

Chapter 1

Key Features of Quantum Mechanics

©B. Zwiebach

Quantum mechanics is now almost one-hundred years old, but we are still discovering some of its surprising features and it remains the subject of much investigation and speculation. The framework of quantum mechanics is a rich and elegant extension of the framework of classical physics. It is also counterintuitive and seems, at first, almost paradoxical.

Quantum physics has replaced classical physics as the correct fundamental description of our physical universe. It is used routinely to describe most phenomena that occur at short distances. Quantum physics is the result of applying the framework of quantum mechanics to different physical phenomena. We thus have Quantum Electrodynamics, when quantum mechanics is applied to electromagnetism, Quantum Optics, when it is applied to light and optical devices, or Quantum Gravity, when it is applied to gravitation. Quantum mechanics indeed provides a remarkably coherent and elegant framework. The era of quantum physics begins in earnest in 1925, with the discoveries of Schrödinger and Heisenberg. The seeds for these discoveries were planted by Planck, Einstein, Bohr, de Broglie, and others. It is a tribute to human imagination that we have been able to discover the counterintuitive and abstract set of rules that define quantum mechanics. Here we aim to explain and provide some perspective on the main features of this framework.

We will begin by discussing the property of linearity, which quantum mechanics shares with electromagnetic theory. This property tells us what kind of theory quantum mechanics is and why, it could be argued, it is simpler than classical mechanics. We then turn to photons, the particles of light. We use photons and polarizers to explain why quantum physics is not deterministic and, in contrast with classical physics, the results of some experiments cannot be predicted. Quantum mechanics is a framework in which we can only predict the *probabilities* for the various outcomes of any given experiment. Our next subject is quantum superpositions, in which a quantum object somehow manages to exist simultaneously in two mutually incompatible states. A quantum light-bulb, for example,

could be in a state in which it is both on and off at the same time!

1.1 Linearity of the equations of motion

In physics a theory is usually described by a set of equations for some quantities called the **dynamical variables** of the theory. After writing a theory, the most important task is finding solutions of the equations. A solution of the equations describes a possible reality, according to the theory. Because an expanding universe is a solution of Albert Einstein's gravitational equations, for example, it follows that an expanding universe is possible, according to this theory. A single theory may have many solutions, each describing a possible reality.

There are linear theories and nonlinear theories. Nonlinear theories are more complicated than linear theories. In a linear theory a remarkable fact takes place: if you have two solutions you obtain a third solution of the theory simply by adding the two solutions. An example of a beautiful linear theory is Maxwell's theory of electromagnetism, a theory that governs the behavior of electric and magnetic fields. A field, as you probably know, is a quantity whose values may depend on position and on time. A simple solution of this theory describes an electromagnetic wave propagating in a given direction. Another simple solution could describe an electromagnetic wave propagating in a different direction. Because the theory is linear, having the two waves propagating simultaneously, each in its own direction and without affecting each other, is a new and consistent solution. The sum is a solution in the sense that the electric field in the new solution is the sum of the electric field in the first solution plus the electric field in the second solution. The same goes for the magnetic field: the magnetic field in the new solution is the sum of the magnetic field in the first solution plus the magnetic field in the second solution. In fact you can add any number of solutions to still find a solution. Even if this sounds esoteric, you are totally familiar with it. The air around you is full of electromagnetic waves, each one propagating oblivious to the other ones. There are the waves of thousands of cell phones, the waves carrying hundreds of wireless internet messages, the waves from a plethora of radio-stations, TV stations, and many, many more. Today, a single transatlantic cable can carry simultaneously millions of telephone calls, together with huge amounts video and internet data. All of that courtesy of linearity.

More concretely, we say that Maxwell's equations are **linear** equations. A solution of Maxwell's equation is described by an electric field \mathbf{E} a magnetic field \mathbf{B} , a charge density ρ and a current density \mathbf{J} , all collectively denoted as $(\mathbf{E}, \mathbf{B}, \rho, \mathbf{J})$. This collection of fields and sources satisfy Maxwell's equations. Linearity implies that if $(\mathbf{E}, \mathbf{B}, \rho, \mathbf{J})$ is a solution so is $(\alpha\mathbf{E}, \alpha\mathbf{B}, \alpha\rho, \alpha\mathbf{J})$, where all fields and sources have been multiplied by the constant α . Given two solutions

$$(\mathbf{E}_1, \mathbf{B}_1, \rho_1, \mathbf{J}_1), \quad \text{and} \quad (\mathbf{E}_2, \mathbf{B}_2, \rho_2, \mathbf{J}_2), \quad (1.1.1)$$

linearity also implies that we can obtain a new solution by adding them

$$(\mathbf{E}_1 + \mathbf{E}_2, \mathbf{B}_1 + \mathbf{B}_2, \rho_1 + \rho_2, \mathbf{J}_1 + \mathbf{J}_2). \quad (1.1.2)$$

The new solution may be called the superposition of the two original solutions.

