4

How Do Physicists Deal with Interference?

So, we have seen, in Chap. 2, one of the two "impossible things that we have to believe before breakfast" namely things being apparently in two different states before being measured or before one looks at them.

There is a mathematical language that allows to predict the "impossible" phenomenon of the double-slit experiment. We shall describe this language without mathematics.[1]

This language should be regarded, for the moment, only as a "recipe" for predicting those phenomena, without worrying about what it "means" physically. The main problem for the reader is probably *not to ask, for the moment*: what does this language mean (beyond being an efficient recipe for predictions of observations)? We shall come to the question of the physical meaning of this language in the next chapter, but in this chapter, we shall only explain how the recipe works.

4.1 The Wave Function

In order to predict the behavior of electrons in the double-slit experiment, we need a notion that is central in the whole of quantum mechanics: the *wave function*.[2] As the name indicates, it is a function, which is usually denoted by

---

[1] See the Appendix for somewhat more precise mathematical treatment; however *it is not necessary to read the Appendix in order to follow the rest of the arguments*.

[2] Note for the advanced reader: throughout this book, we shall not distinguish between the wave function of a physical system and its quantum state, which is a more general notion.

J. Bricmont, *Quantum Sense and Nonsense*, DOI 10.1007/978-3-319-65271-9_4

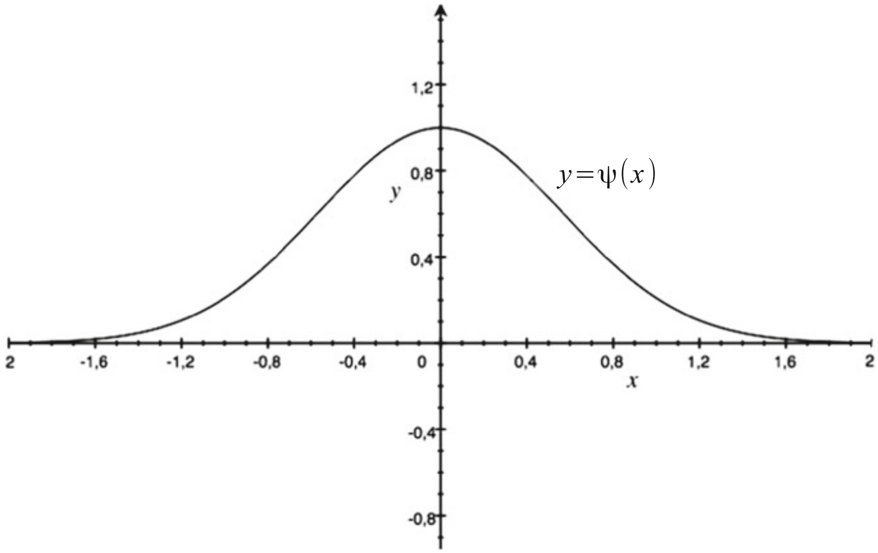

Fig. 4.1 An example of a wave function $\Psi(x)$. Each point on the *curve*, with coordinate (x, y) has a y coordinate equal to $\Psi(x)$

the Greek letter Ψ and such a function is associated in principle to any physical system.

To simplify matters, we shall consider Ψ as being a function of one variable denoted x: $\Psi = \Psi(x)$, which corresponds to a single particle moving on a line.[3] For a graphical representation of wave functions, see Figs. 4.1 and 4.2. We will give other examples of wave functions below.

The physical meaning of $\Psi(x)$, in orthodox quantum mechanics, is simply that the square of the wave function $\Psi(x)^2$, determines the probability of finding the particle somewhere if one "measures" its position: the probability of finding the particle in a set A is given by the shaded area in Fig. 4.3 under the curve $\Psi(x)^2$. We assume that the total area under the curve in Fig. 4.3 is equal to one, so that the probability of the particle being found *somewhere* is equal to one, as it should. Another example of a function $\Psi(x)^2$ is given in Fig. 4.4.[4]

---

[3] In mathematics, a function is usually denoted $f(x)$, but for the wave function the notation $\Psi(x)$ is almost universal.

[4] A caveat is necessary here: the number $\Psi(x)$ is in reality a complex number and one should write $|\Psi(x)|^2$ instead of $\Psi(x)^2$, see the Appendix.

