

## Introduction and revision of quantum mechanics

### Classical and quantum definitions of *state*

In classical mechanics we ask for the *dynamical variables* of a particle (its position, momentum, kinetic energy, angular momentum). All these are functions of just the two conjugate variables of Hamiltonian mechanics, *viz.* position  $\mathbf{r}$  and momentum  $\mathbf{p}$ . So the kinetic energy is

$$T = \frac{1}{2}mv^2 = \frac{p^2}{2m}$$

where  $m$  is the mass of the particle; angular momentum is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

The dynamical variables are given *deterministically* by Newton's laws, once we provide the initial conditions.

In quantum mechanics, we ask for the same quantities of a particle. But they are more elusive to obtain. These are furnished by Schrödinger's equation which is also deterministic, again once the initial conditions are provided. However unlike Newton's laws Schrödinger's equation does not mention  $\mathbf{p}$  or  $\mathbf{r}$  as variables. Instead Schrödinger's equation is a differential equation for the complex wavefunction  $\Psi(\mathbf{r}, t)$  which is a function of  $\mathbf{r}$  and time,  $t$ ,

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi$$

and  $H$  is the Hamiltonian. If the forces on the particle are conservative, then

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V(\mathbf{r})\Psi$$

in which the first term on the right hand side is the kinetic energy (we prove this later).

Our job is to deduce the dynamical variables, in particular  $\mathbf{r}$  and  $\mathbf{p}$  from  $\Psi(\mathbf{r}, t)$ . This is especially difficult because we cannot *even in principle* know both these simultaneously to arbitrary precision. We are hampered by the uncertainty principle.

$\Psi$  represents the *state* of the particle and is sometimes called the *probability amplitude*. This is because Schrödinger's equation does not tell us exactly the position of the particle. It tells us that the probability that the particle may be found at time  $t$  in a small region of space  $d\mathbf{r}$  about  $\mathbf{r}$  is

$$\Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t) d\mathbf{r}$$

Normalisation ensures the probability that the particle is to be found somewhere in space is one,

$$1 = \int \Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t) d\mathbf{r}$$

It's important to establish that a wavefunction *can* be normalised. Generally speaking it must go rapidly to zero at infinity. Many proofs in quantum mechanics rely on this. Therefore free particles and particles subject to periodic boundary conditions such as found in solid state physics need careful handling.

Once normalisation is established at, say,  $t = 0$  the wavefunction remains normalised while being propagated according to Schrödinger's equation.

Confusion can arise over what is meant by the *state* of a particle. Einstein writes,

The attempt to conceive the quantum-theoretical description as the complete description of the individual systems leads to unnatural theoretical interpretations, which become immediately unnecessary if one accepts the interpretation that the description refers to ensembles of systems and not to individual systems.

In classical mechanics the *state* of a particle is indeed uniquely provided by its position and momentum. In quantum mechanics the state  $\Psi$  furnishes us with merely a probability. This is inescapable. The probability refers to an average value measured over a collection (ensemble) of identically prepared particles. So when you consider the probability, say, of your particle having a momentum,  $\mathbf{p}$ , you are expected to imagine a very large number of particles prepared in state  $\Psi$  each having its momentum measured at the same time. The average value of these is the momentum corresponding to the state  $\Psi$ .

It is often convenient to represent a state as a “vector” and instead of writing a wavefunction as an explicit function of  $\mathbf{r}$  and  $t$ , namely  $\Psi(\mathbf{r}, t)$ , we use a more abstract notation and represent it as the *state vector*  $|\Psi\rangle$ .

Imagine a unit vector in three dimensional space,

$$\mathbf{u} = u_1\hat{\mathbf{i}} + u_2\hat{\mathbf{j}} + u_3\hat{\mathbf{k}}$$

If you measure the vector in terms of the three cartesian coordinates, you’ll be able to determine the coefficients  $u_1$ ,  $u_2$  and  $u_3$  which describe exactly all you need to know about the vector. But imagine if you knew that the vector has some such coefficients but whenever you measured them you always found only one of these to be non zero. In other words whenever you measure the vector *you’ll find it’s one of the three cartesian basis vectors*  $\hat{\mathbf{i}}$ ,  $\hat{\mathbf{j}}$  or  $\hat{\mathbf{k}}$ . Your act of measurement has moved it from where it was onto one of the three cartesian axes. Moreover you can’t predict what measurement you’d get. In fact if you prepare many identical copies of this vector and get all your friends to measure one then they may get different answers. You will find when you pool your results that  $\hat{\mathbf{i}}$  was found with a probability  $u_1^2$ ,  $\hat{\mathbf{j}}$  was found with a probability  $u_2^2$  and  $\hat{\mathbf{k}}$  was found with a probability  $u_3^2$ . Strange, but true. That’s what happens in quantum mechanics.