We will consider equations for which the solutions satisfy the following conditions: (i) given a solution, any multiple of the solution is also a solution and, (ii) given two solutions the sum of solutions is also a solution. We will call these **linear equations**,¹ and they take the general form

$$L u = 0, \quad (1.1.3)$$

where u denotes the unknown. The unknown may be a number, a function of time, a function of space, a function of spacetime, essentially anything unknown! In fact, u could represent a collection (u_1, u_2, \dots) of unknowns. The symbol L denotes a **linear operator**, an object that acts on u and satisfies the following two properties:

$$L(au) = aLu, \quad L(u_1 + u_2) = Lu_1 + Lu_2, \quad (1.1.4)$$

with a is a number. When L is a linear operator the equation $Lu = 0$ is a linear equation that satisfies the two conditions listed above. Indeed, $Lu = 0$ implies $L(au) = 0$ using the first equation above, and $Lu_1 = Lu_2 = 0$ implies $L(u_1 + u_2) = 0$ using the second equation above. Note that the linearity of L implies that

$$L(\alpha u_1 + \beta u_2) = \alpha Lu_1 + \beta Lu_2. \quad (1.1.5)$$

showing that if u_1 is a solution ($Lu_1 = 0$) and u_2 is a solution ($Lu_2 = 0$) then $\alpha u_1 + \beta u_2$ is also a solution. We call $\alpha u_1 + \beta u_2$ the **general superposition** of the solutions u_1 and u_2 .

Example. Consider the equation

$$\frac{du}{dt} + \frac{1}{\tau} u = 0, \quad (1.1.6)$$

where τ is a constant with units of time. This differential equation can be cast in the form $Lu = 0$ if we define the operator L to act as follows

$$L u \equiv \frac{du}{dt} + \frac{1}{\tau} u. \quad (1.1.7)$$

One sometimes writes the operator L as

$$L \equiv \frac{d}{dt} + \frac{1}{\tau}, \quad (1.1.8)$$

where it is implicit that the derivative acts on whatever is to the right of L and the constant term $1/\tau$ acts by multiplication. We quickly verify that L is a linear operator. For any arbitrary constant a we have

$$L(au) = \frac{d(au)}{dt} + \frac{1}{\tau} au = a \left(\frac{du}{dt} + \frac{1}{\tau} u \right) = a Lu. \quad (1.1.9)$$

¹Conventional terminology calls equations of the type $Lu = f$, with f independent of u , linear *inhomogeneous* equations. Here we reserve the term ‘linear equations’ for equations in which solutions can be superposed.

Moreover, for two functions u_1 and u_2 we also have

$$\begin{aligned} L(u_1 + u_2) &= \frac{d(u_1 + u_2)}{dt} + \frac{1}{\tau}(u_1 + u_2) = \frac{du_1}{dt} + \frac{du_2}{dt} + \frac{1}{\tau}u_1 + \frac{1}{\tau}u_2 \\ &= \frac{du_1}{dt} + \frac{1}{\tau}u_1 + \frac{du_2}{dt} + \frac{1}{\tau}u_2 = Lu_1 + Lu_2. \end{aligned} \quad (1.1.10)$$

This shows that the differential equation is actually linear. With a little practice, linear equations are easily recognized by inspection.

Einstein's theory of general relativity is a nonlinear theory whose dynamical variable is a gravitational field, the field that describes, for example, how planets move around a star. Being a nonlinear theory, you simply cannot add the gravitational fields of different solutions to find a new solution. This makes Einstein's theory rather complicated, by all accounts much more complicated than Maxwell theory. In fact, classical mechanics, as invented mostly by Isaac Newton, is also a nonlinear theory! In classical mechanics the dynamical variables are positions and velocities of particles, acted by forces. There is no general way to use two solutions to build a third.

Indeed, consider the equation of motion for a particle on a line under the influence of a time-independent potential $V(x)$, which is in general an arbitrary function of x . The dynamical variable in this problem is $x(t)$, the position as a function of time. Letting V' denote the derivative of V with respect to its argument, Newton's second law takes the form

$$m \frac{d^2 x(t)}{dt^2} = -V'(x(t)). \quad (1.1.11)$$

The left-hand side is the mass times acceleration and the right hand side is the force experienced by the particle in the potential. It is probably worth emphasizing that the right hand side is the function $V'(x)$ evaluated for x set equal to $x(t)$:

$$V'(x(t)) \equiv \left. \frac{\partial V(x)}{\partial x} \right|_{x=x(t)}. \quad (1.1.12)$$

While we could have used an ordinary derivative here, we wrote a partial derivative as is commonly done for the general case of time dependent potentials. The reason Newton's second law (1.1.11) is not a linear equation is that the function $V'(x)$ is not linear. In general, for arbitrary functions u and v we expect

$$V'(ax) \neq aV'(x), \quad \text{and} \quad V'(x_1 + x_2) \neq V'(x_1) + V'(x_2). \quad (1.1.13)$$

As a result given a solution $x(t)$, the scaled solution $\alpha x(t)$ is not expected to be a solution. Given two solutions $x_1(t)$ and $x_2(t)$ then $x_1(t) + x_2(t)$ is not guaranteed to be a solution either.

Exercise. What is the most general potential $V(x)$ for which the equation of motion for $x(t)$ is linear?

Quantum mechanics is a linear theory. The signature equation in this theory, the so-called Schrödinger equation is a linear equation for a quantity called the **wavefunction** and it determines its time evolution. The wavefunction is the dynamical variable in quantum mechanics but, curiously, its physical interpretation was not clear to Erwin Schrödinger when he wrote his equation in 1925. It was Max Born, who months later suggested that the wavefunction encodes probabilities. This was the correct physical interpretation, but it was thoroughly disliked by many, including Schrödinger, who remained unhappy about it for the rest of his life. The linearity of quantum mechanics implies a profound simplicity. In some sense quantum mechanics is simpler than classical mechanics. In quantum mechanics solutions can be added to form new solutions.