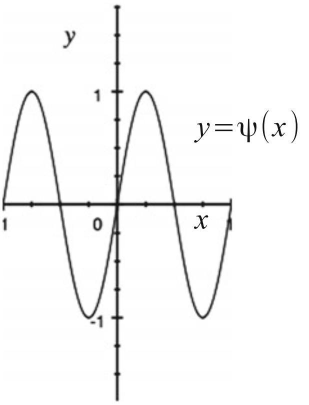

Fig. 4.2 An example of a wave function $\Psi(x)$ that takes both positive and negative values. Each point on the *curve*, with coordinate (x, y) has a y coordinate equal to $\Psi(x)$

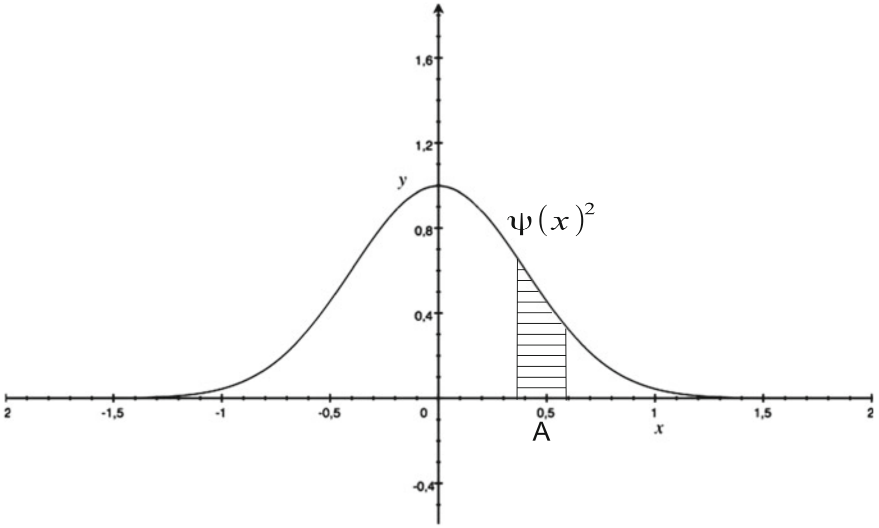

Fig. 4.3 An example of a wave function $\Psi(x)^2$. Each point on the *curve*, with coordinate (x, y) has a y coordinate equal to $\Psi(x)^2$. The function $\Psi(x)^2$ is the square of the function $\Psi(x)$ drawn in Fig. 4.1. The probability of finding the particle in the region A is equal to the *shaded area* (we used the same curve as in Fig. 3.2)

So, if one considers many particles, all having the same wave function $\Psi(x)$, and one detects their individual positions, the statistical distribution of those positions will be given by the curve $\Psi(x)^2$.[5]

---

[5] This fact follows from the law of large numbers, discussed in Sect. 3.4.1.

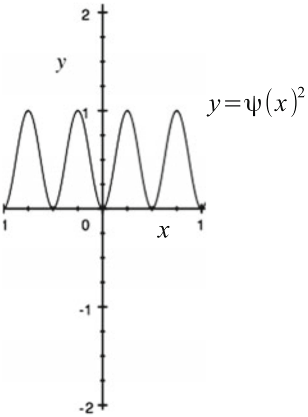

Fig. 4.4 Another example of a wave function $\Psi(x)^2$. Each point on the *curve*, with coordinate (x, y) has a y coordinate equal to $\Psi(x)^2$. The function $\Psi(x)^2$ is the square of the function $\Psi(x)$ drawn in Fig. 4.2

An immediate question may occur: if the physical meaning of $\Psi(x)$ is given by $\Psi(x)^2$, why talk about $\Psi(x)$ instead of talking directly and only about $\Psi(x)^2$? That will become clear soon, when we explain below that, in some circumstances, one must add different wave functions and not their squares.

Another natural question is, why use the square of $\Psi(x)$, and not $\Psi(x)$ itself or its cube $\Psi(x)^3$, or its fourth power $\Psi(x)^4$, or any other power? Why not $\Psi(x)$ is easy to answer: $\Psi(x)$ in general is not a positive number and probabilities have to be positive numbers (see Fig. 4.2)! The same answer holds for $\Psi(x)^3$ (the cube of a negative number is also negative), but for the other powers, there is no easy answer: the fact is that only $\Psi(x)^2$ leads to experimentally correct results. We will simply accept that as a fact.

