectors \mathbf{e}_i are all at right angles, we have the relation

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}. \quad (8.3)$$

This corresponds to the relations (5.25) among the base states i ,

$$\langle i | j \rangle = \delta_{ij}. \quad (8.4)$$

You see now why one says that the base states i are all “orthogonal.”

8-1 Amplitudes and vectors

8-2 Resolving state vectors

8-3 What are the base states of the world?

8-4 How states change with time

8-5 The Hamiltonian matrix

8-6 The ammonia molecule

Review: Chapter 49, Vol. I, *Modes*

There is one minor difference between Eq. (8.1) and the dot product. We have that

$$\langle \phi | \chi \rangle = \langle \chi | \phi \rangle^* \quad (8.5)$$

But in vector algebra,

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}.$$

With the complex numbers of quantum mechanics we have to keep straight the order of the terms, whereas in the dot product, the order doesn't matter.

Now consider the following vector equation:

$$\mathbf{A} = \sum_i \mathbf{e}_i (\mathbf{e}_i \cdot \mathbf{A}). \quad (8.6)$$

It's a little unusual, but correct. It means the same thing as

$$\mathbf{A} = \sum_i A_i \mathbf{e}_i = A_x \mathbf{e}_x + A_y \mathbf{e}_y + A_z \mathbf{e}_z. \quad (8.7)$$

Notice, though, that Eq. (8.6) involves a quantity which is *different* from a dot product. A dot product is just a *number*, whereas Eq. (8.6) is a *vector* equation. One of the great tricks of vector analysis was to abstract away from the equations the idea of a *vector* itself. One might be similarly inclined to abstract a thing that is the analog of a "vector" from the quantum mechanical formula Eq. (8.1)—and one can indeed. We remove the $\langle \chi |$ from both sides Eq. (8.1) and write the following equation (don't get frightened—it's just a notation and in a few minutes you will find out what the symbols mean):

$$| \phi \rangle = \sum_i | i \rangle \langle i | \phi \rangle. \quad (8.8)$$

One thinks of the bracket $\langle \chi | \phi \rangle$ as being divided into two pieces. The second piece $| \phi \rangle$ is often called a *ket*, and the first piece $\langle \chi |$ is called a *bra* (put together, they make a "bra-ket"—a notation proposed by Dirac); the half-symbols $\langle \chi |$ and $| \phi \rangle$ are also called *state vectors*. In any case, they are *not* numbers, and, in general, we want the results of our calculations to come out as numbers; so such "unfinished" quantities are only part-way steps in our calculations.

It happens that until now we have written all our results in terms of numbers. How have we managed to avoid vectors? It is amusing to note that even in ordinary vector algebra we *could* make all equations involve only numbers. For instance, instead of a vector equation like

$$\mathbf{F} = m\mathbf{a},$$

we could always have written

$$\mathbf{C} \cdot \mathbf{F} = \mathbf{C} \cdot (m\mathbf{a}).$$

We have then an equation between dot products that is true for *any* vector \mathbf{C} . But if it is true for any \mathbf{C} , it hardly makes sense at all to keep writing the \mathbf{C} !

Now look at Eq. (8.1). It is an equation that is true for *any* χ . So to save writing, we should just leave *out* the χ and write Eq. (8.8) instead. It has the same information *provided* we understand that it should always be "finished" by "multiplying on the left by"—which simply means reinserting—some $\langle \chi |$ on both sides. So Eq. (8.8) means exactly the same thing as Eq. (8.1)—no more, no less. When you want numbers, you put in the $\langle \chi |$ you want.

Maybe you have already wondered about the ϕ in Eq. (8.8). Since the equation is true for *any* ϕ , why do we keep *it*? Indeed, Dirac suggests that the ϕ also can just as well be abstracted away, so that we have only

$$| = \sum_i | i \rangle \langle i |. \quad (8.9)$$

And this is the great law of quantum mechanics! (There is no analog in vector analysis.) It says that if you put *in* any two states χ and ϕ on the left and right of both sides, you *get back* Eq. (8.1). It is not really very useful, but it's a nice reminder that the equation is true for any two states.

8-2 Resolving state vectors

Let's look at Eq. (8.8) again; we can think of it in the following way. Any state vector $|\phi\rangle$ can be represented as a linear combination with suitable coefficients of a set of base "vectors"—or, if you prefer, as a superposition of "unit vectors" in suitable proportions. To emphasize that the coefficients $\langle i|\phi\rangle$ are just ordinary (complex) numbers, suppose we write

$$\langle i|\phi\rangle = C_i.$$

Then Eq. (8.8) is the same as

$$|\phi\rangle = \sum_i |i\rangle C_i. \quad (8.10)$$

We can write a similar equation for any other state vector, say $|\chi\rangle$, with, of course, different coefficients—say D_i . Then we have

$$|\chi\rangle = \sum_i |i\rangle D_i. \quad (8.11)$$

The D_i are just the amplitudes $\langle i|\chi\rangle$.