The state vector of a quantum mechanical particle can always be written as a linear combination of basis functions. These basis functions are usually solutions of the time independent Schrödinger equation in the potential in which the particle is moving. It’ll be easier to follow this once we do some examples. A good example is the particle in the infinite, one dimensional square well of width  $a$ . Its basis states, or *eigenstates* analogous to the cartesian coordinates  $\hat{\mathbf{i}}$ ,  $\hat{\mathbf{j}}$  or  $\hat{\mathbf{k}}$  are<sup>†</sup>

$$\psi_n(x) = \langle x|n\rangle = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right)$$

---

<sup>†</sup> Here you should recall your level 2 notes. There you used the upper case  $\Psi(\mathbf{r}, t)$  to denote the *state* and the lower case  $\psi_n(\mathbf{r})$  to denote the eigenstate labelled with its quantum number  $n$ . We shall do the same here.

Associated with each eigenstate is the total energy (just kinetic energy in this case since  $V = 0$ ) or *eigenvalue*,

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$

The *state* of a particle in the square well may be any linear combination of eigenstates, but when we measure, say, its energy we will be certain to find it to be one of the  $E_n$  with probabilities depending upon the state it was in before measurement. The most likely outcome is not necessarily the average that we would find upon measuring a large number of replicas of the state. But in the jargon the average value is unfortunately called the “expectation value.”

### Fundamental postulates of quantum mechanics

It would be nice if quantum mechanics as a discipline could be deduced from a small number of agreed postulates. This is the case in electrodynamics, all of which follows from the principle of superposition and Coulomb’s Law; you need also to admit the principles of special relativity and Maxwell’s equations and electromagnetism follow. Likewise thermodynamics is a beautiful and powerful edifice all built on just the first and second laws. Various authors propose various sets of postulates, or axioms, from which quantum mechanics may be developed. Some are quite abstract. In his Lectures on Physics, vol III, R. P. Feynman is clear that all the mysteries of quantum mechanics can be traced back to the double slit experiment. Waves, of course interfere and produce interference patterns. Bullets, which “come in lumps” do not. Electrons come in lumps, but they do produce interference. *But* if you make a measurement of which slit a particular electron went through, the interference goes away and they behave like bullets. In quantum mechanics, what interference means is that when two or more outcomes are possible the total probability is not the sum of the individual probabilities; instead it is the square of the sum of probability amplitudes. Of course, it’s the same in optics; in an interference experiment, such as the double slit, you first add the amplitudes of the interfering waves and *only then* square them to find the intensity. Here’s what Feynman says, starting with his three postulates,

- (1) The probability of an event in an ideal experiment is given by the square of the absolute value of a complex number  $\phi$  which is called the probability amplitude:

$$P = \text{probability}$$

$$\phi = \text{probability amplitude}$$

$$P = |\phi|^2$$

- (2) When an event can occur in alternative ways, the probability amplitude for the event is the sum of the probability amplitudes for each way considered separately. There is interference:

$$\begin{aligned}\phi &= \phi_1 + \phi_2 \\ P &= |\phi_1 + \phi_2|^2\end{aligned}$$

- (3) If an experiment is performed which is capable of determining whether one or another alternative is actually taken, the probability of the event is the sum of the probabilities for each alternative. The interference is lost:

$$P = |\phi_1|^2 + |\phi_2|^2$$

One might still like to ask: “How does it work? What is the machinery behind the law?” No one has found any machinery behind the law. No one can “explain” ... No one will give you any deeper representation of the situation. We have no ideas about a more basic mechanism from which these results can be deduced.

*We would like to emphasise a very important difference between classical and quantum mechanics.* We have been talking about the probability that an electron will arrive in a given circumstance. We have implied that in our experimental arrangement (or even in the best possible one) it would be impossible to predict what would actually happen. We can only predict the odds! That would mean, if it were true, that physics has given up on the problem of trying to predict exactly what will happen in a definite circumstance. Yes! Physics has given up. *We do not know how to predict what would happen in a given circumstance,* and we believe now that it is impossible—that the only thing that can be predicted is the probability of different events. It must be recognised that this is a retrenchment in our earlier ideal of understanding nature. It may be a backward step, but no one has seen a way to avoid it.