There are *two mistakes* to be avoided when one thinks about the wave function in the orthodox fashion:

1. The first mistake is to think that $\Psi(x)^2$ describes some density of "stuff", for example some density of matter or of electric charge. In classical physics, particles were supposed to be "point particles", namely localized at an exact point in space, which works at least as an idealization. It is certainly more intuitive to think of them as being somewhat spread out and if $\Psi(x)^2$ corresponded to a density of matter or of electric charge, that would be appealing (actually this is how Schrödinger first thought of the meaning of $\Psi(x)^2$).

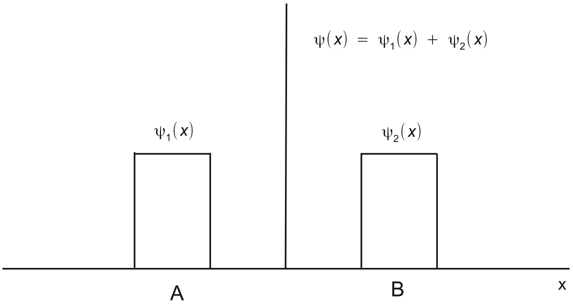

Fig. 4.5 A wave function being the sum of two wave functions that are non-zero over non-overlapping regions

The problem is that, as we see in the example drawn in Fig. 4.5, the wave function $\Psi(x)$ can be a sum of two other functions:

$$\Psi(x) = \Psi_1(x) + \Psi_2(x), \tag{4.1}$$

where the sum of the functions $\Psi_1(x)$ and $\Psi_2(x)$ is easy to understand: we add, for each value of the variable x the numbers $\Psi_1(x)$ and $\Psi_2(x)$. Here, $\Psi_1(x) = 0$ where $\Psi_2(x) \neq 0$ and vice-versa, so that the sum $\Psi(x)$ is simply equal to $\Psi_1(x)$ on the interval A where $\Psi_1(x) \neq 0$, to $\Psi_2(x)$ on the interval B where $\Psi_2(x) \neq 0$, and is equal to 0 everywhere else.

The two regions A where $\Psi_1(x) \neq 0$ and B where $\Psi_2(x) \neq 0$ can be as far apart from each other as one wants (we do not indicate the distance between A and B in Fig. 4.5, but we can imagine it to be large) and then the density of stuff picture does not work, because the particle will always be detected either in the region A where $\Psi_1(x) \neq 0$ or in region B where $\Psi_2(x) \neq 0$, but not in both.

2. The other mistake is to think that $\Psi(x)^2$ determines the probability of the particle *being* in a region like A in Fig. 4.3. But that is *not* what orthodox quantum mechanics says. The latter always defines $\Psi(x)^2$ as determining the probability of the particle *being found in a region like A in Fig. 4.3 if one measures its position.*

   Of course the two notions may look identical: after all, measuring something means that we measure some property (here the particle's position) that is there before one measures it. If one measures the length of a table, one assumes that the table has a length and that, when we measure it, we simply learn what it is.

   But in ordinary quantum mechanics it is not so simple, because, a priori, a measurement might affect the system being "measured" (which of course implies that using the word “measurement” in this context could be misleading).

That is why ordinary quantum mechanics only speaks of *results of measurements*, but says nothing about what happens outside of them or before a measurement. Results of measurement are, by definition, sufficiently macroscopic so that they can be directly perceived by us. Quantum mechanics does not say that particles have a position before being “observed”.[6]

As already mentioned in Chap. 1, someone who expressed that view very clearly was Pascual Jordan, one of the founder of quantum mechanics and an adherent of the Copenhagen interpretation of quantum mechanics, who wrote:

> In a measurement of position, “the electron is forced to a decision. We compel it *to assume a definite position*; previously, it was, in general, neither here nor there; it had not yet made its decision for a definite position [...] If, in another experiment, the *velocity* of the electron is measured, this means: the electron is compelled to decide itself for some exactly defined value of the velocity; and we observe *which* value it has chosen.”
>
> Pascual Jordan [109], quoted and translated by M. Jammer [108, p. 161] (italics in the original)

Putting aside those mistakes (which are made quite naturally), let us come back to the meaning of $\Psi(x)$. Equation (4.1) and Fig. 4.5 is the first example encountered in this book of what is called a *superposition* or a *superposed wave function* which is an expression used whenever the wave function is a sum of wave functions corresponding to different physical situations.