Suppose we had started by abstracting the ϕ from Eq. (8.1). We would have had

$$\langle \chi| = \sum_i \langle \chi|i\rangle \langle i|. \quad (8.12)$$

Remembering that $\langle \chi|i\rangle = \langle i|\chi\rangle^*$, we can write this as

$$\langle \chi| = \sum_i D_i^* \langle i|. \quad (8.13)$$

Now the interesting thing is that we can just *multiply* Eq. (8.13) and Eq. (8.10) to get back $\langle \chi|\phi\rangle$. When we do that, we have to be careful of the summation indices, because they are quite distinct in the two equations. Let's first rewrite Eq. (8.13) as

$$\langle \chi| = \sum_j D_j^* \langle j|,$$

which changes nothing. Then putting it together with Eq. (8.10), we have

$$\langle \chi|\phi\rangle = \sum_{ij} D_j^* \langle j|i\rangle C_i. \quad (8.14)$$

Remember, though, that $\langle j|i\rangle = \delta_{ij}$, so that in the sum we have left only the terms with $j = i$. We get

$$\langle \chi|\phi\rangle = \sum_i D_i^* C_i, \quad (8.15)$$

where, of course, $D_i^* = \langle i|\chi\rangle^* = \langle \chi|i\rangle$, and $C_i = \langle i|\phi\rangle$. Again we see the close analogy with the dot product

$$\mathbf{A} \cdot \mathbf{B} = \sum_i A_i B_i.$$

The only difference is the complex conjugate on D_i . So Eq. (8.15) says that if the state vectors $\langle \chi|$ and $|\phi\rangle$ are expanded in terms of the base vectors $\langle i|$ or $|i\rangle$, the amplitude to go from ϕ to χ is given by the kind of dot product in Eq. (8.15). This equation is, of course, just Eq. (8.1) written with different symbols. So we have just gone in a circle to get used to the new symbols.

We should perhaps emphasize again that while space vectors in three dimensions are described in terms of *three* orthogonal unit vectors, the base vectors $|i\rangle$ of the quantum mechanical states must range over the complete set applicable to any particular problem. Depending on the situation, two, or three, or five, or an infinite number of base states may be involved.

We have also talked about what happens when particles go through an apparatus. If we start the particles out in a certain state ϕ , then send them through

an apparatus, and afterward make a measurement to see if they are in state χ , the result is described by the amplitude

$$\langle \chi | A | \phi \rangle. \quad (8.16)$$

Such a symbol doesn't have a close analog in vector algebra. (It is closer to tensor algebra, but the analogy is not particularly useful.) We saw in Chapter 5, Eq. (5.32), that we could write (8.16) as

$$\langle \chi | A | \phi \rangle = \sum_{ij} \langle \chi | i \rangle \langle i | A | j \rangle \langle j | \phi \rangle. \quad (8.17)$$

This is just an example of the fundamental rule Eq. (8.9), used twice.

We also found that if another apparatus B was added in series with A , then we could write

$$\langle \chi | BA | \phi \rangle = \sum_{ijk} \langle \chi | i \rangle \langle i | B | j \rangle \langle j | A | k \rangle \langle k | \phi \rangle. \quad (8.18)$$

Again, this comes directly from Dirac's method of writing Eq. (8.9)—remember that we can always place a bar ($|$), which is just like the factor 1, between B and A .

Incidentally, we can think of Eq. (8.17) in another way. Suppose we think of the particle entering apparatus A in the state ϕ and coming out of A in the state ψ ("psi"). In other words, we could ask ourselves this question: Can we find a ψ such that the amplitude to get from ψ to χ is always identically and everywhere the same as the amplitude $\langle \chi | A | \phi \rangle$? The answer is yes. We want Eq. (8.17) to be replaced by

$$\langle \chi | \psi \rangle = \sum_i \langle \chi | i \rangle \langle i | \psi \rangle. \quad (8.19)$$

We can clearly do this if

$$\langle i | \psi \rangle = \sum_j \langle i | A | j \rangle \langle j | \phi \rangle = \langle i | A | \phi \rangle, \quad (8.20)$$

which determines ψ . "But it doesn't determine ψ ," you say; "it only determines $\langle i | \psi \rangle$." However, $\langle i | \psi \rangle$ *does* determine ψ , because if you have all the coefficients that relate ψ to the base states i , then ψ is uniquely defined. In fact, we can play with our notation and write the last term of Eq. (8.20) as

$$\langle i | \psi \rangle = \sum_j \langle i | j \rangle \langle j | A | \phi \rangle. \quad (8.21)$$

Then, since this equation is true for all i , we can write simply

$$| \psi \rangle = \sum_j | j \rangle \langle j | A | \phi \rangle. \quad (8.22)$$

Then we can say: "The state ψ is what we get if we start with ϕ and go through the apparatus A ."

One final example of the tricks of the trade. We start again with Eq. (8.17). Since it is true for any χ and ϕ , we can drop them both! We then get†

$$A = \sum_{ij} | i \rangle \langle i | A | j \rangle \langle j |. \quad (8.23)$$

What does it mean? It means no more, no less, than what you get if you put back the ϕ and χ . As it stands, it is an "open" equation and incomplete. If we multiply it "on the left" by $| \phi \rangle$, it becomes

$$A | \phi \rangle = \sum_{ij} | i \rangle \langle i | A | j \rangle \langle j | \phi \rangle, \quad (8.24)$$

† You might think we should write $|A|$ instead of just A . But then it would look like the symbol for "absolute value of A ," so the bars are usually dropped. In general, the bar ($|$) behaves much like the factor one.

which is just Eq. (8.22) all over again. In fact, we could have just dropped the j 's from that equation and written

$$|\psi\rangle = A|\phi\rangle. \quad (8.25)$$

The symbol A is neither an amplitude, nor a vector; it is a new kind of thing called an *operator*. It is something which “operates on” a state to produce a new state—Eq. (8.25) says that $|\psi\rangle$ is what results if A operates on $|\phi\rangle$. Again, it is still an open equation until it is completed with some bra like $\langle x|$ to give

$$\langle x|\psi\rangle = \langle x|A|\phi\rangle. \quad (8.26)$$

The operator A is, of course, described completely if we give the matrix of amplitudes $\langle i|A|j\rangle$ —also written A_{ij} —in terms of any set of base vectors.