Even more abstract and mathematical postulates are made by L. E. Ballentine in “Quantum mechanics—a modern development.” His postulates are these:

1. *To each dynamical variable there is a Hermitian operator whose eigenvalues are the possible values of the dynamical variable.*

This is consistent with the remarks above about the measurement of the kinetic energy of a particle in the square well.

2. *To each state there corresponds a unique state operator, which must be Hermitian, non negative and of unit trace.*

We cannot pursue this further here, it becomes too mathematical and at our level of quantum mechanics we can avoid dealing with the so called density operator, or density matrix, altogether. (But if you go into research in computational condensed matter theory or quantum chemistry you will run into the density matrix all the time.) Ballentine’s postulates are sufficient for him to *derive* the two most central planks (no pun intended) of quantum mechanics, namely Schrödinger’s equation, and the representation of the momentum as the differential operator

$$p_x = -i\hbar \frac{\partial}{\partial x}$$

from the Galilean transformations that apply to homogeneous space. In this way, Ballentine’s approach is the most powerful and economical.

Our purposes are best served by the less economical postulates in Cassels’ book “Basic quantum mechanics.” Here they are.

1. **Wave functions and probability density.** *The state of a particle is represented by a complex function,  $\Psi(\mathbf{r}, t)$ , such that  $|\Psi|^2 d\mathbf{r}$  is the probability of finding the particle at the time  $t$  in the element of volume  $d\mathbf{r}$  at the point  $\mathbf{r}$ .*
2. **Linear superposition of states.** *If two possible states of the particle are represented by  $\Psi_1$  and  $\Psi_2$ , then a possible state of the particle is also represented by  $\Psi = c_1\Psi_1 + c_2\Psi_2$ , where the coefficients  $c_1$  and  $c_2$  are arbitrary complex numbers, independent of  $\mathbf{r}$  but possibly functions of  $t$ .*

- 3. Observables and operators.** *Observables are represented by linear operators which multiply, differentiate, or otherwise act on the wave function to produce a new function.*
- 4. Predictions of the result of measuring an observable.** *When  $\hat{A}$  operates on a particular wavefunction,  $\psi_n$ , say, it may produce a function which differs from  $\psi_n$ , only by a constant multiplicative factor,  $A_n$ :*

$$\hat{A}\psi_n = A_n\psi_n$$

If so, then a measurement of  $A$  is certain to yield the numerical value  $A_n$ . The particle is said to be in a *eigenstate* of  $A$  belonging to the eigenvalue  $A_n$  and the wave function  $\psi_n$  is said to be an eigenfunction of  $A$  belonging to the eigenvalue  $A_n$ .

- 5. Hermitian operators.** *Only a certain kind of linear operator is suitable for representing an observable. If  $\Psi_1$  and  $\Psi_2$  are any two wavefunctions, then  $\hat{A}$  must satisfy*

$$\int \Psi_1^* (\hat{A}\Psi_2) \, d\mathbf{r} = \int (\hat{A}\Psi_1^*) \Psi_2 \, d\mathbf{r}$$

*where the integrals are taken over the space available to the particle. Operators which satisfy this relation are called *Hermitian*.*

- 6. Momentum and energy operators.** *The operator which represents momentum is assumed to be*

$$\hat{\mathbf{p}} = -i\hbar\nabla$$

In cartesian coordinates,

$$\begin{aligned}\hat{p}_x &= -i\hbar \frac{\partial}{\partial x} \\ \hat{p}_y &= -i\hbar \frac{\partial}{\partial y} \\ \hat{p}_z &= -i\hbar \frac{\partial}{\partial z}\end{aligned}$$

If the kinetic energy is  $p^2/2m$  then obviously the *kinetic energy operator* is

$$\hat{T} = -\frac{\hbar^2}{2m} \nabla^2$$

In one dimension

$$\hat{T} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$$

hence the form of the Hamiltonian given at the beginning of these notes. In fact, this postulate is redundant. The operator form of  $\mathbf{p}$  may be deduced from Schrödinger's equation by finding the time derivative of the average value of  $\mathbf{r}$  (see Griffiths, section 1.5).

