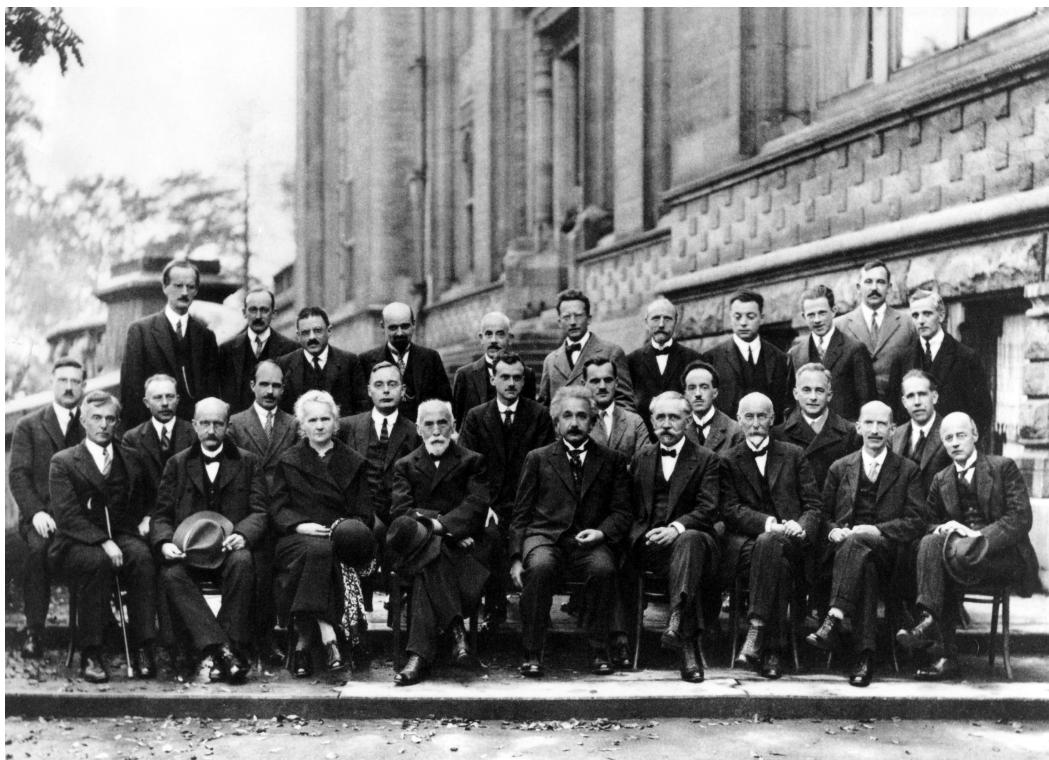


Quantum Physics Lecture Notes



Understanding the Schrödinger Equation

Sebastian de Haro

Amsterdam University College, Fall Semester, 2014

Cover illustration: Wikipedia

Contents

Introduction	2
1 Motivating the Schrödinger Equation	3
1.1 Classical waves	3
1.2 Enter the quantum	4
1.3 The wave function	7
2 Four Steps to Solve the Schrödinger Equation	10
2.1 The Problem	10
2.2 Step 1: Reduce to TISE	10
2.3 Step 2: General Solution of TDSE	11
2.4 Step 3: Impose Initial Condition $\Psi_0(x)$	12
2.5 Step 4: Plug Back into $\Psi(x, t)$	14
2.6 Example: Gaussian Wave Function	14
3 One-Dimensional Potentials	16
3.1 General Theorems	16
3.2 The TISE with one-dimensional potentials	17
4 Fourier Integrals and the Dirac Delta	19
4.1 The Dirac delta	19
4.2 Fourier transformations	20
5 The Formalism of Quantum Mechanics	21
5.1 Why?	21
5.2 The Postulates of Quantum Mechanics	22
5.3 Linear Algebra	23
5.4 Continuous spectra	24
6 Dirac Notation	26
6.1 Base-free notation	26
6.2 Closure	27
6.3 Bras and kets	28
7 The Interpretation of Quantum Mechanics	29
7.1 EPR and Hidden Variables	30
7.2 Bohr's Reply to EPR	31
7.3 The Measurement Problem	32
7.4 Other Approaches	35
7.4.1 Decoherence	35
7.4.2 Many Worlds	36
A Mathematical Formulas and Tricks	37
A.1 Gaussian integration	37

B Technicalities of Quantum Mechanical Measurements	38
B.1 Time Evolution Operators	38
B.2 The Measurement Operator U	39
References	40

Introduction

Welcome to the fascinating world of quantum mechanics! We are going to learn how to do computations in quantum mechanics and how to interpret our results physically. I say quantum mechanics is fascinating because it is so different from any other physical theory we have ever seen before. In classical physics, particles travel along trajectories that can be drawn in space and time. In quantum mechanics, particles don't have trajectories and sometimes we can't even say where they are located in space! In fact, quantum mechanics is so weird that some of the scientists who made major contributions to it—notoriously, Albert Einstein and Erwin Schrödinger—never believed the interpretation of quantum theory that has now become standard. Richard Feynman famously said, “I think I can safely say that nobody understands quantum mechanics.” So, remember: if you get confused, you are in good company. However, I don't think Feynman was entirely right on this one. We don't seem to have quantum mechanical intuition wired into our brains and bodies (as we do learn to appreciate weight, height, etc., intuitively). Feynman was certainly right about that. But we can learn to work with the theory and progressively develop a rational intuition for it, based on what we learn from the formulas and the physical principles they embody. By logical thinking and some guesswork, by looking at the experimental facts (black body radiation, photoelectric effect, double slit experiment, etc.) we can try to find the necessary physical principles that will allow us to construct the theory of quantum mechanics. We can't *derive* quantum mechanics, let me be clear about that, just like we can't *derive*, from sheer thinking about the concept of mass, the fact that Newton's law of gravity decreases with the square of the distance. But we can look for the minimal set of principles and equations that will reproduce and unify all of the facts we know about quantum mechanics. This is what Schrödinger's equation does for us. Once we have motivated why it should be true, we have to learn how to work with it and to interpret it. Indeed this is the right order: first learn to work with it, then gain deeper understanding of what it means. Remember that in the short period in which the ‘new’ quantum theory was developed (basically, Christmas of 1925 to the summer of 1926) Heisenberg and Schrödinger worked on the math and that only later on did they focus on its meaning. We will be doing both things as simultaneously as possible—developing the theory while discussing its interpretation.

These lecture notes are to be used as a complement to a textbook on quantum mechanics such as Griffiths' [1]. I will focus on a number of selected issues that I think are important to understand quantum mechanics. In the first chapter I motivate how the Schrödinger equation comes to be in the first place, rather than just throwing it at you and asking you to trust quantum mechanics. Whereas the Schrödinger equation cannot be *derived* from classical mechanics, it can be *motivated* from semi-classical considerations. I also expand, in later sections, on mathematical explanations and include study tools.

Writing down a mathematical theory of quantum mechanics assumes knowledge of the **basic experiments** that led to this theory and the broad principles that were derived from them. Therefore I recommend that, before you start reading these lecture notes, you refresh your memory (or catch up, whichever may be the case) on the most relevant historical experiments and the physical principles that were drawn from them. I recommend chapter 37 of Giancoli [2]. If you want a more thorough exposition, you can read

chapter 1 of [3]. I will refer to these experiments regularly.

1 Motivating the Schrödinger Equation

1.1 Classical waves

We recall some concepts from classical physics that will both serve fix notation and motivate the way we introduce waves in quantum mechanics. Consider a one-dimensional wave, say:

$$y(x) = A \sin kx . \quad (1)$$

This is a static, sinusoidal wave extending in the x -direction. A , the maximum height, is called the *amplitude* of the wave. The *wavelength* is the distance between two identical points on the wave (e.g. two crests, two troughs, two nodes) and in the above example it is given by $\lambda = \frac{2\pi}{k}$ because this is the smallest number for which the wave repeats itself, i.e. $y(x + \lambda) = y(x)$. k is usually called the *wave number*. Consider now a wave traveling at speed v :

$$y(x, t) = A \sin(k(x - vt)) = A \sin\left(\frac{2\pi}{\lambda}(x - vt)\right) . \quad (2)$$

The period T of this wave is the time it takes for the wave to go back to itself in time, analogously to the wavelength: $y(x, t + T) = y(x, t)$, hence $T = \lambda/v$. You see that the interpretation of v as the speed is right, as v is indeed the wavelength divided by the period. The frequency is defined as $\nu = \frac{1}{T}$, and the angular frequency is $\omega = 2\pi\nu = \frac{2\pi}{T}$. The advantage of using the wave number and the angular frequency is that they include the periodicity of the sine and cosine and the factors of 2π nicely disappear:

$$y(x, t) = A \sin(kx - \omega t) . \quad (3)$$

This is a notation that we will use quite often. I am deliberately using ν for the frequency here and not f , because this is the standard notation in more advanced texts. We will use the Greek alphabet extensively in this course, so you better get used to it!: $\alpha, \beta, \gamma, \delta, \epsilon, \dots$

Complex waves. In quantum mechanics we will always use complex waves, for reasons that will become clear later. A complex wave looks like:

$$\Psi(x, t) = A e^{i(kx - \omega t)} . \quad (4)$$

Now recall the following result:

Euler's formula. $e^{i\varphi} = \cos \varphi + i \sin \varphi$.

Proof. This can be proven by using the following relations for the sine and cosine, which you should know:

$$\begin{aligned} \sin \varphi &= \frac{e^{i\varphi} - e^{-i\varphi}}{2i} \\ \cos \varphi &= \frac{e^{i\varphi} + e^{-i\varphi}}{2} . \end{aligned} \quad (5)$$

From this, we get:

$$\cos \varphi + i \sin \varphi = \frac{1}{2} e^{i\varphi} + \frac{1}{2} e^{-i\varphi} + i \frac{1}{2i} e^{i\varphi} - i \frac{1}{2i} e^{-i\varphi} = e^{i\varphi}, \quad (6)$$

which proves Euler's formula.

Remark. If you don't recognise formula (5) at all, try to prove it using the Taylor expansions (Taylor series) of the sine, cosine, and exponential functions, which you *should* know:

$$\begin{aligned} \sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \\ \cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \\ e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!}. \end{aligned} \quad (7)$$

As you can see, the cosine is an even function of x , hence it only has terms with even powers, x^{2n} , whereas the sine is odd and only has odd powers x^{2n+1} . By taking combinations of $e^{i\varphi}$ and $e^{-i\varphi}$ with plus and minus sign, respectively, in (5), we cancel the odd/even terms and are left with the cosine or sine, respectively.

Because of Euler's formula, we see that the cosine is the real part of the wave while the sine is its imaginary part:

$$\begin{aligned} \cos \varphi &= \operatorname{Re}(e^{i\varphi}) \\ \sin \varphi &= \operatorname{Im}(e^{i\varphi}). \end{aligned} \quad (8)$$

Hence, we can always replace sinusoidal waves such as (3) by complex waves (4) and take the real or imaginary part as needed.

Having made these remarks on complex waves we now go back to the physical meaning of the quantities involved in a wave. I want to make the following point about the amplitude A . When we have a wave of the type $A \sin x$ or $A e^{ix}$, the quantity that is of physical interest is not A itself, but $|A|^2 \equiv AA^*$. Mathematically, the reason is that A could be made to be negative by simply changing our coordinate by $x \rightarrow -x$ in $\sin x$, and obviously that shouldn't affect the amplitude, which is independent of the orientation of the coordinate. Also, in a complex wave $e^{i\varphi}$ we could shift $\varphi \rightarrow \varphi + \alpha$ by some constant α and generate a *complex phase* $e^{i\alpha}$ which then would become part of A . But also that is a change of variables (namely, choosing a different zero point for φ) which should not affect the amplitude. For these reasons, the physical quantities of interest depend on the absolute value (squared) of the amplitude $|A|^2$ and not on A itself.

1.2 Enter the quantum

Experiments such as the two-slit experiment make clear that particles sometimes behave as waves. This is true not only for light, but also for material particles such as electrons.

Indeed, in a flash of genius Louis de Broglie hypothesized that not only electromagnetic radiation has a dual nature as waves and as particles (photons), but that also matter, usually believed to be of corpuscular nature, should possess wave-like properties.

For photons, Planck's 1900 hypothesis—clarified and generalized by Einstein's 1905 description of the photoelectric effect—had been that the energy is quantized in units of h (the *Planck constant*): $E = h\nu = hc/\lambda$. The quantity $p = E/c = h/\lambda$ had been identified as the photon's momentum, which played an important role in the impressive Compton effect, where a single photon could hit an electron at rest and impart it non-zero momentum.

de Broglie now proposed that, analogously to the case of light, matter waves also have a frequency ν and a wavelength λ , given by:

$$\begin{aligned}\nu &= \frac{E}{h} \\ \lambda &= \frac{h}{p}.\end{aligned}\tag{9}$$

Indeed it is one of the great features of Planck's constant h that, by introducing a fundamental constant with units $\text{J}\cdot\text{s}$, such relations between frequency and energy and between wavelength and momentum can be written down. For the same reasons for which one introduces the angular frequency and the wave number, we will often use the *reduced Planck constant*

$$\hbar = \frac{h}{2\pi},\tag{10}$$

which is usually simply referred to as ‘Planck's constant’. In terms of these, the above read:

$$\begin{aligned}E &= \hbar\omega \\ p &= \hbar k.\end{aligned}\tag{11}$$

Furthermore, remember the classical relation $E = T + V$. If we ignore the potential, we have, classically:

$$E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m},\tag{12}$$

where m is the mass of the electron. We will use this expression a lot in what follows.

