

Quantum states



1.1 Introduction

Quantum mechanics describes the behaviour of matter and light at the atomic scale (*i.e.* at distances $d \sim 10^{-10}m$), where physical objects behave very differently from what we experience in everyday's life. Because the atomic behaviour is so unusual, we need to develop new, abstract tools that allow us to *compute* the expected values of physical observables starting from the postulates of the theory. The theory can then be *tested* by comparing theoretical predictions to experimental results. In this course we will focus on presenting the basic postulates, and in developing the computational tools needed to study elementary systems. As we venture on this journey, we can only develop some intuition by practice; solving problems is an essential component for understanding quantum mechanics.

Experimental results at the beginning of the 20th century first highlighted behaviours at atomic scales that were inconsistent with classical mechanics. It took a lot of effort until the combined efforts of many physicists – Planck, Bohr, Schrödinger, Heisenberg, Born (who has been a professor in Edinburgh) amongst others– led to a consistent picture of the new dynamics.

Despite its unintuitive aspects, quantum mechanics describes very concrete features of the world as we know it, like e.g. the stability of the hydrogen atom. Many of its predictions have now been tested to great accuracy.

In fact, there are numerous experimental results that provide evidence in favour of quantum mechanics, e.g.

- double-slit experiments,
- photoelectric effect,
- stability of the H atom,
- black-body radiation.

Some of these experiments have been discussed already in previous courses, and we do not want to review them here. You can find exhaustive discussions in most textbooks of Quantum Mechanics – see the bibliography for the course for a list of suggested readings.

We shall introduce the basic ideas of quantum mechanics by discussing briefly the **double-slit** experiment, which was first performed by Young in 1801, in order to resolve the question of the corpuscular nature of light ¹. As we shall see some of the key concepts emerged

¹ An interesting account of Young's original experiment has been published in *The Physics Teacher*, **24** 217, 1986, and can be found at <http://cavendishscience.org/phys/tyoung/tyoung.htm>.

from the comparison of the quantum behaviour of electrons with the more familiar behaviour of classical particles and waves.

1.1. INTRODUCTION

1.1.1 Experiment with classical particles

Let us consider first a source of *classical* particles, emitting projectiles (bullets) in random directions. In front of the source we have a wall with two holes that are denoted **1** and **2**. A detector is placed behind the wall, counting the bullets that pass through the holes. The apparatus is schematically represented in Fig. 1.1. We do not want to get into the details of the experiment here, neither into a precise computation of the probability distributions. We should focus instead on the logical steps.

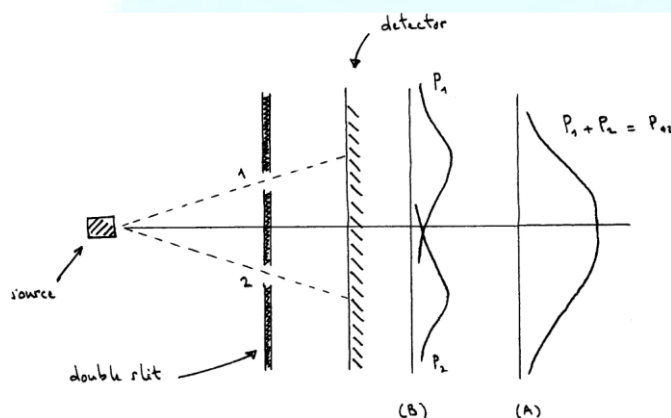


Figure 1.1: Double-slit experiment with classical particles.

When both holes are open the distribution of bullets on the detector is given by a function that resembles the curve P_{12} in (A). When **2** is closed, and therefore the bullets can only pass through **1**, we obtain the distribution labelled P_1 in (B). Similarly, when the bullets can only go through **2**, the distribution is given by P_2 . The important feature, which is typical of *classical* mechanics, is that $P_{12} = P_1 + P_2$. We call this result an observation of *no interference*, particles that go through **1** do not interfere with those that pass through **2**, and the probabilities add in an intuitive way. Each electron that arrives into the detector must have gone through *either* hole **1** *or* **2**. Therefore the chance of arrival at some position x on the detector must be the sum of the probability of getting to x via **1** *plus* the probability of getting to x via **2**.