### 7. Time dependence of $\Psi$ . The rate of change of $\Psi$ is given by

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi$$

### Stationary states

Generally speaking there are two steps involved in solving a quantum mechanical problem once a Hamiltonian is given (basically once the potential is known in which the particle is moving). First Schrödinger's equation is solved subject to the boundary conditions to find the *state* as a function of position and time. Secondly one finds the *average value* of the operator one is interested in. Any *individual* measurement of an observable, typically position, momentum or energy, will always result in one of the eigenvalues of its operator  $\hat{\mathbf{r}}$ ,  $\hat{\mathbf{p}}$  or  $\hat{H}$ . And if two operators do not commute (for example  $\hat{\mathbf{r}}$  and  $\hat{\mathbf{p}}$ ) then we cannot determine the eigenvalues of both simultaneously. On the other hand measurement of an observable over a large number of identically prepared particles will result in an average value ("expectation value"). For example a measurement of the position will yield an average value,

$$\begin{aligned}\langle \mathbf{r}(t) \rangle &= \int \Psi^*(\mathbf{r}, t) \mathbf{r} \Psi(\mathbf{r}, t) \, d\mathbf{r} \\ &= \langle \Psi | \mathbf{r} | \Psi \rangle\end{aligned}$$

The second line is an alternative way of writing it, in terms of the state vector  $|\Psi\rangle$  and its hermitian conjugate  $\langle\Psi|$ .

The dynamical variables  $\mathbf{r}$  and  $\mathbf{p}$  are called “conjugate” in classical Hamiltonian mechanics. In quantum mechanics their  $x$ ,  $y$  and  $z$  components do not commute. You can easily prove that using

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$$

$$\hat{x}\hat{p}_x - \hat{p}_x\hat{x} = [\hat{x}, \hat{p}_x] = i\hbar$$

(You do this by introducing some arbitrary function of  $x$  and employing the chain rule—please be sure you can do it).

Niels Bohr called the pair of variables whose operators do not commute *complementary* (see Cassels, page 32; Schiff, page 8). Another set of non commuting observables that we will discuss in detail in the lectures are the three cartesian components of the angular momentum vector. One simply cannot know the  $x$ ,  $y$  and  $z$  components of an electron’s spin. I want to quote here some paragraphs from Griffiths even though we have not yet covered the details of spin and its eigenvalues. You may come back to it again during the lectures on angular momentum. For now what you need to know is that any of the three components of spin of a spin- $\frac{1}{2}$  particle can take on one of only two eigenvalues, namely  $\pm\frac{1}{2}\hbar$ . Since only one can be known at any time because their operators,  $\hat{S}_x$ ,  $\hat{S}_y$  and  $\hat{S}_z$  do not commute, by convention we choose the  $z$ -component as special (it’s simply a matter of how we choose our cartesian coordinate system) and we write its two eigenvalue equations as

$$\hat{S}_z\chi_+ = \frac{1}{2}\hbar ; \quad \hat{S}_z\chi_- = -\frac{1}{2}\hbar$$

in which  $\chi_+$  and  $\chi_-$  are the two eigenvectors; sometimes they’re written simply  $|+\rangle$ , and  $|-\rangle$ , or  $|\uparrow\rangle$ , and  $|\downarrow\rangle$  to denote up and down spin. The operator corresponding to the  $x$ -component of spin has eigenvectors denoted by superscripts,  $\chi_+^{(x)}$  and  $\chi_-^{(x)}$ , and the same eigenvalues as  $S_z$ ,

I’d like now to walk you through an imaginary measurement scenario involving spin  $\frac{1}{2}$ , because it serves to illustrate in very concrete terms some of the abstract ideas we discussed back in Chapter 1. Let’s say we start out with a particle in the state  $\chi_+$ . If someone asks, “What is the  $z$ -component of that particular particle’s spin angular momentum?”, we could answer unambiguously  $+\frac{1}{2}\hbar$ . For a measurement of  $S_z$  is certain to return that value. But if our interrogator asks instead, “What is the  $x$ -component of that particle’s spin angular momentum?” we are obliged to equivocate: If you measure  $S_x$ , the chances are fifty-fifty of getting either  $\frac{1}{2}\hbar$  or  $-\frac{1}{2}\hbar$ . If the questioner is a classical physicist, or a “realist” (in the sense of Section 1.2), he will regard this as an inadequate—not to say impertinent—response: “Are you telling me that you *don’t know* the true state of the particle?” On the contrary, I know *precisely* what the state of the particle is:  $\chi_+$ . “Well then, how come you can’t tell me what the  $x$ -component of the spin is?” Because it simply *does not have* a particular  $x$ -component of spin. Indeed it cannot, for if both  $S_x$  and  $S_z$  were well-defined, the uncertainty principle would be violated.