Remark. The above formulas (11)-(12) also imply a relation between the frequency and the wavelength of the wave (viz. between the angular frequency and the wave number): $\omega(k) = \frac{\hbar k^2}{2m}$. Such a relation is called a *dispersion relation* because it determines the effect of dispersion of the wave as it travels through a medium¹. The above dispersion

¹Namely, by giving the speed as a function of the wavelength.

relation holds for a matter particle. A photon, on the other hand, is massless and has the dispersion relation² $E = pc$.

The central question that Schrödinger asked, then, is as follows. We know that, at least semi-classically, electrons have energy given by (12). But they are also waves of the type:

$$\Psi(x, t) = A e^{i(kx - \omega t)} . \quad (13)$$

So how can such a wave have energy (12)? The answer from optics would be: write down a wave equation! So following de Broglie's lead, Schrödinger had the idea writing down a *wave* equation that would reproduce (12). The equation in question is the following:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} . \quad (14)$$

As you can see by filling (13) into this equation, (14) looks like:

$$i\hbar(-i\omega)A e^{i(kx - \omega t)} = -\frac{\hbar^2}{2m}(-k^2)A e^{i(kx - \omega t)} , \quad (15)$$

which holds if: $\omega = \frac{\hbar k^2}{2m}$, which is precisely the condition (12). So the expression for the energy does appear as a consequence of Schrödinger's wave equation (13): the wave equation forces on us a relation between the angular frequency ω and the wave number k , and this relation is nothing but the classical energy relation (12). This is nice: waves seem to describe electrons too!

This is the one-dimensional Schrödinger equation for a *free* particle (free electron), i.e. one with zero potential energy hence satisfying (12). Now if there is potential energy V around, (12) will change to:

$$E = \frac{p^2}{2m} + V , \quad (16)$$

again a result from classical mechanics. So it is natural to modify the Schrödinger equation as follows:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x)\Psi(x) \equiv H\Psi(x) . \quad (17)$$

Here, $V(x)$ is the *potential function*, that is the function from which (in classical mechanics) the force can be derived. And the Hamiltonian H was defined as:

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) . \quad (18)$$

²Ironically, quantum theory originated from considerations of the wave vs. particle nature of photons, but the Schrödinger equation only describes matter particles such as electrons and protons which, because they are massive, travel at low speeds. Photons travel at the speed of light and one needs to take relativistic effects into account in order to describe them quantum mechanically, which in turn means that one needs to generalize the Schrödinger equation and replace it by an appropriate equation incorporating special relativity. The reason the Schrödinger equation works for massive particles such as the electron is that it is based on the *non-relativistic* limit of the energy (12), which can be conveniently ‘quantized’, as I show below.

The Hamiltonian is an *operator* (a kind of derivative) that *represents* energy in quantum mechanics. If you have studied the Hamiltonian formalism in classical mechanics, you might remember that classically the Hamiltonian is the sum of the kinetic and the potential energy *functions*, $T + V$. The Hamiltonian we just defined is not a function, but an *operator*: just like the derivative $\partial/\partial x$ (also an operator), it acts on functions: $H\psi(x) \equiv -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi(x)$. That is, given a (wave) function $\psi(x)$, an operator assigns to it a *new* wave function $H\psi(x)$ defined by the equation above. In quantum mechanics, classical quantities such as the energy are replaced by operators. A quantum system may not always have a definite energy, but the energy operator can always be defined. We will see later on under what conditions we reproduce classical formulas like (16).

The above is not a *derivation* of the Schrödinger equation. All I have done is motivate that, given classical formulas such as (12), and given the assumption that particles are also waves like (13), we can write down a formula that reproduces (12) and hence encompasses both principles. But I have not derived it. Then, based on classical mechanics, I made it plausible that (17) is the right generalization to the case of non-vanishing potential. Much of what we will do in this course is solving the Schrödinger equation (17) for various forms of the potential $V(x)$, and finding out precisely under what conditions we can make predictions of the type (16).

1.3 The wave function

We will now look in more detail at the interpretation of the wave function. Before we proceed, let me remark that the wave function Ψ is necessarily a complex quantity. We cannot stick to its real or to its imaginary part (as we would do with, for instance, classical electromagnetic waves) because the time evolution will normally develop an imaginary part, even if we start with a purely real wave function. The reason is in the form of the Schrödinger equation (17). Taking the complex conjugate of this equation, we get a *different* equation:

$$-i\hbar \frac{\partial \Psi^*}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi^*}{\partial x^2} + V\Psi^*. \quad (19)$$

Notice the difference in sign on the left-hand side compared to (17). Thus, Ψ and Ψ^* satisfy different equations, which means that the real and the imaginary part of the wave function contain different information. Roughly speaking, Ψ^* is the mirror image of Ψ under $t \rightarrow -t$, and if Ψ propagates ‘forward’ in time, we can say that Ψ^* propagates ‘backward’. In other words, it is not enough to look for real solutions of the Schrödinger equation as we would be losing essential information.

Interpretation. Schrödinger originally thought—although he was careful enough not to write this in his paper—that the absolute value squared of the wave function, $|\Psi|^2 = \Psi\Psi^*$, could be interpreted as some kind of charge density or particle density distributed over the space. This was analogous to the classical theory of light, where the intensity of the light is proportional to the square of the field. As it turns out, this interpretation does not conflate well with quantum mechanics. Consider doing the double-slit experiment with a *single* particle. Repeating the experiment many times the pattern that appears

is described by the distribution $|\Psi(x, t)|^2$. Since every time we do the experiment there is just one particle in the set-up, this means that the wave function is not the average particle density in a single experiment, but rather an average over many experiments. This is usually called an *ensemble average*. It is Max Born who is responsible for the interpretation of $|\Psi|^2$ as a probability density. $|\Psi(x, t)|^2 dx$ is the probability to find, upon detection, a particle within a small neighbourhood³ $(x, x + dx)$ of x at time t . Notice the insistence on detection: this is necessary because before we measure the particle we cannot make any assumptions about its location. In the double slit experiment that we discussed earlier, we cannot say which slit the particle went through unless we measure it. This is another reason why $|\Psi(x, t)|^2$ cannot be interpreted as a ‘particle density’: the particles cannot be localized until we measure them.

Let me now further motivate why the probability grows like the *square* of the wave function, instead of the wave function itself. I have already mentioned that, for waves, only *amplitudes* such as $|\Psi|$, and not Ψ itself, can have a physical meaning. This is because Ψ can become negative, even complex (something we don’t want for a probability). Now we also must explain why $|\Psi|^2$, and not $|\Psi|$, is the relevant quantity. This follows from the wavelike nature of the wave function, and I will motivate it with three examples:

1. *The electromagnetic field.* The intensity of radiation (the intensity of the radiation emitted by, say, an antenna, or the amount of light emitted by a light bulb—both manifestations of the electromagnetic field) grows proportionally to the square of the electromagnetic field E , not linearly with the field⁴. And it is the electromagnetic field which satisfies the differential equations of electromagnetism (called Maxwell’s equations), something analogous to our Schrödinger equation. This intensity of light is proportional to the number of photons with the given frequency (and, hence, to the density of these photons). It makes sense that the probability for a particle to have a given frequency be the quantum mechanical quantity corresponding to the number of particles with a given frequency in classical electromagnetism.

2. *The harmonic oscillator.* We have seen that the Schrödinger equation could be interpreted as reproducing the relation between the energy and the momentum (12) (for the case of a simple wave like (13)). We extended it to (17) in the case of non-zero potential. So let us actually compute this energy for a familiar classical system with non-zero potential, the harmonic oscillator. The potential is $V = \frac{1}{2} kx^2$ and the energy function is $H = \frac{1}{2} m\dot{x}^2 + \frac{1}{2} kx^2$. Solving Newton’s equation $x = A \sin(\omega t + \phi)$, plugging this back into

³It is incorrect to say that $|\psi(x)|^2$ is the probability to find the particle at point x . The probability to find a particle exactly at point x is in fact zero. On the other hand, the probability to find the particle in the neighborhood $(x, x + dx)$ is infinitesimal and non-zero. The probability to find a particle in a finite interval (a, b) is $P(a, b) = \int_a^b |\psi(x)|^2 dx$. From the latter formula we see that, for any point x , $P(a, x) = \int_a^x |\psi(y)|^2 dy$ (notice the different name for the dummy variable y), from which it follows that $\frac{\partial P}{\partial x} = |\psi(x)|^2$. Hence $|\psi(x)|^2$ itself can be interpreted as the spatial *rate of change*, or gradient, of the probability at point x .

⁴The intensity of radiation is defined as the amount of energy emitted per time per unit area, that is, the power transported across a given unit area perpendicular to the flow. It is measured in units of W/m^2 .

the energy function and using $\omega = \sqrt{k/m}$, we get the familiar result: $E = \frac{1}{2} kA^2$, indeed the square of the amplitude. So, again, classical energies behave like squares of amplitudes.

3. *Conservation of probability.* Taking $P \propto |\Psi|^2$, one can prove a ‘probability conservation theorem’ completely analogous to the charge conservation theorem in electrodynamics (if charge disappears, it must be taken away by a current, i.e. the time derivative of the charge density is minus the gradient of the current, $\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{J}$). The analogous result for probabilities involves the time derivative of the probability (see problem 1.14 of [1]). This only works if one takes the *square* of the amplitude to be the probability.

4. *Interference.* If, in view of these arguments, we accept the identification of the probability distribution with the *square* of the wave function, we still have to explain how interference appears. What I will show now is that interference appears when we add up the wave functions, $\psi = \psi_1 + \psi_2$, rather than the probabilities. Consider again the two-slit experiment, where we now close one of the slits. Call $\psi_1(x)$ the wave function at point x on the screen when only one slit is open (the left one, say) and $\psi_2(x)$ when the other is open. According to the superposition principle of waves, when both slits are open the wave function is $\psi(x) = \psi_1(x) + \psi_2(x)$. Now because $P \propto |\psi_1(x) + \psi_2(x)|^2$, this is different from the sum of the probabilities. There is a cross term in the square, and this cross term is responsible for interference. You can easily simulate this yourself by writing the following mathematica program:

```
a = 0.5; b = 5;
psi1[x_] := Exp[-(x+a)^2/2];
psi2[x_] := Exp[-(x-a)^2/2];
psi[x_] := psi1[x] - psi2[x];
Plot[psi1[x], {x,-b,b}]
Plot[psi2[x], {x,-b,b}]
Plot[psi[x], {x,-b,b}]
Plot[psi1[x]^2+psi2[x]^2, {x,-b,b}]
```

The outcome of the program is depicted in Figure 1. As you see from the picture, there is an interference pattern that follows when we add up the wave functions rather than the probabilities themselves. The result of adding up the probabilities is in the last picture. You see that in that case interference completely disappears. Hence, it is important that we add the wave functions, $\psi = \psi_1 + \psi_2$ (corresponding $P \propto |\psi_1 + \psi_2|^2$) rather than the probabilities ($P \neq P_1 + P_2$).

Hopefully the above arguments have convinced you that we should interpret $|\Psi(x, t)|^2$ as the probability of finding the particle within a region $(x, x + dx)$ at time t . These arguments of course do not provide a derivation; they only make the assertion plausible. Like when we set up the Schrödinger equation, we are now constructing a new theory and we cannot give derivations from first principles. Otherwise this would be no new theory at all! So we write down tentative equations and try to uncover the fundamental principles. In the end, of course, we will know that the structure is right because it works—it solves many problems that would otherwise be intractable, and it agrees with the results of experiment. Both things tell us that quantum mechanics is right!

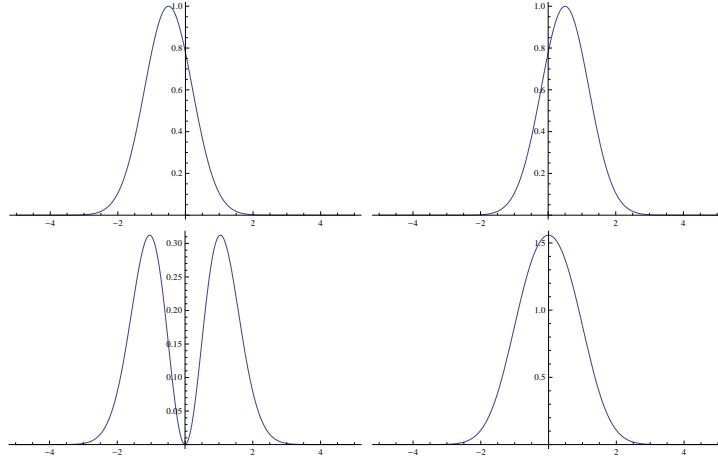


Figure 1: Two slit experiment (from left to right, and from top to bottom): 1) with only the left slit open; 2) with only the right slit open; 3) interference pattern when both slits are open; 4) classical result one obtains by simply adding up the probabilities.

2 Four Steps to Solve the Schrödinger Equation

In this section I will summarize and further explain the steps you have to take to solve the time-dependent Schrödinger equation as outlined in [1].