We have really added nothing new with all of this new mathematical notation. One reason for bringing it all up was to show you the way of writing pieces of equations, because in many books you will find the equations written in the incomplete forms, and there's no reason for you to be paralyzed when you come across them. If you prefer, you can always add the missing pieces to make an equation between numbers that will look like something more familiar.

Also, as you will see, the “bra” and “ket” notation is a very convenient one. For one thing, we can from now on identify a state by giving its state vector. When we want to refer to a state of definite momentum p we can say: “the state $|p\rangle$ ”. Or we may speak of some arbitrary state $|\psi\rangle$. For consistency we will always use the ket, writing $|\psi\rangle$, to identify a state. (It is, of course an arbitrary choice; we could equally well have chosen to use the bra, $\langle\psi|$.)

8-3 What are the base states of the world?

We have discovered that any state in the world can be represented as a superposition—a linear combination with suitable coefficients—of base states. You may ask, first of all, *what* base states? Well, there are many different possibilities. You can, for instance, project a spin in the z -direction or in some other direction. There are many, many different *representations*, which are the analogs of the different *coordinate systems* one can use to represent ordinary vectors. Next, *what* coefficients? Well, that depends on the physical circumstances. Different sets of coefficients correspond to different physical conditions. The important thing to know about is the “space” in which you are working—in other words, what the base states mean physically. So the first thing you have to know about, in general, is what the base states are like. Then you can understand how to describe a situation in terms of these base states.

We would like to look ahead a little and speak a bit about what the general quantum mechanical description of nature is going to be—in terms of the now current ideas of physics, anyway. First, one decides on a particular representation for the base states—different representations are always possible. For example, for a spin one-half particle we can use the plus and minus states with respect to the z -axis. But there's nothing special about the z -axis—you can take any other axis you like. For consistency we'll always pick the z -axis, however. Suppose we begin with a situation with one electron. In addition to the two possibilities for the spin (“up” and “down” along the z -direction), there is also the momentum of the electron. We pick a set of base states, each corresponding to one value of the momentum. What if the electron doesn't have a definite momentum? That's all right; we're just saying what the *base* states are. If the electron hasn't got a definite momentum, it has some amplitude to have one momentum and another amplitude to have another momentum, and so on. And if it is not necessarily spinning up, it has some amplitude to be spinning up going at this momentum, and some amplitude to be spinning down going at that momentum, and so on. The complete description of an electron, *so far as we know*, requires only that the base states be described by the *momentum* and the *spin*. So one acceptable set of base states $|i\rangle$ for a single electron refer to different values of the momentum and

whether the spin is up or down. Different mixtures of amplitudes—that is, different combinations of the C 's describe different circumstances. What any particular electron is doing is described by telling with what amplitude it has an up-spin or a down-spin and one momentum or another—for all possible momenta. So you can see what is involved in a complete quantum mechanical description of a single electron.

What about systems with more than one electron? Then the base states get more complicated. Let's suppose that we have two electrons. We have, first of all, four possible states with respect to spin: both electrons spinning up, the first one down and the second one up, the first one up and the second one down, or both down. Also we have to specify that the first electron has the momentum p_1 , and the second electron, the momentum p_2 . The base states for two electrons require the specification of two momenta and two spin characters. With seven electrons, we have to specify seven of each.

If we have a proton and an electron, we have to specify the spin direction of the proton and its momentum, and the spin direction of the electron and its momentum. At least that's approximately true. *We do not really know* what the correct representation is for the world. It is all very well to start out by supposing that if you specify the spin in the electron and its momentum, and likewise for a proton, you will have the base states; but what about the "guts" of the proton? Let's look at it this way. In a hydrogen atom which has one proton and one electron, we have many different base states to describe—up and down spins of the proton and electron and the various possible momenta of the proton and electron. Then there are different combinations of amplitudes C_i which together describe the character of the hydrogen atom in different states. But suppose we look at the whole hydrogen atom as a "particle." If we didn't know that the hydrogen atom was made out of a proton and an electron, we might have started out and said: "Oh, I know what the base states are—they correspond to a particular momentum of the hydrogen atom." No, because the hydrogen atom has internal parts. It may, therefore, have various states of different internal energy, and describing the real nature requires more detail.

The question is: Does a proton have internal parts? Do we have to describe a proton by giving all possible states of protons, and mesons, and strange particles? We don't know. And even though we suppose that the electron is simple, so that all we have to tell about it is its momentum and its spin, maybe tomorrow we will discover that the electron also has inner gears and wheels. It would mean that our representation is incomplete, or wrong, or approximate—in the same way that a representation of the hydrogen atom which describes only its momentum would be incomplete, because it disregarded the fact that the hydrogen atom could have become excited inside. If an electron could become excited inside and turn into something else like, for instance, a muon, then it would be described not just by giving the states of the new particle, but presumably in terms of some more complicated internal wheels. *The main problem in the study of the fundamental particles today* is to discover what are the correct representations for the description of nature. At the present time, we *guess* that for the electron it is enough to specify its momentum and spin. We also guess that there is an idealized proton which has its π -mesons, and k -mesons, and so on, that all have to be specified. Several dozen particles—that's crazy! The question of what *is* a fundamental particle and what *is not* a fundamental particle—a subject you hear so much about these days—is the question of what is the final *representation* going to look like in the ultimate quantum mechanical description of the world. Will the electron's momentum still be the right thing with which to describe nature? Or even, should the whole question be put this way at all! This question must always come up in any scientific investigation. At any rate, we see a problem—how to find a representation. We don't know the answer. We don't even know whether we have the "right" problem, but if we do, we must first attempt to find out whether any particular particle is "fundamental" or not.