The wavefunction Ψ depends on time and may also depend on space. The Schrödinger equation is a partial differential equation that takes the form

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi, \quad (1.1.14)$$

where i is the imaginary unit ($i = \sqrt{-1}$), \hbar is Planck's constant, and \hat{H} is the Hamiltonian or energy operator. The wavefunction is a complex number, as we will elaborate below. The Hamiltonian \hat{H} is a linear operator that can act on wavefunctions:

$$\hat{H}(a\Psi) = a\hat{H}\Psi, \quad \hat{H}(\Psi_1 + \Psi_2) = \hat{H}(\Psi_1) + \hat{H}(\Psi_2), \quad (1.1.15)$$

with a a constant that in fact need not be real; it can be a complex number. Of course, \hat{H} itself does not depend on the wavefunction! To check that the Schrödinger equation is linear we cast it in the form $L\Psi = 0$ with L defined as

$$L\Psi \equiv i\hbar \frac{\partial \Psi}{\partial t} - \hat{H}\Psi. \quad (1.1.16)$$

It is now a simple matter to verify that L is a linear operator. Physically this means that if Ψ_1 and Ψ_2 are solutions to the Schrödinger equation, then so is the superposition $\alpha\Psi_1 + \beta\Psi_2$, where α and β are both complex numbers.

1.2 Complex numbers are essential

Quantum mechanics is the first physics theory that truly requires the use of *complex* numbers. The numbers most of us use for daily life (integers, fractions, decimals) are *real* numbers. The set of complex numbers is denoted by \mathbb{C} and the set of real numbers is denoted by \mathbb{R} . Complex numbers appear when we combine real numbers with the imaginary unit i , defined to be equal to the square root of minus one: $i \equiv \sqrt{-1}$. Being the square root of minus one, it means that i squared must give minus one: $i^2 = -1$. Complex numbers are fundamental in mathematics. An equation like $x^2 = -4$, for an unknown x cannot be solved if x has to be real. No real number squared gives you minus one. But if we allow for complex numbers, we have the solutions $x = \pm 2i$. Mathematicians have shown that all polynomial equations can be solved in terms of complex numbers.

A complex number z , in all generality, is a number of the form

$$z = a + ib \in \mathbb{C}, \quad a, b \in \mathbb{R}. \quad (1.2.1)$$

Here a and b are real numbers, and ib denotes the product of i with b . The number a is called the real part of z and b is called the imaginary part of z :

$$\operatorname{Re} z = a, \quad \operatorname{Im} z = b. \quad (1.2.2)$$

The complex conjugate z^* of z is defined by

$$z^* = a - ib. \quad (1.2.3)$$

You can quickly verify that a complex number z is real if $z^* = z$, and it is purely imaginary if $z^* = -z$. For any complex number $z = a + ib$ one can define the *norm* $|z|$ of the complex number to be a *positive, real* number given by

$$|z| = \sqrt{a^2 + b^2}. \quad (1.2.4)$$

You can quickly check that

$$|z|^2 = zz^*. \quad (1.2.5)$$

Complex numbers are represented as vectors in a two dimensional “complex plane”. The real part of the complex number is the x component of the vector and the imaginary part of the complex number is the y component. If you consider the unit length vector in the complex plane making an angle θ with the x axis has x component $\cos \theta$ and y component $\sin \theta$. The vector is therefore the complex number $\cos \theta + i \sin \theta$. Euler’s identity relates this to the exponential of $i\theta$:

$$e^{i\theta} = \cos \theta + i \sin \theta. \quad (1.2.6)$$

A complex number of the form $e^{i\chi}$, with χ real is called a *pure phase*.

While complex numbers are sometimes useful in classical mechanics or Maxwell theory, they are not strictly needed. None of the dynamical variables, which correspond to measurable quantities, is a complex number. In fact, complex numbers can’t be measured at all: all measurements in physics result in real numbers. In quantum mechanics, however, complex numbers are fundamental. The Schrödinger equation involves complex numbers. Even more, the wavefunction, the dynamical variable of quantum mechanics it itself a complex number:

$$\Psi \in \mathbb{C}. \quad (1.2.7)$$

Since complex numbers cannot be directly measured the relation between the wavefunction and a measurable quantity must be somewhat indirect. Born’s idea to identify probabilities, which are always positive real numbers, with the square of the norm of the wavefunction was very natural. If we write the wavefunction of our quantum system as Ψ , the probabilities for possible events are computed from $|\Psi|^2$. The mathematical framework required to express the laws of quantum mechanics consists of complex vector spaces. In any vector space we

have objects called vectors that can be added together. In a complex vector space a vector multiplied by a complex number is still a vector. As we will see in our study of quantum mechanics it is often useful to think of the wavefunction Ψ as a vector in some complex vector space.

1.3 Loss of determinism

Maxwell's crowning achievement was the realization that his equations of electromagnetism allowed for the existence of propagating waves. In particular, in 1865 he conjectured that light was an electromagnetic wave, a propagating fluctuation of electric and magnetic fields. He was proven right in subsequent experiments. Towards the end of the nineteenth century physicists were convinced that light was a wave. The certainty, however, did not last too long. Experiments on blackbody radiation and on the photo-emission of electrons suggested that the behavior of light had to be more complicated than that of a simple wave. Max Planck and Albert Einstein were the most prominent contributors to the resolution of the puzzles raised by those experiments.

In order to explain the features of the photoelectric effect, Einstein postulated (1905) that in a light beam the energy comes in quanta – the beam is composed of packets of energy. Einstein essentially implied that light was made up of particles, each carrying a fixed amount of energy. He himself found this idea disturbing, convinced like most other contemporaries that, as Maxwell had shown, light was a wave. He anticipated that a physical entity, like light, that could behave both as a particle and as a wave could bring about the demise of classical physics and would require a completely new physical theory. He was in fact right. Though he never quite liked quantum mechanics, his ideas about particles of light, later given the name *photons*, helped construct this theory.