2.1 The Problem

The basic problem is to solve the time-dependent Schrödinger equation

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \Psi(x, t) &= H \Psi(x, t) , \\ H &\equiv -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) , \end{aligned} \quad (20)$$

given a known initial wave function (also called the ‘initial condition’) at time⁵ $t = 0$:

$$\Psi(x, 0) = \Psi_0(x) , \quad (21)$$

where $\Psi_0(x)$ is a *given function*. I make the distinction between Ψ and Ψ_0 because $\Psi(x, t)$ is the function we want to solve for (we only know it at one particular point in time, $t = 0$), whereas Ψ_0 is a given function of x . Griffiths doesn’t use Ψ_0 , only $\Psi(x, 0)$.

The solution to the above problem consists in a four-step procedure.

2.2 Step 1: Reduce to TISE

Look for special solutions of the type

$$\Psi(x, t) = \varphi(t) \psi(x) . \quad (22)$$

⁵One may as well choose the initial condition at any other initial time $t = t_{\text{in}}$.

These solutions are called *separable*. Filling this ansatz into the Schrödinger equation, we completely solve for φ and find another equation for ψ :

$$\varphi(t) = e^{-\frac{i}{\hbar} Et} \quad (23)$$

$$H \psi(x) = E \psi(x) \Rightarrow -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x) \psi(x) = E \psi(x) . \quad (24)$$

Notice that I write partial derivatives $\partial/\partial t$ for functions such as $\Psi(x, t)$ that depend on several variables, and regular derivatives d/dx for functions $\psi(x)$ that depend on a single variable.

Equation (24) is called the *time-independent Schrödinger equation*. Filling the ansatz (22) into the time dependent equation, we were able to solve for φ in (23). This gave us an integration constant E , which reappears in (24) and can be interpreted as an energy (though not ‘the energy of the system’, a concept which may not always be well defined). Solving (24) will give us $\psi(x)$ as well as the allowed values of the energy E , which is called the *spectrum* of the Hamiltonian H . The spectrum depends on the potential $V(x)$. In general, (24) has many solutions (usually an infinite number of them). For instance, $\psi(x) = \sin nx$ for positive integers $n = 1, 2, 3, \dots$. In this case, the spectrum is called **discrete** (as there are discrete energy levels labeled by $n = 1, 2, 3, \dots$). Accordingly, we label both the energies and the wave function by n in this case:

$$H \psi_n(x) = E_n \psi_n(x) . \quad (25)$$

The main problem now is to solve this equation, imposing the appropriate boundary conditions (this will be dealt with in chapter (3) for different types of potentials). For instance, for the infinite square well it turns out that $E_n = \frac{\hbar^2 \pi^2 n^2}{2ma^2}$.

If the spectrum is **continuous** rather than discrete, the energies are labeled by a continuous variable that we usually call k , the wave number. For instance, for the free particle: $E_k = \frac{\hbar^2 k^2}{2m}$, and the wave functions are $\psi_k(x) = e^{ikx}$. In this case, k is the wave number, with $k > 0$ for right-moving, $k < 0$ for left-moving waves. Remember that the wave number relates to the momentum of the wave by the de Broglie formula (11). In this case, the energy is given by (12).

To summarize this step, we have reduced the time-dependent Schrödinger equation (a partial differential equation in t and x) to the *time-independent* Schrödinger equation (a second order differential equation in x). It is useful to think of (25) as an eigenvalue problem. H is an operator (an object which, like a matrix, acts on a vector space) and ψ_n are the eigenfunctions, with corresponding eigenvalues E_n . Solving this eigenvalue problem gives us, like in linear algebra, both the eigenvalues and the eigenfunctions. We will develop this point of view further when we discuss the formalism.

2.3 Step 2: General Solution of TDSE

Once you have the solutions $\psi_n(x)$ and E_n of the TISE (25), you can move on to get the general solution of the TDSE (20). It is given by:

$$\Psi(x, t) = \sum_n c_n e^{-\frac{i}{\hbar} E_n t} \psi_n(x) . \quad (26)$$

It is straightforward to show that, if the individual solutions (22) solve (20), then a superposition of them like in (26) also is a solution. The reason is that (20) is a *linear* equation, and its solutions obey the superposition principle: if we have a set of solutions (22), we can add them together with arbitrary coefficients c_n and the result is also a solution. But the key point to understand here is that (26) in fact gives us the *most general solution* of the TDSE (20). Indeed, whereas the separation of variables (22) was an *ansatz* (an assumption, in other words, we only obtained a *specific* solution), one can show that *any* solution of the TDSE is in fact of the form (26).

To show this in general is a little tricky because it actually depends on the form of the potential function $V(x)$. But the idea, which we will later substantiate for specific potentials $V(x)$, is as follows. (20) is a first-order partial differential equation in t and x which we are solving in a two-step procedure of first solving for t and then for x . So we regard this equation as an ordinary, linear differential equation in t (i.e. we treat x as a constant). Since the equation is first order (only one t derivative), it depends on a single ‘integration constant’, which is our initial function $\Psi_0(x)$ (which I have not yet specified; therefore, for the time being Ψ_0 is a generic function). The theory of first-order differential equation tells us that, *if* I can choose c_n in (26) such that $\Psi(x, t)$ satisfies the initial condition $\Psi_0(x)$, *then* (26) is the unique solution associated with that $\Psi_0(x)$. Remember, a first order linear differential equation has only one integration constant/boundary condition (in this case, Ψ_0). Since $\Psi_0(x)$ was generic to start with, this is the same as saying that (26) is the most general solution (i.e. for *any boundary condition*) of the Schrödinger equation regarded as a first-order differential equation in t .

In the continuous case, we replace $n \rightarrow k$, $c_n \rightarrow \frac{1}{\sqrt{2\pi}} \phi(k)$, and the sum becomes an integral:

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \phi(k) e^{-\frac{i}{\hbar} E_k t} \psi_k(x) . \quad (27)$$

For the simple case of a free particle, $\psi_k(x) = e^{ikx}$ and this becomes:

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \phi(k) e^{i(kx - \frac{\hbar k^2}{2m} t)} . \quad (28)$$

2.4 Step 3: Impose Initial Condition $\Psi_0(x)$

In this step, we choose the coefficients c_n in (26) such that our initial condition (21) is satisfied. If we can do this, then we are left with the unique solution of the Schrödinger equation. Physically, what we are doing is we are imposing that the most general solution of the Schrödinger equation $\Psi(x, t)$ agrees with our experimental situation at time $t = 0$. We have prepared our system (e.g. a system of electrons with spin up) in a particular state $\Psi_0(x)$ at time $t = 0$. Given *this* particular state, $\Psi(x, t)$ tells us, via the Schrödinger equation, how the system evolves in time. For instance, H might contain an interaction between the spins and an external magnetic field, and so some of the spin states of the electrons may change in time. So, H encodes the dynamics of the system.

All we have to do is set $t = 0$ in (26) or (27) for the discrete/continuous case, respectively. We get:

$$\begin{aligned}\Psi_0(x) &= \sum_n c_n \psi_n(x), && \text{discrete spectrum} \\ \Psi_0(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \phi(k) \psi_k(x), && \text{continuous spectrum}\end{aligned}\quad (29)$$

So in this step, the task is to **find** c_n or $\phi(x)$ such that (29) is satisfied. It is here where we have to use the specific knowledge about ψ_n and ψ_k . Once we have done this, we have shown that (26)-(27) satisfy the given boundary condition.

Before I show how to do this in practice, we must reflect on the following question: does (29) always have a solution? Can we always be sure that this is somehow the case? Can I always find c_n and $\phi(k)$ to match $\Psi_0(x)$ on the l.h.s.? This is the same as asking whether $\Psi(x, t)$ was in fact the most general solution of the TDSE, and I already anticipated that the answer was yes. Now we see how this works: indeed, the reason I can always find c_n and $\phi(k)$ to solve (29) is that $\psi_n(x)$ and $\psi_k(x)$ form what is called a *complete set of functions*. This means that any (nice enough) function $\Psi_0(x)$ can be written as a linear superposition of them, precisely in the form (29). Mathematicians have shown this⁶, and we will see that in some specific cases as well during the course.

Remember that to solve (29) explicitly we need to know $\psi_n(x)$ and $\psi_k(x)$, which means we have made a choice of potential function $V(x)$ and solved (24). A very useful example to work out is the free particle, where $\psi_k(x) = e^{ikx}$. Equation (29) then becomes the **Fourier decomposition** of $\Psi_0(x)$:

$$\Psi_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \phi(k) e^{ikx}. \quad (30)$$

We say that $\phi(k)$ is the *Fourier transform* of $\Psi_0(x)$. Fourier theory now tells us, via Plancherel's theorem, how to find $\phi(k)$ for given $\Psi_0(x)$:

Plancherel's Theorem. Let $f(x)$ and $F(k)$ be two square integrable functions over the real line, i.e. the integrals $\int_{-\infty}^{\infty} dx |f(x)|^2$ and $\int_{-\infty}^{\infty} dk |F(k)|^2$ are *finite*. Then the following holds:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk F(k) e^{ikx} \Leftrightarrow F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx f(x) e^{-ikx}. \quad (31)$$

This condition precisely allows us to solve for $\phi(k)$ in (30): assuming that Ψ_0 and $\phi(k)$ are square integrable (in fact, since $|\Psi_0(x)|^2$ is a probability density, it should integrate to one!), we see that (30) is of the form of the l.h.s. of (31) if we identify $f(x) = \Psi_0(x)$ and $\phi(k) = F(k)$. Therefore Plancherel's theorem tells us that the r.h.s. is true:

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \Psi_0(x) e^{-ikx}. \quad (32)$$

⁶This is valid under certain conditions for the potential $V(x)$, which we will normally assume to be polynomial.

This is how step 3 looks like for a free particle. Given any boundary condition $\Psi_0(x)$ at $t = 0$, we find the unique infinite set of coefficients $\phi(k)$ that solve the boundary condition (29).

2.5 Step 4: Plug Back into $\Psi(x, t)$

In the previous step we found the coefficients c_n or $\phi(k)$ such that the boundary condition (29) is satisfied. We now have to plug this back into the general solution of the Schrödinger equation (26) or (27) to get the full solution. Of course, to obtain a completely explicit solution one has to carry out the summation or the integral.

Don't take the above four-step procedure as if it were written in stone. Sometimes you may simply be interested in the wave functions $\psi_n(x)$ and will not bother to calculate the coefficients c_n , sometimes step 4 is trivial because only a few terms contribute or sometimes you will directly jump to step 3 because you already know the wave functions.

2.6 Example: Gaussian Wave Function

We will now apply this procedure to the case of a free particle, which means $V(x) = 0$. We will take as our initial wave function a Gaussian distribution:

$$\Psi_0(x) = \left(\frac{2a}{\pi} \right)^{1/4} e^{-ax^2}. \quad (33)$$

Some comments about the physics before we start. This wave function represents a particle localized around $x = 0$ with standard deviation $\sigma = \frac{1}{2\sqrt{a}}$. So when a is large, the particle is well localized around $x = 0$, whereas if $a \rightarrow 0$, the probability distribution flattens and the particle becomes more and more delocalized in space. Using the basic Gaussian integral (85), we check that the wave function is indeed normalized to one: $\int_{-\infty}^{\infty} dx |\Psi_0(x)|^2 = 1$.

Step 1. The TISE (24) reduces to:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} = E \psi(x). \quad (34)$$

This equation is readily solved. The solutions are sinusoidal. We can write them as $\psi(x) = A \sin kx + B \cos kx$ or alternately as complex exponentials, as we saw in section 1.1. We will take exponentials:

$$\begin{aligned} \psi_k(x) &= e^{ikx} \\ k &= \pm \frac{\sqrt{2mE_k}}{\hbar} \Leftrightarrow E_k = \frac{\hbar^2 k^2}{2m}. \end{aligned} \quad (35)$$

The second line follows from filling the wave function into (34).

Step 2. We fill this into the general solution of the Schrödinger equation, which gives us the formula I wrote earlier:

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \phi(k) e^{i(kx - \frac{\hbar k^2}{2m} t)} . \quad (36)$$

Here, we took both the $k > 0$ and the $k < 0$ solutions into account⁷ by integrating from minus infinity to plus infinity with arbitrary coefficients $\phi(k)$.