A superposition is usually described by saying that the particle is *both in A and B*, which means only that the particle will be detected either in A or in B if ones measures its position and that the theory refuses to claim that it is either in A or B before being measured. This, of course, gives a special role to measurements and is the source of this special role in the quantum formalism.

The latter statement is very similar to the claim that the particle goes through both slits in the double-slit experiment; we will discuss that experiment using the language of wave functions in the next section.

---

[6] In Chap. 8 we shall see that $\Psi(x)^2$ can actually be understood as being related to the probability of the particle *being* at point x, but that will be possible only within a more complete theory than ordinary quantum mechanics.

At this point, we beg the reader *not to try to understand* what the wave function means, beyond what is said here: it gives the probability distribution of results of measurements and that's it![7]

The wave function $\Psi(x)$ also changes with time, more or less like a wave, hence the name wave function. The equation that governs the way $\Psi(x)$ changes in time, *when no measurements are made* is famously known as *Schrödinger's equation*. To indicate the fact that Ψ evolves in time, we write Ψ as a function of both the position x and the time t: $\Psi(x, t)$.

We shall not write down Schrödinger's equation, but we will list here two fundamental properties of the way $\Psi(x, t)$ evolves in time:

1. This evolution is *deterministic*, i.e. if, at some initial time, noted 0, we give ourselves a function $\Psi(x, 0)$, then this determines a unique function $\Psi(x, t)$ for all later times t. We explained the notion of determinism in Chap. 3. The function $\Psi(x, 0)$ corresponds to what we called the initial condition in that chapter. Once it is given, the functions $\Psi(x, t)$ for later times are determined in a unique way.
2. This evolution is *linear*: if the initial wave function at time 0 is a sum of two other wave functions $\Psi_1(x, 0)$ and $\Psi_2(x, 0)$, as in (4.1):

   $$\Psi(x, 0) = \Psi_1(x, 0) + \Psi_2(x, 0),$$

   then for all later times,

   $$\Psi(x, t) = \Psi_1(x, t) + \Psi_2(x, t), \tag{4.2}$$

   where $\Psi_1(x, t)$ is the result of the deterministic evolution with initial wave function $\Psi_1(x, 0)$, and $\Psi_2(x, t)$ is the result of the deterministic evolution with initial wave function $\Psi_2(x, 0)$.

This describes all we need to know about the time evolution *when no measurements are made*.

But what happens to $\Psi(x, t)$ if we measure the position of the particle? To give a simple example, suppose that we have, as in Figs. 4.5 and 4.6, a wave function $\Psi(x, t) = \Psi(x) = \Psi_1(x) + \Psi_2(x)$, with the region A, where $\Psi_1(x) \neq 0$, and the region B, where $\Psi_2(x) \neq 0$, being different and with t being the time when the measurement is made.

---

[7] We will discuss in the next chapter some apparently natural ways to understand what the wave function means (and see that they run into problems). We will later give a physical meaning to the wave function, in Chap. 8.

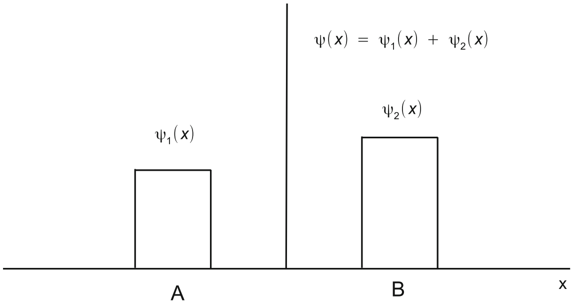

Fig. 4.6 A wave function being the sum of two wave functions that are non-zero over non-overlapping regions, but with unequal probabilities for the particle to be found in region A and in region B

Then, for $\Psi(x)$ as in Fig. 4.5, we shall find the particle in the region A with probability $\frac{1}{2}$, and in region B with probability $\frac{1}{2}$ (because of the symmetry in Fig. 4.5). After the measurement, the wave function *"collapses" or "is reduced"* to either $\Psi_1(x)$ or $\Psi_2(x)$, depending on the result. For an example where there is no symmetry and the probabilities to find the particle in region A and in region B are not equal, see Fig. 4.6.