It took physicists until 1925 to accept that light could behave like a particle. The experiments of Arthur Compton (1923) eventually convinced most skeptics. Nowadays, particles of light, or photons, are routinely manipulated in laboratories around the world. Even if mysterious, we have grown accustomed to them. Each photon of visible light carries very little energy – a small laser pulse can contain many billions of photons. Our eye, however, is a very good photon detector: in total darkness, we are able to see light when as little as ten photons hit upon our retina. When we say that light behaves like a particle we mean a quantum mechanical particle: a packet of energy and momentum that is not composed of smaller packets. We *do not* mean a classical point particle or Newtonian corpuscle, which is a zero-size object with definite position and velocity.

As it turns out, the energy of a photon depends only on the color of the light. As Einstein discovered the energy E and frequency ν for a photon are related by

$$E = h\nu, \tag{1.3.1}$$

where h is Planck's constant and is equal to $2\pi\hbar$, where \hbar is the modified version of the constant that enters the Schrödinger equation. The frequency of a photon determines the wavelength λ of the light through the relation $\nu\lambda = c$, where c is the speed of light. All

green photons, for example, have the same energy. To increase the energy in a light beam while keeping the same color, one simply needs more photons.

As we now explain, the existence of photons implies that Quantum Mechanics is not deterministic. By this we mean that the result of an experiment cannot be determined, as it would in classical physics, by the conditions that are under the control of the experimenter.

Consider a polarizer whose preferential direction is aligned along the \hat{x} direction, as shown in Figure 1.1. Light that is linearly polarized along the \hat{x} direction namely, light whose electric field points in this direction, goes through the polarizer. If the incident light polarization is orthogonal to the \hat{x} direction the light will not go through at all. Thus light linearly polarized in the \hat{y} direction will be totally absorbed by the polarizer. Now consider light polarized along a direction forming an angle α with the x -axis, as shown in Figure 1.2. What happens?

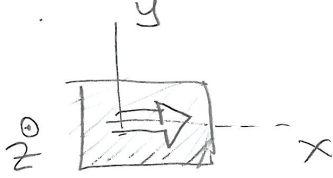


Figure 1.1: A polarizer that transmits light linearly polarized along the \hat{x} direction.

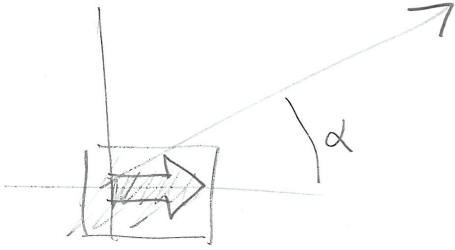


Figure 1.2: Light linearly polarized along the direction at an angle α hitting the polarizer.

Thinking of the light as a propagating wave, the incident electric field \mathbf{E}_α makes an angle α with the x -axis and therefore takes the form

$$\mathbf{E}_\alpha = E_0 \cos \alpha \hat{x} + E_0 \sin \alpha \hat{y}. \quad (1.3.2)$$

This is an electric field of magnitude E_0 . In here we are ignoring the time and space dependence of the wave; they are not relevant to our discussion. When this electric field hits the polarizer, the component along \hat{x} goes through and the component along \hat{y} is absorbed. Thus

$$\text{Beyond the polarizer:} \quad \mathbf{E} = E_0 \cos \alpha \hat{x}. \quad (1.3.3)$$

You probably recall that the energy in an electromagnetic wave is proportional to the square of the magnitude of the electric field. This means that the fraction of the beam's energy that goes through the polarizer is $(\cos \alpha)^2$. It is also well known that the light emerging from the polarizer has the same *frequency* as the incident light.

So far so good. But now, let us try to understand this result by thinking about the photons that make up the incident light. The premise here is that all photons in the incident beam are identical. Moreover the photons do not interact with each other. We could even imagine sending the whole energy of the incident light beam one photon at a time. Since all the light that emerges from the polarizer has the same frequency as the incident light, and thus the same frequency, we must conclude that each individual photon either goes through or is absorbed. If a fraction of a photon went through it would be a photon of lower energy and thus lower frequency, which is something that does not happen.

But now we have a problem. As we know from the wave analysis, a fraction $(\cos \alpha)^2$ of the photons must go through, since that is the fraction of the energy that is transmitted. Consequently a fraction $1 - (\cos \alpha)^2$ of the photons must be absorbed. But if all the photons are identical, why is it that what happens to one photon does not happen to all of them?

The answer in quantum mechanics is that there is indeed a loss of determinism. No one can predict if a photon will go through or will get absorbed. The best anyone can do is to predict probabilities. In this case there would be a probability $(\cos \alpha)^2$ of going through and a probability $1 - (\cos \alpha)^2$ of failing to go through.

Two escape routes suggest themselves. Perhaps the polarizer is not really a homogeneous object and depending exactly on where the photon hits it either gets absorbed or goes through. Experiments show this is not the case. A more intriguing possibility was suggested by Einstein and others. A possible way out, they claimed, was the existence of *hidden variables*. The photons, while apparently identical, would have other *hidden* properties, not currently understood, that would determine with certainty which photon goes through and which photon gets absorbed. Hidden variable theories would seem to be untestable, but surprisingly they can be tested. Through the work of John Bell and others, physicists have devised clever experiments that rule out most versions of hidden variable theories. No one has figured out how to restore determinism to quantum mechanics. It seems to be an impossible task.