Step 3. Impose initial condition $\Psi_0(x)$. We use Plancherel's theorem directly for the initial wave function at hand:

$$\begin{aligned} \phi(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \Psi_0(x) e^{-ikx} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \left(\frac{2a}{\pi} \right)^{1/4} e^{-ax^2} e^{-ikx} \\ &= \left(\frac{a}{2\pi^3} \right)^{1/4} \int_{-\infty}^{\infty} dx e^{-ax^2 - ikx} . \end{aligned} \quad (37)$$

The integral can once again be done using the basic Gaussian integral (85). To bring it to this form, we need to cancel the $-ikx$ in the exponent. This can be done by a change of variables. One readily checks that the following does the job⁸:

$$x = y - \frac{ik}{2a} . \quad (38)$$

One way to see why this trick works is to rewrite the term in the exponent as follows:

$$-ax^2 - ikx = -ax \left(x + \frac{ik}{a} \right) \equiv a(y^2 + C) . \quad (39)$$

We impose the last equality because we want to get a Gaussian integral, up to a constant C but with no linear term in x . From this form it is more or less obvious that we can complete the square as follows:

$$\begin{aligned} -ax \left(x + \frac{ik}{a} \right) &= -a \left(x + \frac{ik}{2a} - \frac{ik}{2a} \right) \left(x + \frac{ik}{2a} + \frac{ik}{2a} \right) = -a \left(y - \frac{ik}{2a} \right) \left(y + \frac{ik}{2a} \right) \\ &= -a \left(y^2 - \left(\frac{ik}{2a} \right)^2 \right) = -a \left(y^2 + \frac{k^2}{4a^2} \right) = -ay^2 - \frac{k^2}{4a} . \end{aligned} \quad (40)$$

In other words, to find the trick (38) I just add a constant to x such that in the end I get $(y - c)(y + c) = y^2 - c^2$. This way of reducing integrals of exponentials containing a

⁷As in the sinusoidal representation, the general solution for given k is actually $\psi_k(x) = Ae^{ikx} + Be^{-ikx}$. However, as mentioned in the text, the negative solution is automatically taken into account by the fact that we integrate over both positive *and* negative k . It is also unnecessary to include the normalization constant A , as overall normalizations are taken care of by $\phi(k)$ with which this wave function is multiplied (in other words, A can always be reabsorbed in $\phi(k)$).

⁸One should worry about the fact that this change of variables involves an *imaginary* shift of the integration variable, by a factor of $-\frac{ik}{2a}$. However, for real and positive a the integral is everywhere finite and it follows from complex analysis that this change of variables can be done.

Gaussian piece to a pure Gaussian is an important trick. We also have $dx = dy$ and the integration range is the same. Hence:

$$\begin{aligned}\phi(k) &= \left(\frac{a}{2\pi^3}\right)^{1/4} \int_{-\infty}^{\infty} dx e^{-ax^2 - ikx} = \left(\frac{a}{2\pi^3}\right)^{1/4} \int_{-\infty}^{\infty} dy e^{-ay^2 - \frac{k^2}{4a}} = \left(\frac{a}{2\pi^3}\right)^{1/4} e^{-\frac{k^2}{4a}} \sqrt{\frac{\pi}{a}} \\ \Rightarrow \phi(k) &= \left(\frac{1}{2\pi a}\right)^{1/4} e^{-\frac{k^2}{4a}}.\end{aligned}\quad (41)$$

To summarize the trick, we have derived the following generalized Gaussian formula:

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2 - i\beta x} = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}}. \quad (42)$$

Step 4. Plug back into $\Psi(x, t)$. Now we can plug this back into the solution of the TDSE (28):

$$\begin{aligned}\Psi(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \phi(k) e^{i(kx - \frac{\hbar k^2}{2m} t)} = \left(\frac{1}{8a\pi^3}\right)^{1/4} \int_{-\infty}^{\infty} dk e^{-\frac{k^2}{4a}} e^{i(kx - \frac{\hbar k^2}{2m} t)} \\ &= \left(\frac{1}{8a\pi^3}\right)^{1/4} \int_{-\infty}^{\infty} dk e^{ikx - k^2(\frac{1}{4a} + \frac{i\hbar t}{2m})}.\end{aligned}\quad (43)$$

This last integral is of the *same* type as the one before, it is a generalized Gaussian integral. The only difference is that we now integrate over k instead of x , but we can apply formula (42) all the same since x and k are dummy variables. We get (see exercise 3.2):

$$\Psi(x, t) = \left(\frac{2a}{\pi}\right)^{1/4} \frac{1}{\sqrt{1 + \frac{2ai\hbar t}{m}}} e^{-\frac{ax^2}{1 + \frac{2ai\hbar t}{m}}}. \quad (44)$$

This is the final result for the wave function for the given boundary condition (33). As you see, it is a completely explicit function of time. For its physical interpretation, see exercise 3.2.

3 One-Dimensional Potentials

3.1 General Theorems

There are some useful theorems that save you work when solving the time-dependent Schrödinger equation with one-dimensional potentials $V(x)$. These theorems are presented in problems 2.1 and 2.2 of [1]. I summarize them here.

1) The assumption that (whenever this is possible) solutions of the Schrödinger equation should be normalizable implies the following:

- a) $E \in \mathbb{R}$. That is, we can always take the energy to be positive.

b) $E > V_{\min}$. The energy is bounded from below.

Remark. Normalizable solutions do not always exist. For instance, plane waves $\psi(x) = e^{ikx}$ are not normalizable in the range $r \in (-\infty, \infty)$ because $\int_{-\infty}^{\infty} dx |\psi(x)|^2 = \int_{-\infty}^{\infty} dx = \infty$. But we use normalizable solutions whenever they exist. Whenever a non-normalizable solution is used, one must have a physical motivation for doing that. For instance, in the case of plane waves, it is not surprising that they are non-normalizable because they fill all of space. We regard them as approximations to more realistic situations that can be built by superpositions of plane waves. Superposing plane waves we *can* get a normalizable solution, as in (28). Remember that we always only need to normalize the *total* wave function $\Psi(x, t)$. In this case, plane waves are like ‘fundamental building blocks’ of more physical solutions of the TDSE and as such they are useful.

2) In the TISE, $\psi(x)$ can always be taken to be real. A complex solution of the TISE is always a linear combination of real solutions.

3) If the potential is even, $V(x) = V(-x)$, then the even and odd solutions of the TISE can be analyzed separately. In this case, any solution $\psi(x)$ of the TISE can be written as a linear combination of an even and an odd solution. This is useful when we consider bound states because it means that we can solve the Schrödinger equation to one side of the potential ($x > 0$, say) and we automatically obtain the other side by use of the symmetry. On the other hand, we often cannot apply this to scattering states, because there is an explicit breaking of the symmetry by the boundary conditions (which are generically different at plus and minus infinity). See the next section.

3.2 The TISE with one-dimensional potentials

As we saw in chapter 2, the first step in solving the TDSE is reducing to the TISE. Here we concentrate on this step. There are a few things to remember:

- **Normalizability** imposes $E > V_{\min}$ (1b).

Further, one distinguishes **bound states** and **scattering states** because their physical behavior is completely different and they give independent contributions to the final wave function $\Psi(x, t)$.

Also, we solve the Schrödinger equation separately in each region in which the potential is a continuous, differentiable function (a smooth function).

- **Bound states** have the following *generic* behavior which you should check:
 - ✓ Exponential behavior (damped/exponential growth) in the **exterior regions**.
 - ✓ The behavior in the **interior regions** (if there are any) depends on the details of the potential. It can be oscillatory (sine, cosine or plane wave) or exponential. See the example in Figure 2.
 - ✓ **Symmetric potential:** in that case, it is useful to separate even/odd solutions, which leads to sines and cosines (or hyperbolic sines and cosines) in the **interior region**. Use

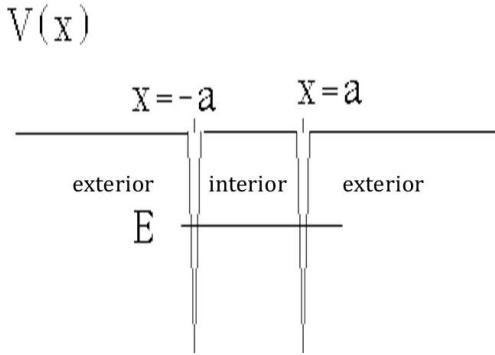


Figure 2: Potential symmetric around $x = 0$, with narrow spikes at $x = a$ and $x = -a$.

exponentials in the exterior regions.

✓ **Discrete spectrum:** in that case, the values of E are limited (sometimes a only a handful of solutions or none, or an infinite number of them).

✓ Solve the energy equation **graphically** to obtain the spectrum.

- **Scattering states** have the following *generic* properties:

✓ Oscillatory behavior. Use plane waves.

✓ Transmission and reflection coefficients are interesting quantities.

✓ There is an interesting physical interpretation in terms of waves with which one can probe a potential.

✓ Algebraic manipulations can be heavy if one has to match different regions.

- Impose **(dis-)continuity** conditions at each point at which the potential is non-differentiable (it has a kink or it is singular at those points):

✓ If the potential is everywhere **finite**, then ψ and $\frac{d\psi}{dx}$ are continuous. Such a potential can be non-smooth though.

✓ If the potential has **singularities** (points where it is infinite) then ψ is continuous but $\frac{d\psi}{dx}$ will have discontinuities at the singular points. The discontinuities are known and can be obtained by integrating the Schrödinger equation around the singular points, as in the steps leading to [2.125]. Obviously, an infinite potential is only an idealization of a very large one.

✓ If the potential is infinite in a whole region (more than a single point), the wave function must vanish there. For example: the infinite potential well.

4 Fourier Integrals and the Dirac Delta

4.1 The Dirac delta

Consider the following Gaussian function with unit area:

$$f(x) = \frac{1}{\sqrt{\pi\varepsilon}} e^{-\frac{(x-x_0)^2}{\varepsilon}}. \quad (45)$$

Using (85), one easily sees that $\int_{-\infty}^{\infty} dx f(x) = 1$. The standard deviation of this Gaussian is $\sigma = \sqrt{\varepsilon/2}$ (see Exercise 1.3b) of [1]), so clearly $\sqrt{\varepsilon}$ determines the width of the distribution and $f(x_0) = 1/\sqrt{\pi\varepsilon}$ its height.

Consider what happens when we make ε smaller and smaller; the peak clearly becomes higher and higher (in order for the function to preserve its area 1) and the distribution becomes narrower and narrower. In the limit, the height of the distribution is infinite and its width goes to zero, but the area stays equal to one. We can picture what happens in the limit as follows. Replace the Gaussian by a square of height $\sqrt{\varepsilon}$ and height $1/\sqrt{\varepsilon}$ such that the total area under the square is $\sqrt{\varepsilon} \times \frac{1}{\sqrt{\varepsilon}} = 1$, i.e. a finite limit. Now since in the limit the height of the distribution is infinite but the width goes to zero, we can think of a function $\delta(x)$ that takes the following values:

$$\begin{aligned} \delta(x - x_0) &= \begin{cases} 0 & \text{for } x \neq x_0 \\ \infty & \text{for } x = x_0 \end{cases} \\ \int_{-\infty}^{\infty} dx \delta(x - x_0) &= 1. \end{aligned} \quad (46)$$

This is called the Dirac delta distribution. It is called a distribution because it is not a proper mathematical function (it has an explicit divergence at one point; this divergence, however, is mild enough that it can be treated in the more general theory of distributions, which goes beyond the scope of these notes). Our treatment here has been mathematically heuristic, but sufficient. We argued that this function is the following limit:

$$\delta(x - x_0) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\sqrt{\pi\varepsilon}} e^{-\frac{(x-x_0)^2}{\varepsilon}}. \quad (47)$$

To check the integration property (46), we first perform the integral and then take the limit $\varepsilon \rightarrow 0$. It now also follows that for any function $f(x)$:

$$\int_{-\infty}^{\infty} dx \delta(x - x_0) f(x) = f(x_0). \quad (48)$$

From the fact that the Dirac delta vanishes everywhere except at $x = x_0$, we have: $\delta(x - x_0) f(x) = \delta(x - x_0) f(x_0) \forall x$. Now we can take $f(x_0)$ outside the integral, and evaluate the integral using (46). The result is (48).

The Dirac delta is paramount to many applications in physics, where we want to mimic situations with ‘spikes’. If we identify $t = \varepsilon$ as time, it turns out that (45) satisfies the heat equation. This Gaussian for instance describes heating up an aluminium bar. Initially

at $t = 0$, the temperature is very high at the point where the material is being heated ($x = x_0$), but as the material goes to thermal equilibrium the temperature quickly drops and the temperature profile looks like a blob centered around the origin but spreading in time, as heat exits the center. See Figure 3.

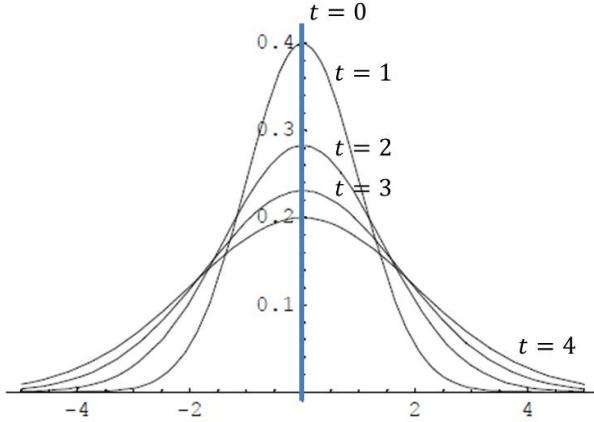


Figure 3: Heat distribution, all energy concentrated at $x = 0$ at $t = 0$, spreading in space as time increases according to a Gaussian distribution.

4.2 Fourier transformations

Plancherel's theorem (31) tells us that, for every square-integrable function $f(x)$, there is a unique square integrable function $F(k)$ from which $f(x)$ can be reconstructed. $F(k)$ is called the *Fourier transform* of $f(x)$. Its physical significance is that it gives an expansion of the function $f(x)$ in terms of plane waves e^{ikx} with a well-defined wavelength $\lambda = \frac{2\pi}{k}$. Thus, $f(x)$ is written as a superposition of (an infinite number of) monochromatic waves, one for each frequency. Fourier analysis is used in many branches of physics. Decomposing a sound signal into a spectrum of frequencies of sound is an example of Fourier analysis.