In the nonrelativistic quantum mechanics—if the energies are not too high, so that you don't disturb the inner workings of the strange particles and so forth—

you can do a pretty good job without worrying about these details. You can just decide to specify the momenta and spins of the electrons and of the nuclei; then everything will be all right. In most chemical reactions and other low-energy happenings, nothing goes on in the nuclei; they don't get excited. Furthermore, if a hydrogen atom is moving slowly and bumping quietly against other hydrogen atoms—never getting excited inside, or radiating, or anything complicated like that, but staying always in the ground state of energy for internal motion—you can use an approximation in which you talk about the hydrogen atom as one object, or particle, and not worry about the fact that it *can* do something inside. This will be a good approximation as long as the kinetic energy in any collision is well below 10 electron volts—the energy required to excite the hydrogen atom to a different internal state. We will often be making an approximation in which we do not include the possibility of inner motion, thereby decreasing the number of details that we have to put into our base states. Of course, we then omit some phenomena which would appear (usually) at some higher energy, but by making such approximations we can simplify very much the analysis of physical problems. For example, we can discuss the collision of two hydrogen atoms at low energy—or any chemical process—without worrying about the fact that the atomic nuclei could be excited. To summarize, then, when we can neglect the effects of any internal excited states of a particle we can choose a base set which are the states of definite momentum and z-component of angular momentum.

One problem then in describing nature is to find a suitable representation for the base states. But that's only the beginning. We still want to be able to say what "happens." If we know the "condition" of the world at one moment, we would like to know the condition at a later moment. So we also have to find the laws that determine how things change with time. We now address ourselves to this second part of the framework of quantum mechanics—how states change with time.

8-4 How states change with time

We have already talked about how we can represent a situation in which we put something through an apparatus. Now one convenient, delightful "apparatus" to consider is merely a wait of a few minutes; that is, you prepare a state ϕ , and then before you analyze it, you just let it sit. Perhaps you let it sit in some particular electric or magnetic field—it depends on the physical circumstances in the world. At any rate, whatever the conditions are, you let the object sit from time t_1 to time t_2 . Suppose that it is let out of your first apparatus in the condition ϕ at t_1 . And then it goes through an "apparatus," but the "apparatus" consists of just delay until t_2 . During the delay, various things could be going on—external forces applied or other shenanigans—so that something is happening. At the end of the delay, the amplitude to find the thing in some state χ is no longer exactly the same as it would have been without the delay. Since "waiting" is just a special case of an "apparatus," we can describe what happens by giving an amplitude with the same form as Eq. (8.17). Because the operation of "waiting" is especially important, we'll call it U instead of A , and to specify the starting and finishing times t_1 and t_2 , we'll write $U(t_2, t_1)$. The amplitude we want is

$$\langle \chi | U(t_2, t_1) | \phi \rangle. \quad (8.27)$$

Like any other such amplitude, it can be represented in some base system or other by writing it

$$\sum_{ij} \langle \chi | i \rangle \langle i | U(t_2, t_1) | j \rangle \langle j | \phi \rangle. \quad (8.28)$$

Then U is completely described by giving the whole set of amplitudes—the matrix

$$\langle i | U(t_2, t_1) | j \rangle. \quad (8.29)$$

We can point out, incidentally, that the matrix $\langle i | U(t_2, t_1) | j \rangle$ gives much more detail than may be needed. The high-class theoretical physicist working in

high-energy physics considers problems of the following general nature (because it's the way experiments are usually done). He starts with a couple of particles, like a proton and a proton, coming together from infinity. (In the lab, usually one particle is standing still, and the other comes from an accelerator that is practically at infinity on atomic level.) The things go crash and out come, say, two k -mesons, six π -mesons, and two neutrons in certain directions with certain momenta. What's the amplitude for this to happen? The mathematics looks like this: The ϕ -state specifies the spins and momenta of the incoming particles. The χ would be the question about what comes out. For instance, with what amplitude do you get the six mesons going in such-and-such directions, and the two neutrons going off in these directions, with their spins so-and-so. In other words, χ would be specified by giving all the momenta, and spins, and so on of the final products. Then the job of the theorist is to calculate the amplitude (8.27). However, he is really only interested in the special case that t_1 is $-\infty$ and t_2 is $+\infty$. (There is no experimental evidence on the details of the process, only on what comes in and what goes out.) The limiting case of $U(t_2, t_1)$ as $t_1 \rightarrow -\infty$ and $t_2 \rightarrow +\infty$ is called S , and what he wants is

$$\langle \chi | S | \phi \rangle.$$

Or, using the form (8.28), he would calculate the matrix

$$\langle i | S | j \rangle,$$

which is called the S -matrix. So if you see a theoretical physicist pacing the floor and saying, "All I have to do is calculate the S -matrix," you will know what he is worried about.

How to analyze—how to specify the laws for—the S -matrix is an interesting question. In relativistic quantum mechanics for high energies, it is done one way, but in nonrelativistic quantum mechanics it can be done another way, which is very convenient. (This other way can also be done in the relativistic case, but then it is not so convenient.) It is to work out the U -matrix for a small interval of time—in other words for t_2 and t_1 close together. If we can find a sequence of such U 's for successive intervals of time we can watch how things go as a function of time. You can appreciate immediately that this way is not so good for relativity, because you don't want to have to specify how everything looks "simultaneously" everywhere. But we won't worry about that—we're just going to worry about non-relativistic mechanics.