At this point our challenger grabs the test-tube and *measures* the  $x$ -component of its spin: let's say he gets the value  $+\frac{1}{2}\hbar$ . "Aha!" (he shouts in triumph), "You *lied!* This particle has a perfectly well-defined value of  $S_x$ :  $\frac{1}{2}\hbar$ ." Well, sure—it does *now*, but that doesn't prove it *had* that value, prior to your measurement. "You have obviously been reduced to splitting hairs. And anyway what happened to your uncertainty principle? I now know both  $S_x$  and  $S_z$ ." I'm sorry, but you do *not*: In the course of your measurement, you altered the particle's state; it's now in the state  $\chi_+^{(x)}$ , and whereas you know the value of  $S_x$ , you no longer know the value of  $S_z$ . "But I was extremely careful not to disturb the particle when I measured  $S_x$ ." Very well, if you don't believe me, *check it out*: Measure  $S_z$ , and see what you get. (Of course he *may* get  $+\frac{1}{2}\hbar$ , which will be embarrassing to my case—but if we repeat this whole scenario over and over, half the time he will get  $-\frac{1}{2}\hbar$ .)

To the layman, the philosopher, or the classical physicist, a statement of the form "this particle doesn't have a well-defined position" (or momentum, or  $x$ -component of spin angular momentum, or whatever) sounds vague, incompetent or (worst of all) profound. It is none of these. But its precise meaning is, I think, almost impossible to convey to anyone who has not studied quantum mechanics in some depth. If you find your own comprehension slipping, from time to time (if you *don't*, you probably haven't understood the problem), come back to the spin- $\frac{1}{2}$  system: It is the simplest and cleanest context for thinking through the conceptual paradoxes of quantum mechanics.

*Stationary states* are those for which average values of observables do not depend on time. They are best discovered by first solving the time independent Schrödinger equation which will yield a set of energy eigenstates,  $\psi_n$ , with eigenvalues,  $E_n$ . An example, again, is the square well, or the simple harmonic oscillator. As long as the Hamiltonian *does not mention time* then a general stationary state is

$$\Psi(x, t) = \psi_n(x) e^{-iE_n t/\hbar}$$

which follows from the time dependent Schrödinger equation by separation of variables. This looks at first if it depends on time (as indeed it does through its phase factor). But when you form  $|\Psi|^2$  the result is independent of time because the exponential times its complex conjugate is one.

### **Coherent superpositions of stationary states—get the particle moving!**

The simplest example of a *non* stationary state is a superposition of two stationary states. Suppose a particle is in a potential (it could for example be the square well potential) which has eigenstates  $\psi_n(x)$  and eigenvalues  $E_n$ . We could prepare it in a linear combination of

two of these, say the ground state and first excited state, at time  $t = 0$ ,

$$\Psi(x, 0) = c_1\psi_1(x) + c_2\psi_2(x)$$

This is called a “coherent superposition” because it will exhibit interference in the sense of Feynman’s postulates (above). As long as the hamiltonian is not itself dependent on time we know that each of these eigenstates is multiplied by a time dependent phase factor to make a stationary state. In the same way the time dependence of our state is such that at at time  $t$ ,

$$\Psi(x, t) = c_1\psi_1(x)e^{-iE_1t/\hbar} + c_2\psi_2(x)e^{-iE_2t/\hbar}$$

Now we ask what is the probability of the particle being found at position  $x$  and time  $t$ . This is

$$\begin{aligned} |\Psi(x, t)|^2 &= \left( c_1\psi_1(x)e^{iE_1t/\hbar} + c_2\psi_2(x)e^{iE_2t/\hbar} \right) \left( c_1\psi_1(x)e^{-iE_1t/\hbar} + c_2\psi_2(x)e^{-iE_2t/\hbar} \right) \\ &= c_1^2\psi_1^2 + c_2^2\psi_2^2 + 2c_1c_2\psi_1\psi_2 \cos(E_2 - E_1)t/\hbar \end{aligned}$$

Note how the first parenthesis is the complex conjugate of the second as required when you take the square of the absolute value of a complex number. Compare this with the Feynman postulates. The probability is the sum of the uncorrelated probabilities plus the third term which describes *interference*. Now the particle is moving sinusoidally in time; it is no longer stationary.

If at some time  $t$  the energy of the particle is measured what will you find? Are you surprised that you will get *either*  $E_1$  with probability  $c_1^2$  *or*  $E_2$  with probability  $c_2^2$ ? No other result is possible. The measurement has destroyed the coherence; we’d better stop here—we’re getting into quantum information theory. But does it worry you that the measurement apparently violates the conservation of energy? Then remember Einstein’s statement on page 2. If I have prepared many replicas of the same system then the *average* over all my measurements will be  $c_1^2E_1 + c_2^2E_2$ , which is the same as the expectation value of the energy of the state at  $t = 0$ .