When we try to describe photons quantum mechanically we could use wavefunctions, or equivalently the language of states. A photon polarized along the $\hat{\mathbf{x}}$ direction is not represented using an electric field, but rather we just give a name for its *state*:

$$|\text{photon}; x\rangle. \quad (1.3.4)$$

We will learn the rules needed to manipulate such objects, but for the time being you could think of it like a vector in some space yet to be defined. Another state of a photon, or vector is

$$|\text{photon}; y\rangle, \quad (1.3.5)$$

representing a photon polarized along $\hat{\mathbf{y}}$. These states are the wavefunctions that represent the photon. We now claim that the photons in the beam that is polarized along the direction

α are in a state $|\text{photon}; \alpha\rangle$ that can be written as a superposition of the above two states:

$$|\text{photon}; \alpha\rangle = |\text{photon}; x\rangle \cos \alpha + |\text{photon}; y\rangle \sin \alpha . \quad (1.3.6)$$

This equation should be compared with the decomposition (1.3.2) of the electric field. While there are some similarities –both are superpositions– one refers to electric fields and the other to states of a single photon. Any photon that emerges from the polarizer will necessarily be polarized in the \hat{x} direction and therefore it will be in the state

$$\text{Beyond the polarizer: } |\text{photon}; x\rangle . \quad (1.3.7)$$

This can be compared with the electric field (1.3.3) beyond the polarizer. This field, with the factor $\cos \alpha$ carries information about the amplitude of the wave. Here, for a single photon, there is no room for such a factor.

In the famous Fifth Solvay International Conference of 1927 the world’s most notable physicists gathered to discuss the newly formulated quantum theory. Seventeen out of the twenty nine attendees were or became Nobel Prize winners. Einstein, unhappy with the uncertainty in quantum mechanics stated the nowadays famous quote: “God does not play dice”, to which Niels Bohr is said to have answered: “Einstein, stop telling God what to do.” Bohr was willing to accept the loss of determinism, Einstein was not.

1.4 Quantum superpositions

We have already discussed the concept of linearity; the idea that the sum of two solutions representing physical realities represents a new, allowed, physical reality. This superposition of solutions has a straightforward meaning in classical physics. In the case of electromagnetism, for example, if we have two solutions, each with its own electric and magnetic field, the sum solution is simply understood: its electric field is the sum of the electric fields of the two solutions and its magnetic field is the sum of the magnetic fields of the two solutions. In quantum mechanics, as we have explained, linearity holds. The interpretation of a superposition, however, is very surprising.

One interesting example is provided by a Mach-Zehnder interferometer; an arrangement of beam splitters, mirrors, and detectors used by Ernst Mach and Ludwig Zehnder in the 1890’s to study interference between two beams of light.

A beam splitter, as its name indicates, splits an incident beam into two beams, one that is reflected from the splitter and one that goes through the splitter. Our beam splitters will be balanced: they split a given beam into two beams of equal intensity (Figure 1.3). The light that bounces off is called the reflected beam, the light that goes through is called the transmitted beam. The incident beam can hit the beam splitter from the top or from the bottom.

The Mach-Zehnder configuration, shown in Figure 1.4, has a left beam splitter (BS1) and a right beam splitter (BS2). In between we have the two perfect mirrors, M1 on the top and M2 on the bottom. An incoming beam from the left is split by BS1 into two beams,

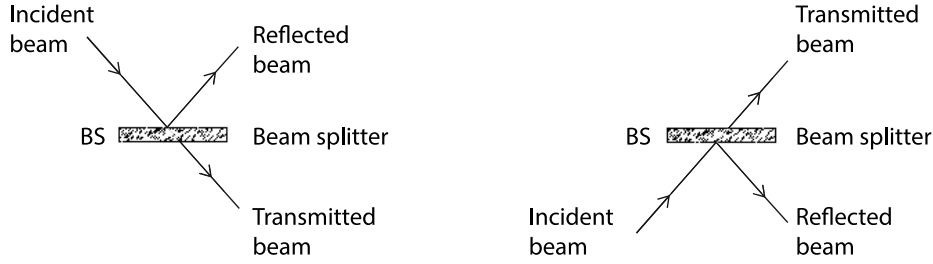


Figure 1.3: An incident beam hitting a beam-splitter results in a reflected beam and a transmitted beam. Left: incident beam coming from the top. Right: incident beam coming from the bottom.

each of which hits a mirror and is then sent into BS2. At BS2 the beams are recombined and sent into two outgoing beams that go into photon detectors D0 and D1.

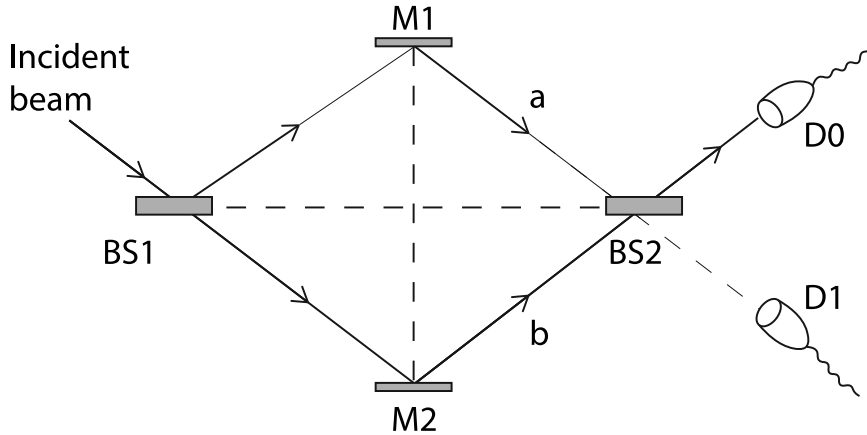


Figure 1.4: A Mach-Zehnder interferometer consists of two beam splitters BS1 and BS2, two mirrors M1 and M2, and two detectors D0 and D1. An incident beam will be split into two beams by BS1. One beam goes through the upper branch, which contains M1, the other beam goes through the lower branch, which contains M2. The beams on the two branches recombine at BS2 and are then sent into the detectors. The configuration is prepared to produce an interference so that all incident photons end at the detector D0, with none at D1.