Filling the second of (31) into the first, we find the following relation:

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikx} \int_{-\infty}^{\infty} dx' f(x') e^{-ikx'} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dx' f(x') e^{ik(x-x')} \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx' f(x') \int_{-\infty}^{\infty} dk e^{ik(x-x')} . \end{aligned} \quad (49)$$

The integral shifts are allowed because we were careful about distinguishing x from the dummy variable x' that is integrated over (and anything to the *right* of the integral is integrated over).

Now comparing (49) with (48), we conclude that⁹

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x-x')} , \quad (50)$$

where in comparing the two we have relabeled $x \leftrightarrow x_0, x' \leftrightarrow x$.

Let us evaluate this expression. When $x = x'$ the integrand is one, and the integral diverges. When $x \neq x'$, we can evaluate the integral explicitly:

$$\delta(x - x') = \frac{1}{2\pi} \frac{e^{ik(x-x')}}{i(x - x')} \Big|_{-\infty}^{\infty} = 0 . \quad (51)$$

Hence this agrees with our previous definition (46).

The expression (50) we have obtained for the Dirac delta is a very useful one. Comparing it to Plancherel's theorem (31), taking $f(x - x') = \delta(x - x')$, we find from the left-hand side of (31) that:

$$F(k) = \frac{1}{\sqrt{2\pi}} . \quad (52)$$

In other words, the Dirac delta is the Fourier transform of a constant! This agrees with our intuition about position vs. momentum space: a function that is completely localized at some point in space corresponds, under Fourier analysis, to a constant function—i.e. a superposition of all wavelengths, all with the same amplitude.

5 The Formalism of Quantum Mechanics

5.1 Why?

If you have read a text on the formalism of quantum mechanics, you might have been left with the question: how is this useful? Why study it?

There are some important reasons why it pays off to learn quantum mechanics in this high-brow language of Hilbert spaces:

- 1) For the simple case of finite Hilbert spaces, the formalism simply reduces to the algebra of matrices (diagonalization of matrices, etc.).
- 2) In general, the formalism allows for a unified conceptual treatment of continuous and discrete cases by means of linear algebra.
- 3) The formalism is independent of the particular basis you choose. Describing your system in terms of its *positions* or in terms of *momenta* does not make any difference, just as you can describe a wave either by specifying its spatial profile or by specifying the elementary frequencies that the wave is built up from. It is a choice of *basis*, and the

⁹This equation certainly makes (49) and (48) compatible. But is this the unique solution? In fact, it is. Assume that the two sides of (50) differed by some function $g(x - x')$. This function would have to be such that $\int_{-\infty}^{\infty} dx' f(x') g(x - x') = 0 \forall f \in L^2(\mathbb{R})$. But the only such function is $g(x) = 0$. Hence the solution (50) is unique (even if for the reasons explained the Dirac delta is not a proper function).

formalism shows that quantum mechanics does not depend on a choice of basis.

4) It is conceptually simple: it allows to formulate the postulates of quantum mechanics in a clear and concise way, and in particular it allows for a natural introduction of the *measurement postulate*. Some approaches to quantum mechanics do not need a measurement postulate (or they claim they don't), but in the standard Copenhagen interpretation this is now the connection between theory and experiment is made.

5.2 The Postulates of Quantum Mechanics

Von Neumann gave an axiomatic formulation of quantum mechanics, similarly to what Einstein had done for special relativity. The great advantage of this approach is that, if you want to generalize the theory, all you have to do is modify one or several of its postulates. Similarly, if something turns out to be 'wrong' with the theory, the axiomatic structure makes it easier to trace the inconsistency back to one or several of the axioms. By explicitly including a projection postulate we are not only making it possible to interpret measurements, but also making explicit something that might turn out to be a weakness of the theory—rather than hiding it under the rug of the theoretical machinery. So here are the postulates¹⁰ [3]:

Postulate 1. State of the System:

An ensemble of physical systems is completely described by a *wave function* or *state function*. The wave function lives in a Hilbert space, and it may be multiplied by an arbitrary complex number without altering its physical significance.

Postulate 2. Observables:

Observables are represented by Hermitian (or self-adjoint) operators, $\hat{Q} = \hat{Q}^\dagger$.

Justification: Hermiticity is the natural notion of 'reality' for operators. An Hermitian operator has real eigenvalues hence real expectation values (see the next postulate).

Postulate 3. Measurement postulate:

The only result of a precise measurement of \hat{Q} is one of the eigenvalues q_n of \hat{Q} .

Justification: when we measure any quantity in the laboratory, we never observe an actual superposition¹¹. We always find a unique value for the observed outcome of the experiment—the energy, say. Since the eigenvalues are well-defined real numbers associated with an Hermitian operator, they are good candidates for measurement outcomes. The probabilities that, upon measurement, we find q_n , are given by $|c_n|^2$.

¹⁰Different authors sometimes rank the postulates in a different order.

¹¹Occasionally you will find a text, even in prestigious journals, where it is claimed that a superposition has been *observed*. Beware of such *metaphorical* phrasing! So far no superposition has been observed as an *actual* experimental outcome. We observe interference patterns after repeating an experiment many times, from which we *infer* that the state of the system was a superposition. The difference is subtle, but important. The detector either clicks, or it doesn't!

Postulate 4. Schrödinger postulate:

The time evolution of the system is governed by the Schrödinger equation.

Postulate 5. Projection postulate:

After a precise measurement of \hat{Q} with outcome q_n , the system is (shortly after measurement) in the state $|\psi_{q_n}\rangle$.

The justification of this postulate is that repeated measurements of \hat{Q} should give the same result if the time interval between them is small. The only state with $P(q_n) = 1$ is $|\psi_{q_n}\rangle$ itself, therefore after measurement the system must be in state $|\psi_{q_n}\rangle$.

This postulate introduces what we call the ‘collapse of the wave function’. If prior to measurement the system is in the state:

$$|\Psi(t)\rangle = \sum_n c_n(t) |\psi_{q_n}\rangle , \quad (53)$$

and after measurement it is in state:

$$|\Psi\rangle = |\psi_{q_n}\rangle , \quad (54)$$

we see that we have projected onto one of the components of (53) and lost all of the information about the c_n ’s prior to measurement. We can neither predict with certainty which state we will find upon measurement, nor can we, after measurement, use the Schrödinger equation to ‘trace back’ the state (53).

5.3 Linear Algebra

In the table below I summarize the main concepts from linear algebra that are used in the description of quantum systems, including a detailed comparison with inner products in \mathbb{C}^N , with their respective notations:

linear algebra in \mathbb{C}^N	wave function	Hilbert space
vector: $ \alpha\rangle$ basis \mathbf{e}_i :	$\Psi(x, t)$	$ \Psi(t)\rangle$
$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_N = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$	$\psi_n(x)$ ($n = 1, 2, \dots, \infty$)	$ \psi_n\rangle$ or $ n\rangle$
vector $ \alpha\rangle$ in a basis:		
$\alpha = \sum_{i=1}^N \mathbf{e}_i \alpha_i = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix}$	$\Psi(x, t) = \sum_n c_n(t) \psi_n(x)$	$ \Psi(t)\rangle = \sum_n c_n(t) \psi_n\rangle$
inner product: $\langle \alpha \beta \rangle = \sum_{i=1}^N \alpha_i^* \beta_i$	$\int_a^b dx \Psi_1^*(x) \Psi_2(x)$	$\langle \Psi_1 \Psi_2 \rangle$
orthonormal basis: $\langle e_i e_j \rangle = \delta_{ij}$	$\int_a^b dx \psi_n^*(x) \psi_m(x) = \delta_{nm}$	$\langle \psi_n \psi_m \rangle = \delta_{nm}$
linear transformation (matrix):	operator \hat{Q} :	
$\beta = T \cdot \alpha$	$\psi_1 = \hat{Q} \Psi_2$	$ \Psi_1\rangle = \hat{Q} \Psi_2\rangle$
coefficients in a basis: $\alpha_i = \langle e_i \alpha \rangle$	$c_n(t) = \int_a^b dx \psi_n^*(x) \Psi(x, t)$	$c_n(t) = \langle \psi_n \Psi(t) \rangle$
dual vector: $\langle \alpha $	$\Psi^*(x, t)$	$\langle \Psi(t) $
dual basis:		
$\mathbf{e}_i^T = (0, \dots, 0, 1, 0, \dots, 0)$ (1 on i th place)	$\psi_n^*(x)$	$\langle \psi_n $
vector in dual basis:		
$\alpha^T = \sum_{i=1}^N \mathbf{e}_i^T \alpha_i^* = (\alpha_1^*, \dots, \alpha_N^*)$	$\Psi^*(x, t) = \sum_n c_n^*(t) \psi_n^*(x)$	$\langle \Psi(t) = \sum_n c_n^*(t) \langle \psi_n $

The Hilbert space considered here is $L_2(a, b)$, that is the square integrable functions between (a, b) :

$$\int_a^b dx |\Psi(x)|^2 < \infty \quad (55)$$

where a, b can be finite or infinite $(-\infty, \infty)$.

5.4 Continuous spectra

With appropriate modifications, the above formulas also hold in the case of continuous spectra. For definiteness, we will consider the momentum operator $\hat{p} = -i\hbar d/dx$ and the associated wave functions of a free particle moving along $x \in (-\infty, \infty)$. This is a continuous spectrum labeled by $p = \hbar k$. The eigenfunctions of \hat{p} are:

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} px}, \quad (56)$$

where the reason for the normalization will become clear in a moment.

The above wave-functions are not normalizable because $\int_{-\infty}^{\infty} dx |\psi_p(x)|^2 = \infty$. However, as long as we consider different p 's, we do have the following orthonormality condition:

$$\int_{-\infty}^{\infty} dx \psi_{p'}^*(x) \psi_p(x) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx e^{i(p-p')x/\hbar} = \delta(p - p') , \quad (57)$$

where in the last equality we used the representation of the Dirac delta function introduced in (50), after interchanging x and k and introducing $p = \hbar k$ (this also explains the reason for the $1/\sqrt{\hbar}$ in the normalization of (56)). So we get:

$$\langle \psi_{p'} | \psi_p \rangle = \delta(p - p') , \quad (58)$$

which is analogous to the orthonormality condition in the discrete case.

Let us generalize this picture to the eigenfunctions of an arbitrary observable \hat{Q} . Let us assume that the eigenvalues are labeled by $q(z)$, where z is a continuous variable ($z = k$ and $q(z) = \hbar k$, in the previous example). They come from solving the eigenvalue problem:

$$\hat{Q} \psi_z(x) = q(z) \psi_z(x) . \quad (59)$$

The probability of finding a result $q(z)$ in the range $(z, z + dz)$ at time t is then given by $|c(z, t)|^2 dz$, where:

$$\begin{aligned} \Psi(x, t) &= \int dz c(z, t) \psi_z(z) \\ c(z, t) &= \langle \psi_z | \Psi \rangle . \end{aligned} \quad (60)$$

Notice that $c(z, t)$ contains as much information as $\Psi(x, t)$ does. There is a one-to-one correspondence between the two which is brought out by Plancherel's theorem. In the specific case $z = k$, then, $c(z, t)$ has a special name, $\Phi(p, t)$, and it is called the ‘wave function in momentum space’:

$$\begin{aligned} \Psi(x, t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp e^{ipx/\hbar} \Phi(p, t) \\ \Phi(p, t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx e^{-ipx/\hbar} \Psi(x, t) . \end{aligned} \quad (61)$$

Examples

a) If we take $\hat{Q} = H$ for a free particle, then $z = p$ (the momentum) and $q(z) = E(p) = \frac{p^2}{2m}$. So these wave functions are the usual plane waves, $\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$. Alternatively, we use the wave number instead of the momentum, $z = k$ and $E(k) = \frac{\hbar^2 k^2}{2m}$. The result is the same.

b) If we take $\hat{Q} = \hat{p}$ for a particle allowed to move with any momentum p (for instance, a free particle), then the eigenvalue equation is $\hat{p} \psi_p(x) = p \psi_p(x)$. Hence $z = p$ and the eigenvalues are the momenta themselves, $q(z) = p$. Hence also $\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$.

c) If we take $\hat{Q} = \hat{x}$, the eigenvalue equation is $\hat{x} \psi_y(x) = y \psi_y(x)$. Then $z = y$ and $q(z) = y$. The solutions of the eigenvalue equation are delta functions, $\psi_y(x) = \delta(x - y)$.

Once we have the basis $\psi_z(x)$ that solves (59), we can expand the wave function in this basis:

$$\Psi(x, t) = \int dz c(z, t) \psi_z(x) . \quad (62)$$

In the above examples, this formula reproduces known results:

- a) $\Psi(x, t) = \int dp c(p, t) \psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} \int dp c(p, t) e^{ipx/\hbar}$ which is the Fourier transform (61).
- b) Idem.
- c) $\Psi(x, t) = \int dy c(y, t) \psi_y(x) = \int dy c(y, t) \delta(x - y) = c(x, t)$. This is a ‘diagonal’ basis: the expansion of $\Psi(x, t)$ contains just one term, namely $c(x, t) = \Psi(x, t)$ itself.