1.1.2 Experiment with waves

Let us consider now the same experiment, but with a source emitting waves. The detector measures the *intensity* of the waves, i.e. the mean squared height of the wave.

The original wave emitted by the source is diffracted at the holes, and two new circular waves spread from each hole. If we cover each hole at a time, we obtain respectively the intensity profiles I_1 and I_2 sketched in (B) in Fig. 1.2. However when both holes are open, we find the pattern I_{12} in (A).

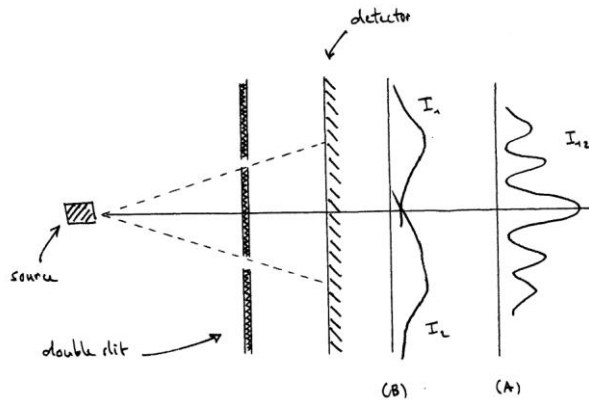


Figure 1.2: Double-slit experiment with waves.

Clearly in the case of diffracted waves $I_{12} \neq I_1 + I_2$. The two waves interfere. The maxima of I_{12} correspond to constructive interference, and viceversa for the minima. The height of the waves at time t can be represented as (the real part of) a complex amplitude $h_1 e^{i\omega t}$, where h_1 is a complex number, and the intensity is given by $I_1 = |h_1|^2$. In the case of wave interference, the complex amplitudes add up, so that $I_{12} = |h_1 + h_2|^2$.

1.1.3 Experiment with electrons

A conclusive double-slit experiment with electron beams was performed by Tonomura et al. at Hitachi (Japan)². This experiment confirmed the observation of an interference pattern similar to the one observed for waves. This experiment shows that the dynamics of the electrons, that can be considered as single indivisible particles, is influenced by the presence

² A. Tonomura et al., *Am. J. Phys.* **57**, 117, 1989.

of both slits; a single particle at the atomic scale “behaves like a wave”. This property is known as *wave-particle duality*. We can find the correct probability distribution by identifying P_{12} with the absolute square of a complex quantity ψ_{12} , the so-called *probability amplitude*. The probability amplitudes satisfy the following relations:

$$P_{12}(x) = |\psi_{12}(x)|^2, \quad (1.1)$$

$$P_1(x) = |\psi_1(x)|^2, \quad P_2(x) = |\psi_2(x)|^2, \quad (1.2)$$

$$\psi_{12}(x) = \psi_1(x) + \psi_2(x). \quad (1.3)$$

Wave-particle duality was first formulated by De Broglie in 1925, who stated that a particle with momentum p behaves like a wave with wavelength $\lambda = h/p$, where h is Planck’s constant:

$$h = 6.62606896(33) \times 10^{-34} \text{ J} \cdot \text{s}. \quad (1.4)$$

Planck’s constant is a fundamental constant of Nature; it has units $[h] = \text{Energy} \times \text{Time}$; as we shall see in more details later it defines the scale where quantum phenomena become relevant. We will often encounter the constant $\hbar = h/(2\pi)$.

In order to avoid contradictions that may arise from this dual use of wave and particle languages, we must be very careful to define properly the statements that are permitted about a given experimental situation. In other words, we need to specify carefully the laws of the new quantum theory. As we will see, there are some crucial differences between the quantum and the classical world.³

1.2 One-dimensional systems

We shall now proceed to summarize the laws of quantum mechanics for one-dimensional systems. Before entering into the details, it is worthwhile to quote Feynman’s lectures⁴:

“In this subject we have, of course, the difficulty that the quantum mechanical behavior of things is quite strange. Nobody has an everyday experience to lean on to get a rough, intuitive idea of what will happen. So there are two ways of presenting the subject: We could either describe what can happen in a rather rough physical way, telling you more or less what happens without giving the precise laws of everything; or we could, on the other hand, give the precise laws in their

³ A more exhaustive discussion of the interpretation of the double-slit experiment can be found in R.P. Feynman, “The concept of probability in Quantum Mechanics”, Proc. Second Berkeley Symp. on Math. Stat. and Prob., 533-541, Univ. of California Press, 1951. The article is available online at <http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.bsmsp/1200500252>. The same discussion is also presented in the first chapter of R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York 1965, Dover Publications 2010.