It is relatively simple to arrange the beam splitters so that the incoming beam, upon splitting at BS1 and recombination at BS2 emerges in the top beam which goes into D0. In this arrangement no light at all goes into D1. This requires a precise interference effect at BS2. Note that we have two beams incident upon BS2; the top beam is called *a* and the lower beam is called *b*. Two contributions go towards D0: the reflection of *a* at BS2 and the transmission from *b* at BS2. These two contributions interfere constructively to give a beam going into D0. Two contributions also go towards D1: the transmission from *a* at BS2 and the reflection from *b* at BS2. These two can indeed be arranged to interfere destructively to give no beam going into D1.

It is instructive to think of the incoming beam as a sequence of photons that we send into the interferometer, one photon at a time. This shows that, at the level of photons, the interference is not interference of one photon with another photon. Each photon must interfere with *itself* to give the result. Indeed interference between two photons is not possible: destructive interference, for example, would require that two photons end up giving no photon, which is impossible by energy conservation.

Therefore, each photon does the very strange thing of going through both branches of the interferometer! Each photon is in a superposition of two states: a state in which the photon is in the top beam or upper branch, added to a state in which the photon is in the bottom beam or lower branch. Thus the state of the photon in the interferometer is a funny state in which the photon seems to be doing two incompatible things at the same time.

Equation (1.3.6) for a photon state is another example of a quantum superposition. The photon state has a component along an x -polarized photon and a component along a y -polarized photon.

When we speak of a wavefunction, we also loosely call it a state, because the wavefunction specifies the state of our quantum system. We also sometimes refer to states as vectors. A quantum state may not be a vector like the familiar vectors in three-dimensional space but it is a vector nonetheless because we can do with states what we do with ordinary vectors: we can add states and we can multiply states by numbers. Linearity in quantum mechanics guarantees that adding wavefunctions or states is a sensible thing to do. Just like any vector can be written as a sum of other vectors in many different ways, we will do the same with our states. By writing our physical state as sums of other states we can learn about the properties of our state.

Consider now two states $|A\rangle$ and $|B\rangle$. Assume, in addition, that when measuring some property Q in the state $|A\rangle$ the answer is always a , and when measuring the same property Q in the state $|B\rangle$ the answer is always b . Suppose now that our physical state $|\Psi\rangle$ is the superposition

$$|\Psi\rangle = \alpha|A\rangle + \beta|B\rangle, \quad \alpha, \beta \in \mathbb{C}. \quad (1.4.1)$$

What happens now if we measure property Q in the system described by the state $|\Psi\rangle$? It may seem reasonable that one gets some intermediate value between a and b , but this is not what happens. A measurement of Q will yield either a or b . There is no certain answer, classical determinism is lost, but the answer is always one of these two values and not an intermediate one. The coefficients α and β in the above superposition affect the probabilities with which we may obtain the two possible values. In fact, the probabilities to obtain a or b

$$\text{Probability}(a) \sim |\alpha|^2, \quad \text{Probability}(b) \sim |\beta|^2. \quad (1.4.2)$$

Since the only two possibilities are to measure a or b , the actual probabilities must sum to one and therefore they are given by

$$\text{Probability}(a) = \frac{|\alpha|^2}{|\alpha|^2 + |\beta|^2}, \quad \text{Probability}(b) = \frac{|\beta|^2}{|\alpha|^2 + |\beta|^2}. \quad (1.4.3)$$

If we obtain the value a , immediate repeated measurements would still give a , so the state after the measurement must be $|A\rangle$. The same happens for b , so we have

$$\begin{aligned} \text{After measuring } a \text{ the state becomes } |\Psi\rangle &= |A\rangle, \\ \text{After measuring } b \text{ the state becomes } |\Psi\rangle &= |B\rangle. \end{aligned} \quad (1.4.4)$$

In quantum mechanics one makes the following assumption: *Superposing a state with itself doesn't change the physics*, nor does it change the state in a non-trivial way. Since superimposing a state with itself simply changes the overall number multiplying it, we have that Ψ and $\alpha\Psi$ represent the same physics for any complex number α different from zero. Thus, letting \cong represent physical equivalence

$$|A\rangle \cong 2|A\rangle \cong i|A\rangle \cong -|A\rangle. \quad (1.4.5)$$

This assumption is necessary to verify that the polarization of a photon state has the expected number of degrees of freedom. The polarization of a plane wave, as one studies in electromagnetism, is described by two real numbers. For this consider an elliptically polarized wave, as shown in Figure 1.5. At any given point, the electric field vector traces an ellipse whose shape is encoded by the ratio a/b of the semi-major axes (the first real parameter) and a tilt encoded by the angle θ (the second real parameter). Consider for this a general photon state $|\Psi\rangle$ formed by superposition of the two independent polarization states $|\text{photon}; x\rangle$ and $|\text{photon}; y\rangle$:

$$|\Psi\rangle = \alpha|\text{photon}; x\rangle + \beta|\text{photon}; y\rangle, \quad \alpha, \beta \in \mathbb{C}. \quad (1.4.6)$$

At first sight it looks as if the superposition is determined by two non-zero complex parameters α and β , or equivalently, four real parameters. But since the overall factor does not matter we can multiply this state by $1/\alpha$ to get the equivalent state that encodes all the physics

$$|\Psi\rangle \simeq |\text{photon}; x\rangle + \frac{\beta}{\alpha} |\text{photon}; y\rangle, \quad (1.4.7)$$

showing that we really have one complex parameter, the ratio β/α . This is equivalent to two real parameters, as expected.