Suppose we think of the matrix U for a delay from t_1 until t_3 which is greater than t_2 . In other words, let's take three successive times: t_1 less than t_2 less than t_3 . Then we claim that the matrix that goes between t_1 and t_3 is the *product* in succession of what happens when you delay from t_1 until t_2 and then from t_2 until t_3 . It's just like the situation when we had two apparatuses B and A in series. We can then write, following the notation of Section 5-6,

$$U(t_3, t_1) = U(t_3, t_2) \cdot U(t_2, t_1). \quad (8.30)$$

In other words, we can analyze any time interval if we can analyze a sequence of short time intervals in between. We just multiply together all the pieces; that's the way that quantum mechanics is analyzed nonrelativistically.

Our problem, then, is to understand the matrix $U(t_2, t_1)$ for an infinitesimal time interval—for $t_2 = t_1 + \Delta t$. We ask ourselves this: If we have a state ϕ now, what does the state look like an infinitesimal time Δt later? Let's see how we write that out. Call the state at the time t , $|\psi(t)\rangle$ (we show the time dependence of ψ to be perfectly clear that we mean the condition at the time t). Now we ask the question: What is the condition after the small interval of time Δt later? The answer is

$$|\psi(t + \Delta t)\rangle = U(t + \Delta t, t) |\psi(t)\rangle. \quad (8.31)$$

This means the same as we meant by (8.25), namely, that the amplitude to

find x at the time $t + \Delta t$, is

$$\langle x | \psi(t + \Delta t) \rangle = \langle x | U(t + \Delta t, t) | \psi(t) \rangle. \quad (8.32)$$

Since we're not yet too good at these abstract things, let's project our amplitudes into a definite representation. If we multiply both sides of Eq. (8.31) by $\langle i |$, we get

$$\langle i | \psi(t + \Delta t) \rangle = \langle i | U(t + \Delta t, t) | \psi(t) \rangle. \quad (8.33)$$

We can also resolve the $|\psi(t)\rangle$ into base states and write

$$\langle i | \psi(t + \Delta t) \rangle = \sum_j \langle i | U(t + \Delta t, t) | j \rangle \langle j | \psi(t) \rangle. \quad (8.34)$$

We can understand Eq. (8.34) in the following way. If we let $C_i(t) = \langle i | \psi(t) \rangle$ stand for the amplitude to be in the base state i at the time t , then we can think of this amplitude (just a *number*, remember!) varying with time. Each C_i becomes a function of t . And we also have some information on *how* the amplitudes C_i vary with time. Each amplitude at $(t + \Delta t)$ is proportional to *all of the other* amplitudes at t multiplied by a set of coefficients. Let's call the U -matrix U_{ij} , by which we mean

$$U_{ij} = \langle i | U | j \rangle.$$

Then we can write Eq. (8.34) as

$$C_i(t + \Delta t) = \sum_j U_{ij}(t + \Delta t, t) C_j(t). \quad (8.35)$$

This, then, is how the dynamics of quantum mechanics is going to look.

We don't know much about the U_{ij} yet, except for one thing. We know that if Δt goes to zero, nothing can happen—we should get just the original state. So, $U_{ii} \rightarrow 1$ and $U_{ij} \rightarrow 0$, if $i \neq j$. In other words, $U_{ij} \rightarrow \delta_{ij}$ for $\Delta t \rightarrow 0$. Also, we can suppose that for small Δt , each of the coefficients U_{ij} should differ from δ_{ij} by amounts proportional to Δt ; so we can write

$$U_{ij} = \delta_{ij} + K_{ij} \Delta t. \quad (8.36)$$

However, it is usual to take the factor $(-i/\hbar)^\dagger$ out of the coefficients K_{ij} , for historical and other reasons; we prefer to write

$$U_{ij}(t + \Delta t, t) = \delta_{ij} - \frac{i}{\hbar} H_{ij}(t) \Delta t. \quad (8.37)$$

It is, of course, the same as Eq. (8.36) and, if you wish, just defines the coefficients $H_{ij}(t)$. The terms H_{ij} are just the derivatives with respect to t_2 of the coefficients $U_{ij}(t_2, t_1)$, evaluated at $t_2 = t_1 = t$.

Using this form for U in Eq. (8.35), we have

$$C_i(t + \Delta t) = \sum_j \left[\delta_{ij} - \frac{i}{\hbar} H_{ij}(t) \Delta t \right] C_j(t). \quad (8.38)$$

Taking the sum over the δ_{ij} term, we get just $C_i(t)$, which we can put on the other side of the equation. Then dividing by Δt , we have what we recognize as a derivative

$$\frac{C_i(t + \Delta t) - C_i(t)}{\Delta t} = -\frac{i}{\hbar} \sum_j H_{ij}(t) C_j(t)$$

or

$$i\hbar \frac{dC_i(t)}{dt} = \sum_j H_{ij}(t) C_j(t). \quad (8.39)$$

[†] We are in a bit of trouble here with notation. In the factor $(-i/\hbar)$, the i means the imaginary unit $\sqrt{-1}$, and *not* the *index* i that refers to the i th base state! We hope that you won't find it too confusing.

You remember that $C_i(t)$ is the amplitude $\langle i | \psi \rangle$ to find the state ψ in one of the base states i (at the time t). So Eq. (8.39) tells us how each of the coefficients $\langle i | \psi \rangle$ varies with time. But that is the same as saying that Eq. (8.39) tells us how the state ψ varies with time, since we are describing ψ in terms of the amplitudes $\langle i | \psi \rangle$. The variation of ψ in time is described in terms of the matrix H_{ij} , which has to include, of course, the things we are doing to the system to cause it to change. If we know the H_{ij} —which contains the physics of the situation and can, in general, depend on the time—we have a complete description of the behavior in time of the system. Equation (8.39) is then the quantum mechanical law for the dynamics of the world.