⁴ R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman lectures on physics - Quantum Mechanics*, Addison-Wesley, 1965.



abstract form. But, then because of the abstractions, you wouldn't know what they were all about, physically. The latter method is unsatisfactory because it is completely abstract, and the first way leaves an uncomfortable feeling because one doesn't know exactly what is true and what is false. [...] Here, we will try to find a happy medium between the two extremes".

1.2.1 Quantum states

Let us begin with

the fundamental law of quantum mechanics

which summarizes the idea of wave-particle duality. The quantum state of a system is described by a **complex**⁵ function Ψ , which depends on the coordinate x and on time:

$$\text{quantum state} \sim \Psi(x,t) \tag{1.5}$$

The wave function does not depend on the momentum of the particle. Compared to classical mechanics, we seem to have lost the symmetry between coordinates and momenta. We shall revisit this issue later. The wave function encodes all the information about the system, albeit in a probabilistic sense. This is a peculiarity of Quantum Mechanics: as postulated by Born, the theory can only predict *the probability of the outcome* of an experiment. This probability can be computed from the wave function. There are cases where a complicated computation is needed, and there are cases where this probability can be obtained very easily.

For instance,

$$|\Psi(x,t)|^2 dx \text{ is the probability that a measurement of the position of the particle yields a result in the interval } x \rightarrow x + dx.$$

Thus $|\Psi(x,t)|^2$ is a *probability per unit length* or *probability density*. The total probability of finding the particle somewhere along the real axis must be unity, thus:

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 dx = 1. \tag{1.6}$$

⁵ QM relies entirely on being able to manipulate correctly complex numbers. If you do not feel confident with elementary complex analysis, please go back to your Pre-Honours notes as quickly as possible.

Any function such that its integral along the real axis is finite can be normalized by multiplying by an appropriate constant. In practice two wave functions that differ by an arbitrary factor $c \in \mathbb{C}$ describe the same physical system.

Mathematical aside

Let us discuss an example of a normalizable function. The function $\psi(x) = e^{-x^2/2}$ is clearly normalizable. Its norm is

$$\|\psi\|^2 = \int_{-\infty}^{\infty} dx e^{-x^2} = \sqrt{\pi}, \quad (1.7)$$

and therefore the *normalized wave function* is:

$$\psi(x) = \frac{1}{\pi^{1/4}} \exp[-x^2/2]. \quad (1.8)$$

On the other hand, the function $e^{x^2/2}$ is non-normalizable, and therefore does not represent a physical state.

In general, if $\int_{-\infty}^{\infty} dx |\psi(x)|^2 = c$, then the normalized wave function is $\frac{1}{\sqrt{c}} \psi(x)$.

Example In order to understand better how the information about the system is encoded in the wave function, we shall start with a simpler version of our one-dimensional system. Let us consider a particle in a discretized space. The particle can only be in a *finite* number of positions along the real axis, as shown in Fig. 1.3. In this particular example, the particle can be in one of *six*⁶ points along the real axis labelled $0, \dots, 5$. The lattice spacing (i.e. the distance between two points) is denoted a .

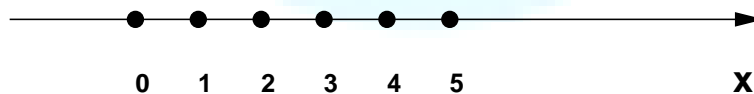


Figure 1.3: Discretized one-dimensional system with six sites. The particle can only be in one of the six points denoted by $0, \dots, 5$. The distance between successive points is a .

⁶ There is no particular significance in the fact that we have chosen here 6 points. The same example could have been worked out with two points, or any finite number of points.