The time evolution "during measurements" has two properties that are the exact opposite of what happens "outside of measurements":

1. This evolution is *non-deterministic*, i.e. it gives only the probabilities for the particle to be found in regions A or B and, depending on the result, the wave function jumps to either $\Psi_1(x)$ or $\Psi_2(x)$. In Fig. 4.5, we chose a symmetric situation, where both probabilities are equal to $\frac{1}{2}$, but in principle any other probabilities could be obtained, see Fig. 4.6.
2. This evolution is *non-linear*: in our example, we start with a wave function $\Psi(x, 0) = \Psi_1(x) + \Psi_2(x)$, see Fig. 4.5. After the measurement, we get $\Psi(x, t) = \Psi_1(x)$ or $\Psi(x, t) = \Psi_2(x)$, where t is a time right after the measurement. This is a *not* a sum as in the linear evolution (4.2).

Now, the fact that we have incompatible rules for the time evolution "outside of measurements" and "during measurements" raises an obvious question: what kind of physical processes qualify as measurements? And why do the physical laws change when measurements occur? As we have said, this question will be with us throughout this book, but we see here why it enters into ordinary quantum mechanics. And, since the word "measurement" implies an "observer" who does the measurement, we see why ordinary quantum mechanics puts the observer on center stage.

The problem posed by this duality of rules was expressed ironically by John Bell:

> What exactly qualifies some physical systems to play the role of "measurer"? Was the wavefunction of the world waiting to jump for thousands of millions of years until a single-celled living creature appeared? Or did it have to wait a little longer, for some better qualified system …with a PhD?
>
> John S. Bell [12, p. 34]

Finally, we should stress that it is *only through this second rule, the one valid "during measurements" that probabilities enter into quantum mechanics*. The Schrödinger evolution, valid when no measurements are made, is perfectly deterministic.

Thus, both the central role of the observer and the apparent indeterminism of quantum mechanics have their roots in this collapse rule. We will see in Chap. 7 that the problem of nonlocality is also related to this rule.

4.2 The Double-Slit Experiment

Let us now discuss the double-slit experiment, using the language of wave functions.

To do that, we shall rely on Fig. 4.10. Let 0 denote the time when the wave passes through one or two slits in the first wall and let $\Psi_1(x, 0)$ be the wave function right beyond the upper slit. Let $\Psi_2(x, 0)$ be the wave function right beyond the lower slit.

To illustrate this via a simple example, consider Fig. 4.7, where we have drawn the wave functions of Fig. 4.5, but with the variable x on the vertical axis: one can think of $\Psi_1(x)$ and $\Psi_2(x)$ as qualitatively similar to the wave functions $\Psi_1(x, 0)$ and $\Psi_2(x, 0)$ after the slits in the two slits experiment of Fig. 4.10.

The wave arriving on the second wall when only the upper slit is open is given by $\Psi_1(x, t)$, where t is the time at which the particles are detected, and where $\Psi_1(x, t)$ is the solution of the usual time evolution when one starts with $\Psi_1(x, 0)$, see Fig. 4.8. The wave coming on the second screen when only the lower slit is open is given by $\Psi_2(x, t)$, where t is the time at which the particles are detected, and where $\Psi_2(x, t)$ is the solution of the usual time evolution when one starts with $\Psi_2(x, 0)$, see Fig. 4.9.

If only the upper slit is open, the density of particles detected at a point x on the screen will be given by $\Psi_1(x, t)^2$, since that density is given by the *square* of the wave function, see the curve on the right of Fig. 4.8. Similarly, if