(We should say that we will always take a set of base states which are fixed and do not vary with time. There are people who use base states that also vary. However, that's like using a rotating coordinate system in mechanics, and we don't want to get involved in such complications.)

8-5 The Hamiltonian matrix

The idea, then, is that to describe the quantum mechanical world we need to pick a set of base states i and to write the physical laws by giving the matrix of coefficients H_{ij} . Then we have everything—we can answer any question about what will happen. So we have to learn what the rules are for finding the H 's to go with any physical situation—what corresponds to a magnetic field, or an electric field, and so on. And that's the hardest part. For instance, for the new strange particles, we have no idea what H_{ij} 's to use. In other words, no one knows the *complete* H_{ij} for the whole world. (Part of the difficulty is that one can hardly hope to discover the H_{ij} when no one even knows what the base states are!) We do have excellent approximations for nonrelativistic phenomena and for some other special cases. In particular, we have the forms that are needed for the motions of electrons in atoms—to describe chemistry. But we don't know the full true H for the whole universe.

The coefficients H_{ij} are called *the Hamiltonian matrix* or, for short, just *the Hamiltonian*. (How Hamilton, who worked in the 1830's, got his name on a quantum mechanical matrix is a tale of history.) It would be much better called the *energy matrix*, for reasons that will become apparent as we work with it. So *the* problem is: Know your Hamiltonian!

The Hamiltonian has one property that can be deduced right away, namely, that

$$H_{ij}^* = H_{ji}. \quad (8.40)$$

This follows from the condition that the total probability that the system is in *some* state does not change. If you start with a particle—an object or the world—then you've still got it as time goes on. The total probability of finding it somewhere is

$$\sum_i |C_i(t)|^2,$$

which must not vary with time. If this is to be true for any starting condition ϕ , then Eq. (8.40) must also be true.

As our first example, we take a situation in which the physical circumstances are not changing with time; we mean the *external* physical conditions, so that H is independent of time. Nobody is turning magnets on and off. We also pick a system for which only one base state is required for the description; it is an approximation we could make for a hydrogen atom at rest, or something similar. Equation (8.39) then says

$$i\hbar \frac{dC_1}{dt} = H_{11}C_1. \quad (8.41)$$

Only one equation—that's all! And if H_{11} is constant, this differential equation is easily solved to give

$$C_1 = (\text{const})e^{-(i/\hbar)H_{11}t} \quad (8.42)$$

This is the time dependence of a state with a definite energy $E = H_{11}$. You see why H_{ij} ought to be called the energy matrix. It is the generalization of the energy for more complex situations.

Next, to understand a little more about what the equations mean, we look at a system which has two base states. Then Eq. (8.39) reads

$$\begin{aligned} i\hbar \frac{dC_1}{dt} &= H_{11}C_1 + H_{12}C_2, \\ i\hbar \frac{dC_2}{dt} &= H_{21}C_1 + H_{22}C_2. \end{aligned} \quad (8.43)$$

If the H 's are again independent of time, you can easily solve these equations. We leave you to try for fun, and we'll come back and do them later. Yes, you can solve the quantum mechanics without knowing the H 's, so long as they are independent of time.

8-6 The ammonia molecule

We want now to show you how the dynamical equation of quantum mechanics can be used to describe a particular physical circumstance. We have picked an interesting but simple example in which, by making some reasonable guesses about the Hamiltonian, we can work out some important—and even practical—results. We are going to take a situation describable by two states: the ammonia molecule.

The ammonia molecule has one nitrogen atom and three hydrogen atoms located in a plane below the nitrogen so that the molecule has the form of a pyramid, as drawn in Fig. 8-1(a). Now this molecule, like any other, has an infinite number of states. It can spin around any possible axis; it can be moving in any direction; it can be vibrating inside, and so on, and so on. It is, therefore, not a two-state system at all. But we want to make an approximation that all other states remain fixed, because they don't enter into what we are concerned with at the moment. We will consider only that the molecule is spinning around its axis of symmetry (as shown in the figure), that it has zero translational momentum, and that it is vibrating as little as possible. That specifies all conditions except one: *there are still the two possible positions for the nitrogen atom*—the nitrogen may be on one side of the plane of hydrogen atoms or on the other, as shown in Fig. 8-1(a) and (b). So we will discuss the molecule as though it were a two-state system. We mean that there are only two states we are going to really worry about, all other things being assumed to stay put. You see, even if we know that it is spinning with a certain angular momentum around the axis and that it is moving with a certain momentum and vibrating in a definite way, there are still two possible states. We will say that the molecule is in the state $|1\rangle$ when the nitrogen is “up,” as in Fig. 8-1(a), and is in the state $|2\rangle$ when the nitrogen is “down,” as in (b). The states $|1\rangle$ and $|2\rangle$ will be taken as the set of base states for our analysis of the behavior of the ammonia molecule. At any moment, the actual state $|\psi\rangle$ of the molecule can be represented by giving $C_1 = \langle 1 | \psi \rangle$, the amplitude to be in state $|1\rangle$, and $C_2 = \langle 2 | \psi \rangle$, the amplitude to be in state $|2\rangle$. Then, using Eq. (8.8) we can write the state vector $|\psi\rangle$ as

$$\begin{aligned} |\psi\rangle &= |1\rangle\langle 1 | \psi \rangle + |2\rangle\langle 2 | \psi \rangle \\ \text{or} \quad |\psi\rangle &= |1\rangle C_1 + |2\rangle C_2. \end{aligned} \quad (8.44)$$

Now the interesting thing is that if the molecule is known to be in some state at some instant, it will *not* be in the same state a little while later. The two C -coefficients will be changing with time according to the equations (8.43)—which hold for any two-state system. Suppose, for example, that you had made some observation—or had made some selection of the molecules—so that you *know* that the molecule is *initially* in the state $|1\rangle$. At some later time, there is some chance that it will be found in state $|2\rangle$. To find out what this chance is, we have to solve the differential equation which tells us how the amplitudes change with time.