According to the probabilistic interpretation of the wave function, the probability for the particle to be at x_i at time t is given by $|\Psi(x_i, t)|^2 \epsilon$. Note the factor that multiplies the modulo square of the wave function. This is needed because the modulo square is the probability density per unit length. Hence to find the probability we need to multiply by the distance between two points. If we redefine

$$\psi_i = \sqrt{\epsilon} \Psi(x_i, t), i = 0, \dots, 5, \tag{1.9}$$

then the whole information about this system is encoded in a six-dimensional *complex vector*:

$$|\Psi\rangle = (\psi_0, \dots, \psi_5). \tag{1.10}$$

Note that the x dependence is not explicitly written in the ket $|\Psi\rangle$. The values of Ψ at different spatial points are the components of the state vector. On the other hand, we will use $|\Psi(t)\rangle$ to indicate the dependence of the state on time when necessary.

In Eq. (1.10) we have used the *ket* notation introduced by Dirac to denote a state vector. We shall keep using this notation throughout these notes.

Example What is the state vector, $|2\rangle$, of a particle localized at the site x_2 ?

A particle localized at x_2 is a particle that has a probability one of being at x_2 , and probability zero to be at any other site. Hence:

$$|2\rangle = (0, 0, 1, 0, 0, 0). \tag{1.11}$$

Actually, from the discussions above, it should be clear that any state of the form:

$$|2\rangle = (0, 0, z, 0, 0, 0) \text{ where } |z|^2 = 1, \tag{1.12}$$

represents a localized solution.

In the (simple) discretized system that we have discussed, we can see explicitly that the state of the system is represented by a vector. This is an important concept to remember.

The normalization condition encodes the fact that the total probability of finding the particle somewhere on the line *must be* one, *i.e.*

$$\sum_{k=0} |\psi_k|^2 = 1, \quad (1.13)$$

i.e. $|\Psi\rangle$ is a *complex* vector of unit norm. We shall frequently use the following convention:

$$\langle \Psi | \Psi \rangle = \sum_{k=0}^5 |\psi_k|^2, \quad (1.14)$$

where we have used Dirac's notation to indicate the scalar product of two vectors:

$$\langle \Phi | \Psi \rangle = \sum_k \varphi_k^* \psi_k. \quad (1.15)$$

The wave function for a continuous system can be seen as the limit of the discretized case where the number of points goes to infinity, while the distance becomes infinitesimally small.

In this limit, instead of a finite-dimensional vector, we obtain an infinite number of coordinates, encoded in a continuous function $\psi(x)$.

We shall still refer to the wave function as the state vector, bearing in mind that in this case the vector space is infinite-dimensional. There is one coordinate $\Psi(x,t)$ for each point x on the real axis, and there is an infinity of points along the real axis. We denote the state vector using a ket $|\Psi(t)\rangle$. Note that the time dependence of the state vector is explicit in Dirac's notation. It should be clear from the discussion above that the spatial coordinate x labels the components of the state vector.

The norm of the state vector can be written as the limit of the norm of the finite-dimensional vector for $\epsilon \rightarrow 0$. Starting from Eq. (1.14), and taking the limit:

$$\langle \Psi(t) | \Psi(t) \rangle = \lim_{\epsilon \rightarrow 0} \sum_k \epsilon |\Psi(x_k, t)|^2 = \int dx |\Psi(x, t)|^2, \quad (1.16)$$

we recover Eq. (1.6). Similarly the scalar product of two wave functions can be defined as the limit of the discrete case:

$$\langle \Phi(t) | \Psi(t) \rangle = \int dx \Phi(x, t)^* \Psi(x, t). \quad (1.17)$$

1.2.2 Superposition principle

State vectors can be combined linearly to obtain new admissible quantum states. If Ψ_1 and Ψ_2 are quantum states, then

$$\Psi(x, t) = c_1 \Psi_1(x, t) + c_2 \Psi_2(x, t), \quad c_1, c_2 \in \mathbb{C}, \quad (1.18)$$

is also a possible state of the system, as long as $\langle \Psi | \Psi \rangle = 1$. In mathematical terms, the space of possible quantum states is called a *vector space*⁷.