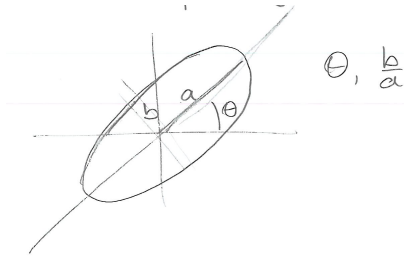


Figure 1.5: Parameters that define an elliptically polarized state.

Let us do a further example of superposition using electrons. Electrons are particles with spin. Classically, we imagine them as tiny balls spinning around an axis that goes through the particle itself. Once an axis is fixed, the electron has two and only two options: its rotation may be clockwise or counterclockwise about the axis, but in both cases it spins at the same fixed rate. These opposite ways of spinning are called *spin up* and *spin down* along the axis (see Figure 1.6). The up and down refer to the direction of the angular momentum associated with the rotation, and it is indicated by an arrow. According to quantum mechanics, and as verified by multiple experiments, the same possibilities, up or down, arise *whatever* axis we use to measure the spin of the electron.

Physicists usually set up coordinate systems in space by choosing three orthogonal directions, the directions of the x , y , and z axes. Let us choose to describe our spinning electrons using the z axis. One possible state of an electron is to be spin up along the z axis. Such a state is described as $|\uparrow; z\rangle$, with an arrow pointing up, and the label z indicating that the spin arrow points along the increasing z direction. Another possible state of an electron is spin down along the z axis. Such a state is described as $|\downarrow; z\rangle$, with an arrow pointing down, meaning this time that the spin points along the decreasing z direction. If these two are possible states, the state $|\Psi\rangle$ represented by the sum is also a possible state:

$$|\Psi\rangle = |\uparrow; z\rangle + |\downarrow; z\rangle.$$

What kind of physics does this sum $|\Psi\rangle$ represent? It represents a state in which a measurement of the spin along the z axis would result in two possible outcomes with equal probabilities: an electron with spin up or an electron with spin down. Since we can only speak of probabilities, any experiment must be repeated multiple times until probabilities can be estimated. Suppose we had a large ensemble of such electrons, all of them in the above state $|\Psi\rangle$. As we measured their spin along z , one at a time, we would find about half of them spinning up along z and the other half spinning down along z . There is no way to predict which option will be realized as we measure each electron. It is not easy to imagine superposition, but one may try as follows. An electron in the above state is in a different kind of existence in which it is able to both be spinning up along z and spinning down along z simultaneously! It is in such a ghostly, eerie state, doing incompatible things simultaneously, until its spin is measured. Once measured, the electron must immediately choose one of the two options; we always find electrons either spinning up or spinning down.

A critic of quantum mechanics could suggest a simpler explanation for the above observations. He or she would claim that the following simpler ensemble results in identical experimental results. In the critic's ensemble we have a large number of electrons with 50% of them in the state $|\uparrow; z\rangle$ and 50% of them in the state $|\downarrow; z\rangle$. He or she would then state, correctly, that such an ensemble would yield the same measurements of spins along z as the ensemble of those esoteric $|\Psi\rangle$ states. The new ensemble could provide a simpler explanation of the result without having to invoke quantum superpositions.

Quantum mechanics, however, allows for further experiments that can distinguish between the ensemble of our friendly critic and the ensemble of $|\Psi\rangle$ states. While it would take us too far afield to explain this, if we measured the spin of the electrons in the x direction,

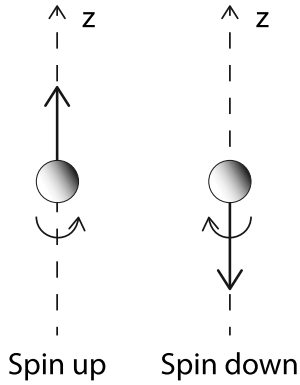


Figure 1.6: An electron with spin along the z axis. Left: the electron is said to have spin up along z . Right: the electron is said to have spin down along z . The up and down arrows represent the direction of the angular momentum associated with the spinning electron.

instead of z direction, the results would be *different* in the two ensembles. In the ensemble of our critic we would find 50% of the electrons up along x and 50% of the electrons down along x . In our ensemble of $|\Psi\rangle$ states, however, one can show that measurement yields a surprisingly simple result: all states pointing up along x . The critic's ensemble is therefore not equivalent to our quantum mechanical ensemble. The critic is wrong in his or her claim that quantum mechanical superpositions are not required.

1.5 Entanglement

When we consider superposition of states of *two* particles we can get the remarkable phenomenon called *quantum mechanical entanglement*. Entangled states of two particles are those in which we can't speak separately of the state of each particle. The particles are bound together in a common state in which they are *entangled* with each other.