6 Dirac Notation

6.1 Base-free notation

In the above examples we have been expanding the same wave function in a different basis of wave functions (eigenfunctions of some Hermitian operator). The Hilbert space formalism allows us to do this in a basis-free way, as announced. To that end we introduce the vector $|\mathcal{S}(t)\rangle \in \mathcal{H}$, where we now think of the Hilbert space as an abstract space without the need to specify a basis. The only requirement is that $|\mathcal{S}(t)\rangle$ satisfies the TDSE. We define for this vector space the $L_2(-\infty, \infty)$ inner product:

$$\langle \mathcal{S}(t) | \mathcal{S}'(t) \rangle = \int_{-\infty}^{\infty} dx \Psi^*(x, t) \Psi'(x, t) , \quad (63)$$

where $\Psi(x, t)$ is the wave function that corresponds to the state $|\mathcal{S}(t)\rangle$ *in the position representation*. We can show, using Fourier transformation, that the above inner product is independent of the basis. Fill in the first of (61), then we can rewrite the above inner product as (try to show this):

$$\langle \mathcal{S}(t) | \mathcal{S}'(t) \rangle = \int_{-\infty}^{\infty} dp \Phi^*(p, t) \Phi'(p, t) . \quad (64)$$

In the discrete case (for example, the harmonic oscillator), we find:

$$\langle \mathcal{S}(t) | \mathcal{S}'(t) \rangle = \sum_n c_n^*(t) c'_n(t) . \quad (65)$$

So this inner product is independent of the basis. Using this, we also find:

$$\begin{aligned} \langle \psi_x | \mathcal{S}(t) \rangle &= \int_{-\infty}^{\infty} dy \psi_x^*(y) \Psi(y, t) = \Psi(x, t) \\ \langle \psi_p | \mathcal{S}(t) \rangle &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx e^{-ipx/\hbar} \Psi(x, t) = \Phi(p, t) \\ \langle \psi_n | \mathcal{S}(t) \rangle &= c_n e^{-\frac{i}{\hbar} E_n t} = c_n(t) . \end{aligned} \quad (66)$$

This gives a nice interpretation: $\Psi(x, t)$, $\Phi(x, t)$ and $c_n(t)$ are the overlaps of the abstract wave function $|\mathcal{S}(t)\rangle$ with the respective basis, i.e. the projections of that vector onto a particular basis. These bases are often simply denoted $|x\rangle$, $|p\rangle$, $|n\rangle$, hence we can write:

$$\begin{aligned}\Psi(x, t) &= \langle x | \mathcal{S}(t) \rangle \\ \Phi(p, t) &= \langle p | \mathcal{S}(t) \rangle \\ c_n(t) &= \langle n | \mathcal{S}(t) \rangle .\end{aligned}\tag{67}$$

In other words, these are the coefficients $c(z, t)$ that determine ‘how much $q(z)$ is in $|\mathcal{S}(t)\rangle$ ’ (for the continuous/discrete case).

With this new notation, we can also write: $\psi_p(x) = \langle x | \psi_p \rangle = \langle x | p \rangle$ and $\psi_y(x) = \langle x | \psi_y \rangle = \langle x | y \rangle$. This makes the notation much more symmetric and simple. $\psi_p(x)$ is simply the overlap $\langle x | p \rangle$, ‘the amount of x in the state $|p\rangle$ ’. $\langle x | y \rangle$ is ‘the amount of x in y ’, given by a Dirac delta function.

Now it is clear that we also have: $\psi_x(p) = \langle p | x \rangle = (\langle x | p \rangle)^* = \psi_p^*(x)$, as it should (notice that in Plancherel’s theorem we have $+ipx$ in one exponential, $-ipx$ in the other).

The overlaps $\langle x | p \rangle$ contain the information about the spectrum of the particular operator. For a free particle, $\langle x | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$. In other cases, however, this can take a completely different form.

Examples

a) Harmonic oscillator: in this case, the spectrum of the Hamiltonian is discrete, $E_n = \hbar\omega(n + \frac{1}{2})$ whereas the spectrum of \hat{x} is continuous and infinite, $x \in (-\infty, \infty)$. The overlaps are now $\langle x | n \rangle = \psi_n(x)$ (the harmonic oscillator wave function).

b) Free particle on a circle, $\phi \in [0, 2\pi]$. Here, the spatial coordinate is continuous ($x = a\phi$ with a the radius of the circle) but momenta are quantized, $p_n = \frac{n\pi}{a}$. The overlaps are: $\langle \phi | n \rangle = \frac{1}{\sqrt{2\pi}} e^{in\phi}$.

6.2 Closure

Since now we have a basis-free notation, we can notice that, for $f, g \in \mathcal{H}$, the following holds:

$$\langle f | g \rangle = \int dx f^*(x) g(x) = \int dx \langle f | x \rangle \langle x | g \rangle = \langle f | \int dx |x\rangle \langle x | g \rangle .\tag{68}$$

In the last step, we pulled the integral inside the inner product because f is base-free, i.e. it does not depend on x ; as well we regard $\int dx |x\rangle \langle x | g \rangle = \int dx g(x) |x\rangle$ as a new vector which we obtain by multiplying the vector $|x\rangle$ with the number $g(x) = \langle x | g \rangle$ and integrating over x . Now let us define:

$$\langle f | \hat{Q} | g \rangle \equiv \langle f | \hat{Q} g \rangle .\tag{69}$$

(68) can now be written as:

$$\langle f | g \rangle = \langle f | \left(\int dx |x\rangle \langle x| \right) |g\rangle .\tag{70}$$

We can regard the part in the middle, $\int dx |x\rangle\langle x|$, as an operator that acts on $|g\rangle$ and produces another vector. Since this equation holds for all $f, g \in \mathcal{H}$, this operator must be the identity:

$$\int dx |x\rangle\langle x| = \mathbb{1}. \quad (71)$$

This is called the **closure relation** and indicates that the basis is complete. Multiplying on the right and left with $|p\rangle$ and $\langle p'|$, this gives:

$$\langle p' | \left(\int dx |x\rangle\langle x| \right) |p\rangle = \int dx \psi_{p'}^*(x) \psi_p(x) = \langle p' | p \rangle = \delta(p - p'), \quad (72)$$

which is indeed true for instance for a free particle. If we multiply with $|y\rangle$, $\langle y'|$ instead, we get: $\int dx \psi_{y'}^*(x) \psi_y(x) = \int dx \delta(x - y') \delta(x - y) = \delta(y - y')$ which is again true.

In the same way we can derive closure for p :

$$\int dp |p\rangle\langle p| = \mathbb{1}. \quad (73)$$

Multiplying both sides with $|x\rangle$, $|x'\rangle$, this is the same as $\int dp \psi_{x'}^*(p) \psi_x(p) = \delta(x - x')$.

6.3 Bras and kets

Something funny has happened. The notation (69) gave rise to operators such as $\int dx |x\rangle\langle x|$. These operators act on a state to give a new state. But what do we really mean by $\langle x|$ or, more generally, $\langle \alpha|$? The definition of these states was given in (69). We can in fact think of $\langle \Psi|$ as the ‘complex conjugate’ of $|\Psi\rangle$. For obvious reasons, $|\Psi\rangle$ is called a ‘ket’ and $\langle \Psi|$ is called a ‘bra’, so that we can form an inner product—a bra-ket—by multiplying the two as in a dot product. Here our analogy with \mathbb{R}^3 and \mathbb{C}^3 comes in handy. In usual \mathbb{R}^3 the standard inner product is in fact the dot product:

$$\langle w|v \rangle = w_x v_x + w_y v_y + w_z v_z = (w_x, w_y, w_z) \cdot \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}. \quad (74)$$

In \mathbb{C}^3 :

$$\langle w|v \rangle = w_x^* v_x + w_y^* v_y + w_z^* v_z = (w_x^*, w_y^*, w_z^*) \cdot \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}. \quad (75)$$

In the same way, via (69) (with $Q = \mathbb{1}$, the identity operator), we regard $\langle \Psi_1|\Psi_2\rangle$ as the ‘dot product’ of $\langle \Psi_1|$ and $|\Psi_2\rangle$: $\langle \Psi_1| \cdot |\Psi_2\rangle \equiv \langle \Psi_1| \mathbb{1} |\Psi_2\rangle = \langle \Psi_1|\Psi_2\rangle$. The bra corresponds to the (complex conjugate) row vector, the ket to the column vector. They are each other’s ‘duals’.

Remark (optional). One can show that the bras form a vector space, in the same way

the kets $|\alpha\rangle$ do. In fact, this is the vector space *dual* to \mathcal{H} , and it is often denoted \mathcal{H}^* . By definition, the dual vector space \mathcal{H}^* of a vector space \mathcal{H} is a linear (non-degenerate) mapping from $\mathcal{H} \rightarrow \mathbb{C}$. Indeed, $\langle\alpha|$ maps vectors in \mathcal{H} to \mathbb{C} by the inner product: $\langle\alpha|\beta\rangle \in \mathbb{C}$. This mapping forms itself a vector space.

Now when we put together $|\alpha\rangle\langle\beta|$, this is an operator as it maps a vector $|\gamma\rangle$ to another vector $(|\alpha\rangle\langle\beta|)|\gamma\rangle = |\alpha\rangle\langle\beta|\gamma\rangle = \langle\beta|\gamma\rangle|\alpha\rangle$. Here, $\langle\beta|\gamma\rangle$ is a number so it doesn't matter if you write it left or right. On the other hand, the order does matter in $|\alpha\rangle\langle\beta|$: $|\alpha\rangle\langle\beta| \neq \langle\beta|\alpha\rangle$!

7 The Interpretation of Quantum Mechanics

We have seen that quantum mechanics gives us a *probabilistic interpretation* of physical quantities: it is not always possible to determine the outcome of a measurement with absolute certainty, but we can predict the possible measurement outcomes and their respective probabilities in measurements on an ensemble of identical systems. The only case where the outcome of a measurement is unique is when the system is in a *determinate state* and we measure the corresponding observable. If the system is in an eigenstate of an operator \hat{Q} with eigenvalue q_m , the formalism tells us that the probability to find the value q_n upon measurement is 1 for $n = m$ and 0 for all other states. If \hat{Q} is the position operator, then the particle is well localized; but the momentum is not well defined in that case, as the uncertainty principle expresses. This is because the operators \hat{x} and \hat{p} do not commute. So the formalism tells us that there is no state that describes all possible observables simultaneously, i.e. no state in which all possible measurable quantities have well-defined values.

We can get used to the above interpretation of quantum mechanics, but there is something unsatisfactory about it. Imagine carrying out a two-slit experiment with photons. Given knowledge of the initial wave function and of the interactions, we can predict the intensity distribution of light on the detection screen. If we decrease the intensity of the source so that only one photon at the time goes through the slits, we can predict the probability that that detection of the photon will take place in a particular region on the screen. But there is something counterintuitive about this. Which slit did the particle *actually* go through? Quantum mechanics does not tell us the answer. If we try to measure which slit the photon goes through, the interference pattern disappears—knowledge of which slit the particle went through destroys the quantum superposition. How can the act of measurement be decisive here? Is a measurement any different from other physical interactions? If the photon did not have a position before it was measured, but it does have a well-defined position when it is detected, it would seem as if the act of measurement is able to give particles properties they did not possess before. How can a measurement be so decisive as to the *presence* of a physical property? What makes it so unlike other physical interactions?

This set of questions, and their proposed solutions, which we will next turn to, is what people usually call the problem of the ‘interpretation of quantum mechanics’¹². It is seen

¹²In writing this chapter, I have drawn ideas from [4]

as problematic because it defies our classical intuitions about how physical systems are supposed to work.

7.1 EPR and Hidden Variables

In a 1935 article that was meant to be a death blow on quantum theory as a fundamental theory of reality, Einstein, Podolsky, and Rosen claimed that quantum mechanics was incomplete. Imagine a pair of particles that are initially in contact, at rest at the origin, and then go separate ways until their mutual distance is very large. For instance, you could think of an isotope that decays into two sub-particles that shoot off to different sides, preserving the total momentum. Notice that, quantum mechanically, we cannot determine the position and momentum of the individual particles because $[\hat{Q}_1, \hat{P}_1] = i\hbar$, $[\hat{Q}_2, \hat{P}_2] = i\hbar$. However, we *can* determine the center of mass position of the system $\hat{Q}_1 + \hat{Q}_2$ as well as the relative momenta $\hat{P}_1 - \hat{P}_2$ because these operators do commute: $[\hat{Q}_1 + \hat{Q}_2, \hat{P}_1 - \hat{P}_2] = [\hat{Q}_1, \hat{P}_1] - [\hat{Q}_2, \hat{P}_2] = i\hbar - i\hbar = 0$ (where we used the fact that the operators of the two particles commute). Now imagine that we measure the position of the first particle, q_1 , once they are far apart. Since we know the location of the center of mass, we can infer from this measurement q_2 , the location of the second particle. But if we simultaneously measure p_2 , the momentum of the second particle, we may also infer p_1 in virtue of the fact that we know the difference in momenta. But now, EPR concluded, since the two particles are very far apart, one measurement cannot influence the other. This would violate the result of relativity that there can be no influences that travel faster than the speed of light (in Einstein's words, there can be no 'spooky action at a distance'). Hence we can, by means of independent measurements, have complete knowledge of the positions and momenta of the two particles. Since the outcome of the location of particle 1 cannot affect (by relativity) the location of particle 2, this means that particle 2 must have had a well-defined position prior to measurement even though quantum mechanics tells us it didn't. A similar argument can be made for the momentum. Of course this is all in contradiction with the predictions of quantum mechanics. So measurements can give us complete knowledge about these properties of particles, but the theory doesn't. Therefore, EPR concluded, quantum mechanics is an incomplete theory.