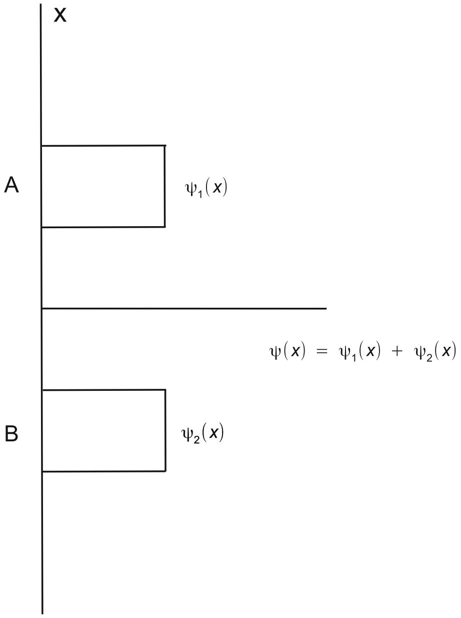

Fig. 4.7 The same wave function as in Fig. 4.5, but drawn vertically

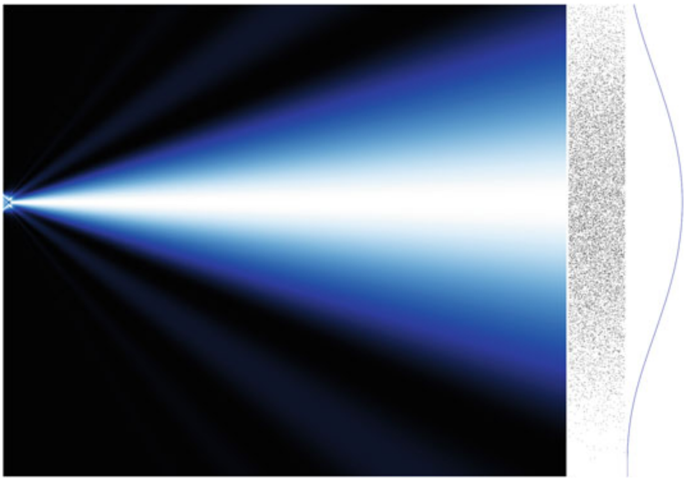

Fig. 4.8 The time evolution of the wave function in the situation of the double-slit experiment when only the upper slit is open (the time evolution goes from *left* to *right*). The *white* and *blue* areas indicate places where the wave function is non zero and their intensity is proportional to the square of the wave function (*white* more intense, *blue* less intense). *Dots* on the right indicate the impact of particles. The *blue curve* on the right indicates the density of such impacts. This figure corresponds to part (a) of Fig. 2.6, but in two dimensions and interpreted in the language of the wave function (A. Gondran cc by-sa 4.0)

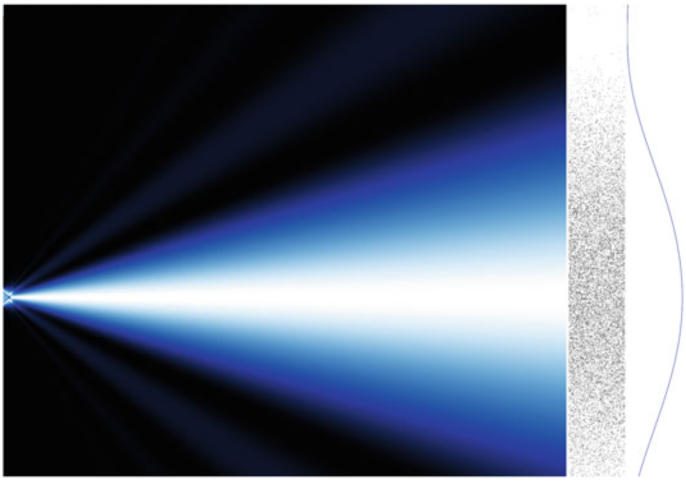

Fig. 4.9 The time evolution of the wave function in the situation of the double-slit experiment when only the lower slit is open (the time evolution goes from *left* to *right*). The *white* and *blue* areas indicate places where the wave function is non zero and their intensity is proportional to the square of the wave function (*white* more intense, *blue* less intense). *Dots* on the right indicate the impact of particles. The *blue curve* on the right indicates the density of such impacts. This figure corresponds to part (b) of Fig. 2.6, but in two dimensions and interpreted in the language of the wave function (A. Gondran cc by-sa 4.0)

only the lower slit is open, the density of particles detected at a point x on the screen will be given by $\Psi_2(x, t)^2$, see the curve on the right of Fig. 4.9.

But if both slits are open, then the initial wave function, just beyond the slits, will be the sum of the wave function associated with the upper slit and the one associated with the lower slit, $\Psi(x, 0) = \Psi_1(x, 0) + \Psi_2(x, 0)$.