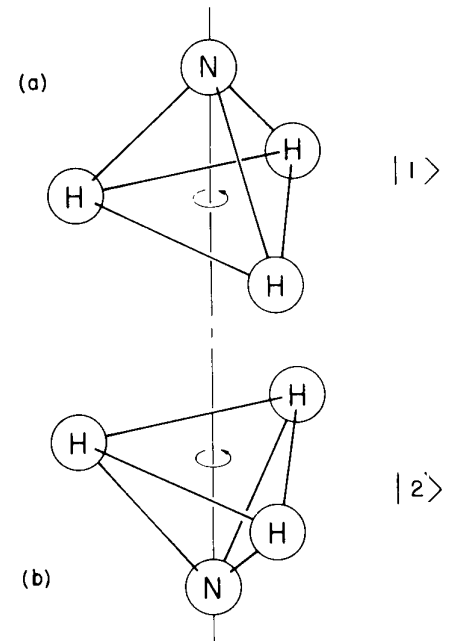


Fig. 8-1. Two equivalent geometric arrangements of the ammonia molecule.

The only trouble is that we don't know what to use for the coefficients H_{ij} in Eq. (8.43). There are some things we *can* say, however. Suppose that once the molecule was in the state $|1\rangle$ there was no chance that it could ever get into $|2\rangle$, and vice versa. Then H_{12} and H_{21} would both be zero, and Eq. (8.43) would read

$$i\hbar \frac{dC_1}{dt} = H_{11}C_1, \quad i\hbar \frac{dC_2}{dt} = H_{22}C_2.$$

We can easily solve these two equations; we get

$$C_1 = (\text{const})e^{-(i/\hbar)H_{11}t}, \quad C_2 = (\text{const})e^{-(i/\hbar)H_{22}t}. \quad (8.45)$$

These are just the amplitudes for *stationary* states with the energies $E_1 = H_{11}$ and $E_2 = H_{22}$. We note, however, that for the ammonia molecule the two states $|1\rangle$ and $|2\rangle$ have a definite symmetry. If nature is at all reasonable, the matrix elements H_{11} and H_{22} must be equal. We'll call them both E_0 , because they correspond to the energy the states would have if H_{12} and H_{21} were zero. But Eqs. (8.45) do not tell us what ammonia really does. It turns out that it is possible for the nitrogen to push its way through the three hydrogens and flip to the other side. It is quite difficult; to get half-way through requires a lot of energy. How can it get through if it hasn't got enough energy? There is *some* amplitude that it *will* penetrate the energy barrier. It is possible in quantum mechanics to sneak quickly across a region which is illegal energetically. There is, therefore, some small amplitude that a molecule which starts in $|1\rangle$ will get to the state $|2\rangle$. The coefficients H_{12} and H_{21} are not really zero. Again, by symmetry, they should both be the same—at least in magnitude. In fact, we already know that, in general, H_{ij} must be equal to the complex conjugate of H_{ji} , so they can differ only by a phase. It turns out, as you will see, that there is no loss of generality if we take them equal to each other. For later convenience we set them equal to a negative number; we take $H_{12} = H_{21} = -A$. We then have the following pair of equations:

$$i\hbar \frac{dC_1}{dt} = E_0C_1 - AC_2, \quad (8.46)$$

$$i\hbar \frac{dC_2}{dt} = E_0C_2 - AC_1. \quad (8.47)$$

These equations are simple enough and can be solved in any number of ways. One convenient way is the following. Taking the sum of the two, we get

$$i\hbar \frac{d}{dt} (C_1 + C_2) = (E_0 - A)(C_1 + C_2),$$

whose solution is

$$C_1 + C_2 = ae^{-(i/\hbar)(E_0 - A)t}. \quad (8.48)$$

Then, taking the difference of (8.46) and (8.47), we find that

$$i\hbar \frac{d}{dt} (C_1 - C_2) = (E_0 + A)(C_1 - C_2),$$

which gives

$$C_1 - C_2 = be^{-(i/\hbar)(E_0 + A)t}. \quad (8.49)$$

We have called the two integration constants a and b ; they are, of course, to be chosen to give the appropriate starting condition for any particular physical problem. Now, by adding and subtracting (8.48) and (8.49), we get C_1 and C_2 :

$$C_1(t) = \frac{a}{2} e^{-(i/\hbar)(E_0 - A)t} + \frac{b}{2} e^{-(i/\hbar)(E_0 + A)t}, \quad (8.50)$$

$$C_2(t) = \frac{a}{2} e^{-(i/\hbar)(E_0 - A)t} - \frac{b}{2} e^{-(i/\hbar)(E_0 + A)t}. \quad (8.51)$$

They are the same except for the sign of the second term.

We have the solutions; now what do they mean? (The trouble with quantum mechanics is not only in solving the equations but in understanding what the solutions mean!) First, notice that if $b = 0$, both terms have the same frequency $\omega = (E_0 - A)/\hbar$. If everything changes at one frequency, it means that the system is in a state of definite energy—here, the energy $(E_0 - A)$. So there is a stationary state of this energy in which the two amplitudes C_1 and C_2 are equal. We get the result that *the ammonia molecule has a definite energy $(E_0 - A)$ if there are equal amplitudes for the nitrogen atom to be “up” and to be “down.”*

There is another stationary state possible if $a = 0$; both amplitudes then have the frequency $(E_0 + A)/\hbar$. So there is another state with the definite energy $(E_0 + A)$ if the two amplitudes are equal but with the opposite sign; $C_2 = -C_1$. These are the only two states of definite energy. We will discuss the states of the ammonia molecule in more detail in the next chapter; we will mention here only a couple of things.