Before trying to rebut the EPR argument, it is natural to ask the following question: could it be that quantum mechanics is indeed missing some piece of information? Is it possible to extend quantum mechanics into a more predictive theory? Such attempts are called 'hidden variable theories': assuming that the position and momenta of particles have precise values *before* measurement amounts to introducing some variable that remains hidden to quantum mechanics, but determines those values before we measure them. One such attempt was carried out by David Bohm in 1952. In Bohm's theory, particles have well-defined values of positions and momenta, and the predictions of quantum mechanics are statistically reproduced. However, in order to achieve this, Bohm has to add a *non-local* interaction potential between the particles. This potential is called non-local because it acts at a distance; when a particle changes momentum or position, its interaction with all other particles changes. Whereas it is not clear whether there is a contradiction with special relativity (as the interactions assumed by Bohm cannot be directly used to transmit information faster than light), it is clearly an unwarranted feature of the theory.

For this reason, theorists have looked for *local* hidden variable theories, that is, hidden variable theories where interactions do not propagate faster than light. The main criticism to Bohm's theory, however, has been that it merely adds theoretical structure without any predictive gain, as its predictions are compatible with quantum mechanics and all experimental results agree with quantum mechanics. Furthermore, in order for the theory to reproduce the results of quantum mechanics, ad hoc assumptions about the distribution of the hidden variables have to be added.

In 1964 John Bell showed that local hidden variable theories are inconsistent with the results of quantum mechanics. That is, that any hidden variable theory of the local type makes experimental predictions which subtly deviate from those of quantum mechanics. The corresponding experiments were carried out in 1982 by Alain Aspect and the predictions of quantum mechanics, rather than those of hidden variable theories, were found to hold. These results have been confirmed in more refined experiments later on.

Hidden variable theories are sometimes identified with *realist* interpretations (see e.g. section 1.2 of [1]). This is a misnomer—whereas the hidden variable theory can certainly be regarded as an (extreme) realist position, there are realist positions that do not require particles to have well-defined properties (see the next section).

7.2 Bohr's Reply to EPR

Bohr's reply to the EPR article is all but a clear-cut physical argument. Instead, it is a piece of philosophical discourse (and rather obscure at that) the main message of which seems to be that what we call 'position' and 'momentum' cannot be determined a priori, but essentially depends on the *measurement context* in which these concepts are defined. The measurement process has, in Bohr's view, an essential influence on the conditions for the definition of physical quantities. Since the conditions of measurement play an essential role in defining what we call the 'physical reality' (and, as part of that, the concepts of position and momentum), one cannot infer a conclusion about the supposed incompleteness of quantum mechanics: as far as the physical phenomena are concerned, there simply *is* nothing else to describe than either position *or* momentum (but not both).

In more practical terms, one could say that Bohr's position amounts to saying that the particular experimental *context* determines whether the concepts 'position' or 'momentum' make sense. If we measure the position, there is no sense in which we can meaningfully talk about the 'momentum' of a particle: this concept is simply not defined.

Second, whereas Bohr denies that there are 'spooky actions at a distance', he remarks that EPR's conclusion that one can infer simultaneously the position and the momentum of the particle is incoherent. Once position is measured on the first particle, this concept is applicable to the second particle as well. A measurement of the momentum of the second particle creates a new measurement context which automatically introduces an uncertainty in the position, which is now no longer applicable. We have therefore not determined the position and momentum of particle two, but simply measured uncorrelated properties of the second particle.

By 1927, Bohr's position had blended with the views of his younger colleagues Heisenberg, Pauli, and Dirac, into the dominant paradigm in the interpretation of quantum mechanics, known as the 'Copenhagen interpretation'. The fact that Bohr's texts on this

matter are rather oracle-like and obscure and that his pupils developed related, but not entirely coinciding accounts of quantum mechanics¹³, have contributed to a lack of clarity as to how ‘the’ Copenhagen interpretation should be understood.

Differences in interpretation between Bohr, Heisenberg, Pauli, and Dirac notwithstanding, the Copenhagen interpretation does seem to offer a genuine reply to the EPR argument—this being the reason that it is widely accepted. The interpretation, however, does not come without a cost. At least two issues in the Copenhagen interpretation require further thought:

1. As a philosophical interpretation, it talks about the *measurement context* as posing restrictions on the class of macroscopic concepts (such as position, momentum, etc.) that can be applied to a microscopic system at any given time. However, the philosophical interpretation does not tell us how this should work. This necessity of invoking the macroscopic context in defining the concepts that are applicable to the microscopic phenomena, whilst upholding that the quantum mechanical description is complete, seems to imply the existence of a fundamental boundary between the macroscopic and the microscopic. It is not at all clear what physical principle, if any, defines this boundary.
2. Bohr’s sketched solution to the EPR paradox does not imply action at a distance, but does require an explanation of how measurement of particle 1 influences the applicability of the concept of ‘position’ to particle 2, and the actual determination of this position. So measurement still seems to play a special role here—it is not simply regarded as an ordinary physical interaction.

7.3 The Measurement Problem

When trying to translate Bohr’s philosophical account of quantum mechanics into a physically workable model, one runs, as mentioned, into the fact that it seems hard to regard measurement as an ordinary physical interaction. Indeed, in Bohr’s view, a measurement does much more than simply ‘giving the value of a variable’: it sets the conditions under which one can meaningfully talk about that variable. In order to get a grasp of how deep this problem runs, we will here give a model of classical and quantum measurements and compare them to each other¹⁴.

Measurement in Classical Mechanics

Consider a system S (for instance, you), a measuring device M (a scale to measure your weight) and the readings R on this device (the value of the scale pointer). Consider also a quantity A corresponding to S (your mass) with possible values $a \in \{a_1, \dots, a_n\} \in \mathbb{R}$ (a_i

¹³For instance, Heisenberg used the Aristotelian concepts of act and potency in his accounts of the measurement problem. This is quite different from the Kantian flavor of Bohr’s remarks, as well as from the instrumentalist interpretation that the ‘Copenhagen interpretation’ has often been given.

¹⁴This model goes back to von Neumann, but I am very much indebted to Jos Uffink for the formulation presented here.

for short). Corresponding to these values there are readings r_i on the scale which indicate your weight. If the scale works properly, then there is an invertible function m that gives the reading r_i for given mass a_i : $r_i = m(a_i)$.

You are standing in front of the scale. Since there is nothing on the scale, the pointer is in its rest position, which we call r_0 . We have a pair of numbers describing this situation, your weight a_i , and the flat scale pointer: (a_i, r_0) . Once you step on the scale, these numbers change. Your weight does not change, but the pointer position does: we have a pair $(a_i, r_i) = (a_i, m(a_i))$. Now the values of a_i and r_i are correlated via the function m . Reading off your weight r_i from the scale, and applying the function m^{-1} to it, you find your mass: $a_i = m^{-1}(r_i) = 70$ kg. Notice that, as an abstract property about you, your mass is unmeasurable; we measure your weight on the scale and use it to infer the mass.

Measurement in Quantum Mechanics

Consider now a quantum system S with a property \mathcal{A} (e.g. energy, or spin) to be measured by a measuring apparatus M (a detector, a phosphor plate, a photon camera). To \mathcal{A} corresponds an operator A with eigenvalues $\{a_1, \dots, a_n\}$. To these eigenvalues correspond states $|a_1\rangle, \dots, |a_n\rangle \in \mathcal{H}_S$, the Hilbert space of the system. Since we want to model measurement physically, we are going to associate an operator R to the measuring apparatus. R gives the possible ‘readings’ of the detector (numbers on a computer screen, dots on a phosphor surface, etc.), which can take values r_0, r_1, \dots, r_n (the value r_0 denoting, as before, ‘no detection’). So the Hilbert space \mathcal{H}_M of the measuring apparatus contains one more basis state: $|r_0\rangle, |r_1\rangle, \dots, |r_n\rangle$. We regard $|r_0\rangle$ as the ground state, the state that indicates that nothing is being measured. We assume $|a_i\rangle$ ($i = 1, \dots, r$) and $|r_i\rangle$ ($i = 0, 1, \dots, n$) to form orthonormal bases of \mathcal{H}_S and \mathcal{H}_M , respectively.

Before detection, S is in a state $|a_i\rangle$ (for some i) and M in the state $|r_0\rangle$. So the total state of the system is the product of states $|a_i\rangle|r_0\rangle$ —this is the quantum analog of the statement that you and the scale are described by the pair (a_i, r_0) .

Now we want to measure A using R . By the measurement interaction, the system M will undergo a change from $|r_0\rangle$ to $|r_i\rangle$, and the latter should be indicative for the state $|a_i\rangle$ of S . For an ideal measurement, the system S itself should not change. So we have the transition:

$$|a_i\rangle|r_0\rangle \rightarrow |a_i\rangle|r_i\rangle . \quad (76)$$

Our task of finding out ‘whether measurement is a physical interaction’ now amounts to asking whether there is a Hamiltonian H that describes (76) as a transition from an initial state $|a_i\rangle|r_0\rangle$ to a final state $|a_i\rangle|r_i\rangle$. In order to eliminate technical details, we will go about this question as follows. It turns out that this can be translated into the question whether there is a linear operator¹⁵ U such that:

$$U(|a_i\rangle|r_0\rangle) = |a_i\rangle|r_i\rangle . \quad (77)$$

We can think of this operator as effecting the Schrödinger time evolution with Hamiltonian H . If U exists, there is a Hamiltonian H that contains the interaction between the system S and the measuring apparatus M .

¹⁵As it turns out, such an operator also has to be unitary, meaning $UU^\dagger = \mathbb{1}$.

In fact, the answer to this question is: yes, such an operator U exists (see the appendix for more details). In principle this is very nice, because it means that:

1. We have succeeded in describing measurement as a physical interaction via a unitary operator.
2. No measurement postulate is needed.
3. There is no distortion of the state of the system a_i .

However, the fact that U is linear has some strange implications. The above really only works if S is in an eigenstate $|a_i\rangle$ of A , as we have assumed above. What if the system S is in a superposition?

$$|\psi\rangle = \sum_j c_j |a_j\rangle . \quad (78)$$

Let us apply our operator to this state. After measurement, we get a final state:

$$|\psi_{\text{final}}\rangle = U(|\psi\rangle|r_0\rangle) = U\left(\sum_j c_j |a_j\rangle|r_0\rangle\right) = \sum_j c_j U(|a_j\rangle|r_0\rangle) = \sum_j c_j |a_j\rangle|r_j\rangle , \quad (79)$$

where we have used linearity of U . You see the important consequence—the system is no longer in a product state $|a_i\rangle|r_i\rangle$ with definite value of i (corresponding to *one* eigenvalue of A) after the measurement! Instead, the final state is a *superposition of the measured system and the measuring apparatus!* This means that you and the scale are forever entangled after you step on it! So which state is this system in? We simply cannot tell: it might be any of $|a_j\rangle|r_j\rangle$ with probabilities $|c_j|^2$.

Could we solve this problem by bringing in another measuring device that will find out what state you and the scale are in? It is not hard to see that that second measuring device will again become entangled with S and M in this way. The inescapable conclusion is that, once we regard measurement interactions to be described by operators U acting as in (77), the whole universe becomes entangled. Of course, this is nothing but Schrödinger's cat in disguise. Once we start with a superposition somewhere in the universe, and if all interactions are given by ‘nice’ linear operators, it will not be long before the whole visible universe will find itself in a superposition that includes S and M . So this does not give us a way of explaining how measurements give definite values after all.

This further motivates von Neumann’s description of measurement not as a ‘normal’ unitary operator U , but as a ‘special’ projection operator P_i that projects unto some particular state: if we start with

$$|\psi\rangle = \sum_j c_j |a_j\rangle \quad (80)$$

then after measurement of eigenvalue a_i the system will project to the following state:

$$P_i |\psi\rangle = c_i |a_i\rangle . \quad (81)$$

Since $|a_j\rangle$ are linearly independent, we see that in order for this to work the projection operator P_i has to act on a basis as:

$$P_i|a_j\rangle = \delta_{ij}|a_i\rangle . \quad (82)$$

We then also get for our final state:

$$P_i|\psi_{\text{final}}\rangle = c_i|a_i\rangle|r_i\rangle . \quad (83)$$

Notice that the P_i 's are not unitary because the final state is not normalized to 1: $c_i|a_i\rangle|r_i\rangle$ has norm $|c_i|^2$, hence this is not a regular physical interaction of the type described by the Schrödinger equation because it does not preserve total probability. We have to rescale the state again in order to normalize it to 1. It can also be shown that projections do not conserve energy.