Mathematical aside

As a consequence of the superposition principle the time evolution of a quantum mechanical system must be determined by a *linear* equation:

$$L\Psi = 0, \quad (1.19)$$

where L is a linear operator, i.e. an operator such that:

$$L(c_1\Psi_1 + c_2\Psi_2) = c_1L\Psi_1 + c_2L\Psi_2. \quad (1.20)$$

⁷If necessary, you can find a brief summary of the properties of vector spaces at [http://en.wikipedia.org/wiki/Vector space](http://en.wikipedia.org/wiki/Vector_space); a better option is to go back to your lecture notes from Pre-Honours.

Example 1 Note that the concept of superposition of states is very different from anything we have encountered in classical mechanics. Consider two quantum states $|A\rangle$ and $|B\rangle$, such that the measurement of an observable O yields the result a when the system is in the state $|A\rangle$, and the result b when the system is in the state $|B\rangle$. The superposition principle states that the state vector:

$$|C\rangle = c_A|A\rangle + c_B|B\rangle, \quad (1.21)$$

where c_A and c_B are complex numbers such that $|c_A|^2 + |c_B|^2 = 1$, describes another possible physical state of the system. It turns out that the measurement of the observable O in the state $|C\rangle$ can only yield the value a or b , with respective probabilities:

$$p_a = \frac{|c_A|^2}{|c_A|^2 + |c_B|^2}, \quad p_b = \frac{|c_B|^2}{|c_A|^2 + |c_B|^2}. \quad (1.22)$$

No other results are possible for the measured value of O in the state $|C\rangle$.

Example 2 Let us consider again the discretized one-dimensional system in Fig. 1.3; we can have a state $|1\rangle$ where the particle is localized e.g. at site 1, and a state $|2\rangle$ where the particle is localized at site 2. Measuring the position of the particle in state $|1\rangle$ yields $x = 1$ with probability 1. Likewise we obtain $x = 2$ with probability 1 for a particle described by the state vector $|2\rangle$.

The state $\frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle$ is an admissible quantum state. Measuring the position of the particle in this latter state, the outcome will be $x = 1$, or $x = 2$ with 50% probability. No other value is allowed in this state.

1.2.3 The uncertainty principle

In classical mechanics the state of a particle in a one-dimensional world is completely determined by the value of its position $x(t)$ and momentum $p(t)$, i.e. by its *trajectory*.

The situation is radically different in quantum mechanics. The probabilistic interpretation of the wave function implies that we can at best obtain the probability density for a particle to be at a given position x at time t . As a consequence the concept of *classical trajectory* used in Newtonian mechanics does not make sense in quantum mechanics. The position and momentum of the particle can be defined, but their values cannot be measured simultaneously.

On the other hand, when the scales in the problem are much larger than the Planck constant h , we expect to recover the classical results.

These two features are summarized in the so-called *uncertainty relations*, first derived by Heisenberg. The uncertainty relations state that, if the position and momentum are measured simultaneously, with respective precisions Δx and Δp , then:

$$\Delta x \cdot \Delta p \geq \frac{\hbar}{2}. \quad (1.23)$$

It is clear from Eq. (1.23) that if the position of the particle is known exactly, then the knowledge of its momentum is completely lost. In general the product of the two uncertainties has to be greater than $\hbar/2$.

It is important to appreciate that Heisenberg's inequalities reflect a physical limitation. A better experimental apparatus would **not** allow a higher precision to be obtained. The uncertainty is a property of the dynamics of the system.

They also encode the idea that in quantum mechanics the measurement of a quantity interferes with the dynamics. If we measure exactly the position of the particle, then we lose all knowledge of its momentum. We shall revisit this issue later in the course.

Hence the concept of determinism is lost during the measurement process. As we will discuss below, the evolution of a quantum system left unperturbed is completely determined by the Schrödinger equation, and therefore is deterministic. You should contrast this with the situation in classical mechanics, where we can assume that measurements do not perturb the state of the system.