We conclude that *because* there is some chance that the nitrogen atom can flip from one position to the other, the energy of the molecule is not just E_0 , as we would have expected, but that there are *two* energy levels $(E_0 + A)$ and $(E_0 - A)$. Every one of the possible states of the molecule, whatever energy it has, is “split” into two levels. We say *every* one of the states because, you remember, we picked out one particular state of rotation, and internal energy, and so on. For each possible condition of that kind there is a doublet of energy levels because of the flip-flop of the molecule.

Let’s now ask the following question about an ammonia molecule. Suppose that at $t = 0$, we *know* that a molecule is in the state $|1\rangle$ or, in other words, that $C_1(0) = 1$ and $C_2(0) = 0$. What is the probability that the molecule will be found in the state $|2\rangle$ at the time t , or will still be found in state $|1\rangle$ at the time t ? Our starting condition tells us what a and b are in Eqs. (8.50) and (8.51). Letting $t = 0$, we have that

$$C_1(0) = \frac{a + b}{2} = 1, \quad C_2(0) = \frac{a - b}{2} = 0.$$

Clearly, $a = b = 1$. Putting these values into the formulas for $C_1(t)$ and $C_2(t)$ and rearranging some terms, we have

$$C_1(t) = e^{-(i/\hbar)E_0 t} \left(\frac{e^{(i/\hbar)At} + e^{-(i/\hbar)At}}{2} \right),$$

$$C_2(t) = e^{-(i/\hbar)E_0 t} \left(\frac{e^{(i/\hbar)At} - e^{-(i/\hbar)At}}{2} \right).$$

We can rewrite these as

$$C_1(t) = e^{-(i/\hbar)E_0 t} \cos \frac{At}{\hbar}, \quad (8.52)$$

$$C_2(t) = ie^{-(i/\hbar)E_0 t} \sin \frac{At}{\hbar}. \quad (8.53)$$

The two amplitudes have a magnitude that varies harmonically with time.

The probability that the molecule is found in state $|2\rangle$ at the time t is the absolute square of $C_2(t)$:

$$|C_2(t)|^2 = \sin^2 \frac{At}{\hbar}. \quad (8.54)$$

The probability starts at zero (as it should), rises to one, and then oscillates back and forth between zero and one, as shown in the curve marked P_2 of Fig. 8-2. The probability of being in the $|1\rangle$ state does not, of course, stay at one. It “dumps” into the second state until the probability of finding the molecule in the first state is zero, as shown by the curve P_1 of Fig. 8-2. The probability sloshes back and forth between the two.

A long time ago we saw what happens when we have two equal pendulums with a slight coupling. (See Chapter 49, Vol. I.) When we lift one back and let go,

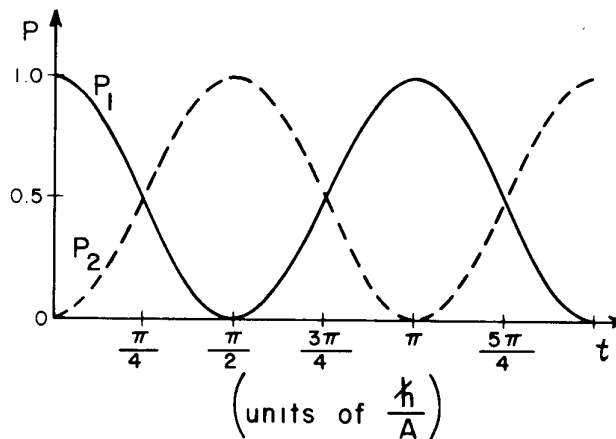


Fig. 8-2. The probability P_1 that an ammonia molecule in state $|1\rangle$ at $t = 0$ will be found in state $|1\rangle$ at t . The probability P_2 that it will be found in state $|2\rangle$.

it swings, but then gradually the other one starts to swing. Pretty soon the second pendulum has picked up all the energy. Then, the process reverses, and pendulum number one picks up the energy. It is exactly the same kind of a thing. The speed at which the energy is swapped back and forth depends on the coupling between the two pendulums—the rate at which the “oscillation” is able to leak across. Also, you remember, with the two pendulums there are two special motions—each with a definite frequency—which we call the fundamental modes. If we pull both pendulums out together, they swing together at one frequency. On the other hand, if we pull one out one way and the other out the other way, there is another stationary mode also at a definite frequency.

Well, here we have a similar situation—the ammonia molecule is mathematically like the pair of pendulums. These are the two frequencies— $(E_0 + A)/\hbar$ and $(E_0 - A)/\hbar$ —for when they are oscillating together, or oscillating opposite.

The pendulum analogy is not much deeper than the principle that the same equations have the same solutions. The linear equations for the amplitudes (8.39) are very much like the linear equations of harmonic oscillators. (In fact, this is the reason behind the success of our classical theory of the index of refraction, in which we replaced the quantum mechanical atom by a harmonic oscillator, even though, classically, this is not a reasonable view of electrons circulating about a nucleus.) If you pull the nitrogen to one side, then you get a *superposition* of these two frequencies, and you get a kind of *beat* note, because the system is *not* in one or the other states of definite frequency. The splitting of the energy levels of the ammonia molecule is, however, strictly a quantum mechanical effect.

The splitting of the energy levels of the ammonia molecule has important practical applications which we will describe in the next chapter. At long last we have an example of a practical physical problem that you can understand with the quantum mechanics!