7.4 Other Approaches

7.4.1 Decoherence

The basic idea of decoherence is quite simple: the reason for the seemingly non-unitary evolution (81) is in the interactions with the environment. The evolution (81) *looks as if* it was non-unitary, but in reality (i.e. if we would be able to include all the very complicated interactions with the environment) we would see that it is unitary. The wave function *appears* to collapse, but in fact it satisfies the Schrödinger equation. Some parts of the wave-function, which should be present in (81), have been projected out simply because their coefficients c_j are very small. So we are dealing with a kind of thermodynamic irreversibility here in which the ‘thermal bath’ provided by the environment washes away some of the information.

Another way of saying this is that, after measurement, the wave function (after rescaling by c_i) does not look like (81), but actually looks like:

$$|a_i\rangle|r_i\rangle + \dots , \quad (84)$$

where the dots relate to terms with low probability. The wave function is actually entangled, but *for all practical purposes* it looks as if just one term in the superposition contributes. The reason for this is that we neglect interactions with the environment in the Hamiltonian.

In my opinion (but some of my colleagues may disagree on this) this is not strictly speaking a solution to the fundamental problem we posed of why we never see superpositions. The magic words above are ‘for all practical purposes’: decoherence is a solution that works in practice, but, in fact, the whole universe is entangled in the state (84) (since the environment includes the whole visible universe). So it does not solve the matter of principle we raised: how do we pass, from a formalism which predicts universal entanglement, to states in which particles have unique properties—determinate states? Simple decoherence just tells us that the determinate states are the most probable ones, but it does not explain how to actually obtain them. Unfortunately, long discussions on texts on decoherence fail to address this fundamental point.

Perhaps the answer to this is that we should not interpret the wave function to describe the world *as it actually is*, but only as *a collection of possibilities*. The wave function tells us what is possible and probable, but not what is actual.

7.4.2 Many Worlds

I just suggested that the wave function maybe only describes possibilities, not actualities. At the other extreme of the spectrum there is the many worlds interpretation, which assigns objective actuality to *all* of the terms in the wave function (whether they decohere or not), i.e. to all the summands in (84). It replaces the collapse of the wave function by its branching off: at each single measurement, the world branches off into distinct possibilities, so that every possible outcome is realized in a different world. This reconciles the appearance of non-deterministic events (such as random decay of an atom) with deterministic equations such as the Schrödinger equation.

This may sound crazy, but you would be surprised by the growing number of physicists who actually support this interpretation. There is some theoretical support for the many-worlds interpretation coming from the ‘histories approach’ to quantum mechanics as well as from some recent puzzles in quantum cosmology. Perhaps the fact that physicists are willing to support such a crazy idea is simply an indication of how bad the measurement problem sits with them.

I see a simple puzzle for the many worlds interpretation which never seems to be discussed in the literature. Consider a harmonic oscillator. It is intuitively clear how the many worlds interpretation would work for this system, since its spectrum is discrete. If the wave function is in a superposition $\sum_n c_n \psi_n$, there is a world with wave function ψ_1 , a world with wave function ψ_2 , etc. (up to infinity). But this only works because the spectrum is discrete. If the spectrum is continuous and labeled by, say, the wave number k , then there is a continuum of worlds between any two finite values of k . This seems nonsensical because it is well-known from mathematics that we cannot label the real numbers using natural numbers. The latter is an infinite but countable set. The former is not only an infinite, but an uncountably infinite set! If we cannot even count the number of worlds, how can they all have separate existence? The concept of ‘continuum’ seems counter to the idea of all these worlds having ‘separate existence’. So to me, this seems like nonsense. A way to resolve this puzzle could be to turn all continuous spectra in quantum mechanics into discrete ones, by introducing a new scale so small that we cannot see it, so that particles ‘seem to have continuous momentum k ’ but are actually always quantized and their energy levels are so dense that we cannot tell them apart. But then again: 1) This is never discussed in the many worlds literature; and 2) It amounts to *changing* quantum mechanics into a different fundamental theory, which runs counter to our initial aim of simply *interpreting* quantum mechanics! (which we assumed to be a complete theory).

A Mathematical Formulas and Tricks

A.1 Gaussian integration

- The basic Gaussian integral is:

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}} . \quad (85)$$

To show this, we compute the integral in two different ways. Consider the area under the function e^{-r^2} on the plane in plane polar coordinates:

$$\int dA e^{-r^2} = \int_0^{\infty} dr \int_0^{2\pi} d\theta r e^{-r^2} = \int_0^{\infty} \frac{du}{2} 2\pi e^{-u} = \pi \int_0^{\infty} du e^{-u} = -\pi e^{-u} \Big|_0^{\infty} = \pi , \quad (86)$$

where we defined $u = r^2$. Now compute this same integral in Cartesian coordinates $x = r \sin \theta, y = r \cos \theta$:

$$\pi = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy e^{-(x^2+y^2)} = \int_{-\infty}^{\infty} dx e^{-x^2} \int_{-\infty}^{\infty} dy e^{-y^2} = \left(\int_{-\infty}^{\infty} dx e^{-x^2} \right)^2 . \quad (87)$$

Taking the square root, it follows that:

$$\int_{-\infty}^{\infty} dx e^{-x^2} = \sqrt{\pi} . \quad (88)$$

To obtain the Gaussian integral with generic width, we simply rescale $x \rightarrow \sqrt{\alpha} x$ in the integrand, which gives back (85).

- From (85) we can also compute the following integral:

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2 - i\beta x} = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}} . \quad (89)$$

For the proof of this, see section 2.4.

- Using (85), we can also compute integrals of the following type:

$$\int_{-\infty}^{\infty} dx x^n e^{-\alpha x^2} . \quad (90)$$

First of all, we notice that this integral vanishes if n is odd. The reason is that in that case the integrand is an odd function under $x \rightarrow -x$. The integral of an odd function over a range $(-a, a)$ is zero. Since we integrate from $(-\infty, \infty)$, the result of this integral is zero.

When n is even, we use the partial integration formula:

$$\int_{\mathcal{D}} u dv = uv \Big|_{\mathcal{D}} - \int_{\mathcal{D}} du v , \quad (91)$$

where \mathcal{D} denotes the domain of integration. Noting that $d(e^{-\alpha x^2}) = -2\alpha x e^{-\alpha x^2} dx$, we can take $u = x$ and $v = e^{-\alpha x^2}/(-2\alpha)$. Applying (91), we get

$$\int_{-\infty}^{\infty} dx x^2 e^{-\alpha x^2} = \frac{x e^{-\alpha x^2}}{-2\alpha} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx \frac{e^{-\alpha x^2}}{-2\alpha} = \frac{1}{2\alpha} \sqrt{\frac{\pi}{\alpha}} = \frac{\sqrt{\pi}}{2\alpha^{3/2}}, \quad (92)$$

where the boundary term vanished because the Gaussian function vanishes at infinity more rapidly than any polynomial. We can get the result for higher powers of x by successive partial integrations. The result is:

$$\int_{-\infty}^{\infty} dx x^{2n} e^{-\alpha x^2} = \frac{(2n)!}{n!} \frac{2\sqrt{\pi}}{(4\alpha)^{n+\frac{1}{2}}}. \quad (93)$$

B Technicalities of Quantum Mechanical Measurements

B.1 Time Evolution Operators

Remember that the problem of quantum mechanics amounts to finding a solution $\Psi(x, t)$ of the TDSE (20) given an initial condition $\Psi_0(x)$. We can recast this problem in the language of operators, as follows: find a linear operator U such that:

$$U : \Psi_0(x) \mapsto \Psi(x, t) = U(t) \Psi_0(x) \quad (94)$$

satisfies the TDSE. Here, U is a time-dependent operator called the *evolution operator*. In a particular basis, and in the discrete case, it maps:

$$\sum_n c_n \psi_n(x) \mapsto \sum_n c_n e^{-\frac{i}{\hbar} E_n t} \psi_n(x). \quad (95)$$

The operator U is easy to find:

$$U(t) \equiv e^{-\frac{i}{\hbar} H t} \equiv \sum_{n=0}^{\infty} \frac{(-\frac{i}{\hbar} H t)^n}{n!}, \quad (96)$$

where the exponential, which includes an operator (the n th power of the Hamiltonian) has been defined by its Taylor expansion. Although this requires some care, Taylor expansions can indeed be defined for linear operators (as for matrices) analogously to how they are defined for numbers:

$$e^{\hat{Q}} \equiv \sum_{n=0}^{\infty} \frac{(\hat{Q})^n}{n!}. \quad (97)$$

Now the claim is that $U(t)$, as defined in (96), does the job in (94): when $U(t)$ is applied to $\Psi_0(x)$, it gives the solution of the TDSE with that initial condition. Let us check this. If $\Psi(x, t)$ is to be given by (94) with U as in (96), it must solve the TDSE:

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = i\hbar \frac{\partial U(t)}{\partial t} \Psi_0(x). \quad (98)$$

Let us compute the time derivative:

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial t} \left(e^{-\frac{i}{\hbar} H t} \right) = \left(-\frac{i}{\hbar} H \right) e^{-\frac{i}{\hbar} H t} = -\frac{i}{\hbar} H U(t) . \quad (99)$$

Filling this back into (98):

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = i\hbar \left(-\frac{i}{\hbar} H \right) U(t) \Psi_0(x) = H \Psi(x, t) , \quad (100)$$

which is precisely the TDSE. Hence, we have shown that (96) does the job of solving for us the TDSE.

Let us now see what these formulas mean in practice. If we are given the initial condition $\Psi_0(x) = \sum_n c_n \psi_n(x)$, then:

$$\begin{aligned} U(t) \Psi_0(x, t) &= U(t) \left(\sum_n c_n \psi_n(x) \right) = \sum_n c_n (U(t) \psi_n(x)) = \sum_n c_n \left(e^{-\frac{i}{\hbar} H t} \psi_n(x) \right) \\ &= \sum_n c_n e^{-\frac{i}{\hbar} E_n t} \psi_n(x) = \Psi(x, t) , \end{aligned} \quad (101)$$

where we have used the fact that when H satisfies the TISE, $H\psi_n = E_n\psi_n$, and hence we can replace any powers of H by powers of E_n when they act on ψ_n (also in the exponential, by using the Taylor expansion we can see this term by term). Now the last expression in (101) is indeed the usual solution of the TDSE.

U happens to be a unitary operator, i.e. $UU^\dagger = U^\dagger U = \mathbb{1}$. This can be shown by computing the adjoint:

$$U^\dagger = \left(e^{-\frac{i}{\hbar} H t} \right)^\dagger = e^{\left(-\frac{i}{\hbar} H t \right)^\dagger} = e^{\frac{i}{\hbar} H t} . \quad (102)$$

where we have used the fact that H is Hermitian and that, again, to compute the adjoint of the exponential, we can use the Taylor expansion, compute the adjoint term by term, and resum the Taylor series. Since we get a plus sign in the exponential, when we multiply it with U which contains a minus sign, the two cancel out and we get the unit matrix. Because U is unitary, it preserves amplitudes (and probabilities):

$$\langle \Psi | \Psi \rangle = \langle U\Psi | U\Psi \rangle = \langle \Psi_0 | U^\dagger U \Psi_0 \rangle = \langle \Psi_0 | \Psi_0 \rangle . \quad (103)$$

B.2 The Measurement Operator U

The measurement operator we were looking for in (77) can be explicitly written as follows:

$$U = \sum_{jk} |a_k\rangle |r_{j+k}\rangle \langle r_j| \langle a_k| . \quad (104)$$

It is not hard to show that U indeed satisfies (77). In fact, this operator is a combination of projectors onto A and onto R :

$$U = \sum_{jk} P_{k,k}^{(a)} P_{j,j+k}^{(r)} \quad (105)$$

where

$$\begin{aligned} P_{i,j}^{(r)} |r_l\rangle &= \delta_{il} |r_j\rangle \\ P_{i,j}^{(a)} |a_l\rangle &= \delta_{il} |a_j\rangle , \end{aligned} \quad (106)$$

and $P^{(a)}$ is defined similarly to act on $|a_l\rangle$.

Proof: We simply let U act on (77):

$$U(|\psi\rangle|r_0\rangle) = \sum_j c_j U(|a_j\rangle|r_0\rangle) . \quad (107)$$

We compute the r.h.s. separately using the decomposition into projectors (we do some relabeling of indices):

$$\begin{aligned} \sum_j c_j U(|a_j\rangle|r_0\rangle) &= \sum_l c_l \sum_{jk} P_{k,k}^{(a)} P_{j,j+k}^{(r)} |a_l\rangle|r_0\rangle \\ &= \sum_{jkl} c_l P_{k,k}^{(a)} |a_l\rangle P_{j,j+k}^{(r)} |r_0\rangle = \sum_k c_k |a_k\rangle|r_k\rangle , \end{aligned} \quad (108)$$

where in the second equality sign we used the fact that $P^{(r)}$ acts only on eigenstates of R and $P^{(a)}$ only on eigenstates of A . In the last formula we used the property (106). This is precisely what we wanted to show.

References

- [1] D.J. Griffiths, “Introduction to Quantum Mechanics”, Pearson, 2nd Edition.
- [2] D.C. Giancoli, “Physics: Principles with Applications”, Pearson, 6th Edition.
- [3] B.H. Bransden and C.J. Joachain, “Quantum Mechanics”, Addison-Wesley, 2nd Edition.
- [4] D. Dieks, “Filosofie/grondslagen van de natuurkunde”, Utrecht University, 2008-2009.