Let us consider two non-interacting particles. Particle 1 could be in any of the states

$$\{|A_1\rangle, |A_2\rangle, \dots\}, \quad (1.5.1)$$

while particle 2 could be in any of the states

$$\{|B_1\rangle, |B_2\rangle, \dots\} \quad (1.5.2)$$

It may seem reasonable to conclude that the state of the full system, including particle 1 and particle 2 would be specified by stating the state of particle 1 and the state of particle 2. If that would be the case the possible states would be written as

$$|A_i\rangle \otimes |B_j\rangle, \quad (1.5.3)$$

for some specific choice of i and j that specify the state of particle one and particle two, respectively. Here we have used the symbol \otimes , which means *tensor* product, to combine

the two states into a single state for the whole system. We will study \otimes later, but for the time being we can think of it as a kind of product that distributes over addition and obeys simple rules, as follows

$$(\alpha_1|A_1\rangle + \alpha_2|A_2\rangle) \otimes (\beta_1|B_1\rangle + \beta_2|B_2\rangle) = \alpha_1\beta_1|A_1\rangle \otimes |B_1\rangle + \alpha_1\beta_2|A_1\rangle \otimes |B_2\rangle \\ + \alpha_2\beta_1|A_2\rangle \otimes |B_1\rangle + \alpha_2\beta_2|A_2\rangle \otimes |B_2\rangle. \quad (1.5.4)$$

The numbers can be moved across the \otimes but the order of the states must be preserved. The state on the left-hand side—expanded out on the right-hand side—is still of the type where we combine a state of the first particle ($\alpha_1|A_1\rangle + \alpha_2|A_2\rangle$) with a state of the second particle ($\beta_1|B_1\rangle + \beta_2|B_2\rangle$). Just like any one of the states listed in (1.5.3) this state is not entangled.

Using the states in (1.5.3), however, we can construct more intriguing superpositions. Consider the following one

$$|A_1\rangle \otimes |B_1\rangle + |A_2\rangle \otimes |B_2\rangle. \quad (1.5.5)$$

A state of two particles is said to be **entangled** if it cannot be written in the factorized form $(\dots) \otimes (\dots)$ which allows us to describe the state by simply stating the state of each particle. We can easily see that the state (1.5.5) cannot be factorized. If it could it would have to be with a product as indicated in (1.5.4). Clearly, involving states like $|A_3\rangle$ or $|B_3\rangle$ that do not appear in (1.5.5) would not help. To determine the constants $\alpha_1, \alpha_2, \beta_1, \beta_2$ we compare the right hand side of (1.5.4) with our state and conclude that we need

$$\alpha_1\beta_1 = 1, \quad \alpha_1\beta_2 = 0, \quad \alpha_2\beta_1 = 0, \quad \alpha_2\beta_2 = 1. \quad (1.5.6)$$

It is clear that there is no solution here. The second equation, for example, requires either α_1 or β_2 to be zero. Having $\alpha_1 = 0$ contradicts the first equation, and having $\beta_2 = 0$ contradicts the last equation. This confirms that our state (1.5.5) is indeed an entangled state. There is no way to describe the state by specifying a state for each of the particles.

Let us illustrate the above discussion using electrons and their spin states. Consider a state of two electrons denoted as $|\uparrow\rangle \otimes |\downarrow\rangle$. As the notation indicates, the first electron, described by the first arrow, is up along z while the second electron, described by the second arrow, is down along z (we omit the label z on the state for brevity). This is not an entangled state. Another possible state is one where they are doing exactly the opposite: in $|\downarrow\rangle \otimes |\uparrow\rangle$ the first electron is down and the second is up. This second state is also not entangled. It now follows that by superposition we can consider the state

$$|\uparrow\rangle \otimes |\downarrow\rangle + |\downarrow\rangle \otimes |\uparrow\rangle. \quad (1.5.7)$$

This is an entangled state of the pair of electrons.

Exercise. Show that the above state (1.5.7) cannot be factorized and is therefore entangled.

In the state (1.5.7) the first electron is up along z if the second electron is down along z (first term), or the first electron is down along z if the second electron is up along z (second

term). There is a correlation between the spins of the two particles; they always point in opposite directions. Imagine that the two entangled electrons are very far away from each other: Alice has one electron of the pair on planet earth and Bob has the other electron on the moon. Nothing we know is connecting these particles but nevertheless the states of the electrons are linked. Measurements we do on the separate particles exhibit correlations. Suppose Alice measures the spin of the electron on earth. If she finds it up along z , it means that the first term in the above superposition is realized, because in that term the first particle is up. As discussed before, the state of the two particles immediately becomes that of the first term. This means that the electron on the moon will *instantaneously* go into the spin down-along- z configuration, something that could be confirmed by Bob, who is sitting in the moon with that particle in his lab. This effect on Bob's electron happens before a message, carried with the speed of light, could reach the moon telling him that a measurement has been done by Alice on the earth particle and the result was spin up. Of course, experiments must be done with an ensemble that contains many pairs of particles, each pair in the same entangled state above. Half of the times the electron on earth will be found up, with the electron on the moon down and the other half of the times the electron on earth will be found down, with the electron on the moon up.

Our friendly critic could now say, correctly, that such correlations between the measurements of spins along z could have been produced by preparing a *conventional* ensemble in which 50% of the pairs are in the state $|\uparrow\rangle \otimes |\downarrow\rangle$ and the other 50% of the pairs are in the state $|\downarrow\rangle \otimes |\uparrow\rangle$. Such objections were dealt with conclusively in 1964 by John Bell, who showed that if Alice and Bob are able to measure spin in *three* arbitrary directions, the correlations predicted by the quantum entangled state are different from the classical correlations of *any* conceivable conventional ensemble. Quantum correlations in entangled states are very subtle and it takes sophisticated experiments to show they are not reproducible as classical correlations. Indeed, experiments with entangled states have confirmed the existence of quantum correlations. The kind of instantaneous action at a distance associated with measurements on well-separated entangled particles does not lead to paradoxes nor, as it may seem, to contradictions with the ideas of special relativity. You cannot use quantum mechanical entangled states to send information faster than the speed of light.