1.2.4 Dynamical variables

We have seen above that the modulo square of the wave function $|\Psi(x,t)|^2$ yields the probability density of finding the particle at position x at time t . We can therefore compute the *mean value* of the particle position in the state Ψ at time t :

$$\langle x \rangle_{\Psi, t} = \int_{-\infty}^{\infty} dx x |\Psi(x, t)|^2. \quad (1.24)$$

The notation $\langle x \rangle_{\Psi, t}$ emphasizes that the mean value is computed in the state described by the wave function Ψ at time t . This average is a property of the state of the system. If the system is in a different state, the mean value of the position will be different. In general we will omit the suffix, however you should keep in mind that mean values depend on the state of the system.

Mathematical aside

Let us consider a random real variable r , characterised by a probability distribution function $p(r)$. The probability of finding r in the interval $I = [a, b]$ is

$$P(r \in I) = \int_a^b dx p(x). \quad (1.25)$$

(Note that x is a dummy variable of integration, we could use any name to identify the integration variable. If this does not sound familiar, think about it...) The n -th moment of the distribution p is defined as:

$$\mu_n = \int_{-\infty}^{\infty} dr p(r) r^n. \quad (1.26)$$

For a properly normalised probability distribution function:

$$\mu_0 = 1. \quad (1.27)$$

The mean value of the random variable r is:

$$\langle r \rangle = \mu_1 = \int_{-\infty}^{\infty} dr p(r) r. \quad (1.28)$$

The variance of the variable r is:

$$\text{Var}[r] = \mu_2 - (\mu_1)^2. \quad (1.29)$$

Compute the mean value and the variance of a Gaussian variable r , i.e. a random variable with probability distribution function:

$$f(r) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - x_0)^2}{2\sigma^2}\right] \quad (1.30)$$

If we define an operator \hat{X} acting on the wave function:

$$\hat{X}\Psi(x,t) = x\Psi(x,t); \quad (1.31)$$

the expectation value above can be written as:

$$\int dx \Psi^*(x,t) \hat{X} \Psi(x,t) = \langle \hat{X} \rangle = \int dx \Psi^*(x,t) x \Psi(x,t) = \langle x \rangle. \quad (1.32)$$

Here we have introduced the mean value of an operator, \hat{X} , in a given state. As discussed above, the expectation value in Eq. (1.32) depends on the state $\Psi(t)$ on which it is computed.

When there is no ambiguity, we shall omit the explicit Ψ dependence, and write simply $\langle \hat{X} \rangle$.

More generally in quantum mechanics each observable O is associated to a *linear operator* \hat{O} acting on the wave function. A linear operator satisfies:

$$\hat{O} (c_1\Psi_1(x,t) + c_2\Psi_2(x,t)) = c_1\hat{O}\Psi_1(x,t) + c_2\hat{O}\Psi_2(x,t), \quad (1.33)$$

where c_1 and c_2 are complex numbers. Note that the result of acting on a state vector with an operator produces *new* state vector.

$$\hat{O}\Psi(x,t) = \Psi^0(x,t), \text{ or equivalently } \hat{O}|\Psi(t)\rangle = |\Psi^0(t)\rangle. \quad (1.34)$$

1.3. SUMMARY

Example For instance the Hamiltonian of the system may be written as $H^\wedge = T^\wedge + V^\wedge$, where the kinetic and potential energy operators are defined by:

$$\hat{T} = \frac{\hat{P}^2}{2m}, \quad \hat{V} = V(\hat{X}). \quad (1.35)$$

The operator $V(X^\wedge)$, which is a function of the position operator X^\wedge , acts as:

$$V(X^\wedge)\Psi(x,t) = V(x)\Psi(x,t). \quad (1.36)$$

Any other operator that is a function of X^\wedge acts in the same way. We shall see later how to define the operator \hat{p} , and more complicated operators.

1.3 Summary

Let us conclude this chapter by summarizing the main concepts that have been introduced.

- Quantum states are represented by a wave function $\Psi(x,t)$, or equivalently by a vector $|\Psi(t)\rangle$ in an infinite-dimensional vector space.
- in an interval $[x, x + dx]$. Hence the normalization $\int |\Psi(x,t)|^2 dx$ yields the probability of finding the particle: $\int dx |\Psi(x,t)|^2 = 1$.

- Superposition principle: a linear combination of wave functions with arbitrary complex coefficients yields a possible quantum state.
- Uncertainty principle.
- Dynamical variables are associated to linear operators acting on the wave functions.