And, because of the fundamental property of *linearity* of the evolution of the wave functions, the wave function arriving on the second screen when both slits are open will be the sum of the two wave functions, $\Psi(x, t) = \Psi_1(x, t) + \Psi_2(x, t)$.

So, the density of particles detected at a point x on the screen will be given by the square of $\Psi(x, t)$, namely by $(\Psi_1(x, t) + \Psi_2(x, t))^2$, see the curve on the right of Fig. 4.10. But that square *is not equal* to the sum of the squares $\Psi_1(x, t)^2 + \Psi_2(x, t)^2$.[8]

---

[8] To see this, consider the following example:

$$(3 + 4)^2 = 7^2 = 49 \neq 3^2 + 4^2 = 9 + 16 = 25. \tag{4.3}$$

In general, for real numbers a and b, we have:

$$(a + b)^2 \neq a^2 + b^2, \tag{4.4}$$

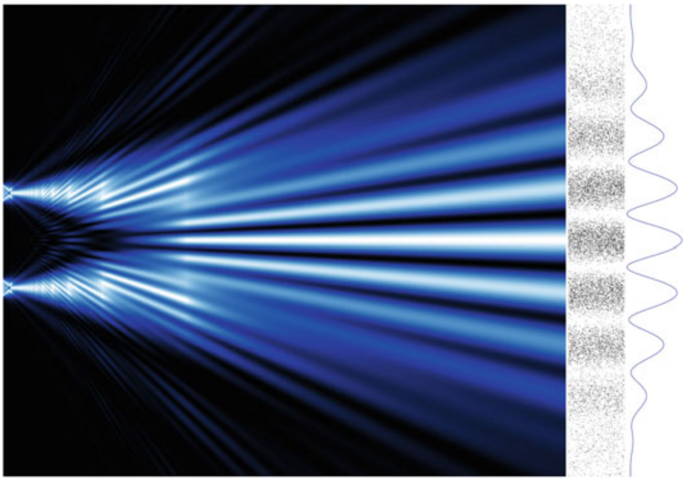

Fig. 4.10 The time evolution of the wave function in the situation of the double-slit experiment when both slits are open (the time evolution goes from *left* to *right*). The *white* and *blue* areas indicate places where the wave function is non zero and their intensity is proportional to the square of the wave function (*white* more intense, *blue* less intense). We see that at some places, the waves coming from both slits interfere constructively, and at other places, destructively. *Dots* on the right indicate the impact of particles. The *blue curve* on the right indicates the density of such impacts. This figure corresponds to part (c) of Fig. 2.6, but in two dimensions and interpreted in the language of the wave function (A. Gondran cc by-sa 4.0)

Since the quantum mechanical rules predict that the distribution of electrons detected on the second screen will be given by the square of the wave function corresponding to one slit being open, or the other one, or both, depending on what is open or not, we can understand, at least qualitatively, why the density given by the curve on the right of Fig. 4.10 can be less than the sum of the densities given by the curves on the right in Figs. 4.8 and 4.9 at some points (and larger at others).

We will illustrate this phenomenon via a simple mathematical example in the Appendix.

On the other hand, if we put a detector behind the first screen, as in Fig. 2.7, we can tell through which slits the particle went, so that the wave function

---

(Footnote 8 continued)
or, with $a = \Psi_1(x,t)$ and $b = \Psi_2(x,t)$,

$$(\Psi_1(x,t) + \Psi_2(x,t))^2 \neq (\Psi_1(x,t))^2 + (\Psi_2(x,t))^2, \tag{4.5}$$

namely the square of a sum is not equal to the sum of the squares!

collapses to either $\Psi_1(x, 0)$ or $\Psi_2(x, 0)$, because the detector performs a measurement and thus the collapse rule applies, and the interference pattern disappears: the resulting density of particles detected on the second wall will be given either by $\Psi_1(x, t)^2$ or $\Psi_2(x, t)^2$.

In the situation of Fig. 2.7, if we *do not* detect the particle going through the lower slit, it means that we "know" that it went through the upper one and thus the wave function collapses to $\Psi_1(x, 0)$ and the density of particles detected on the second wall will be given by $\Psi_1(x, t)^2$.

Finally, in the delayed-choice situation of Sect. 2.2, when no detection plate is inserted, the wave functions pass each other and, if t is now the time of arrival at the counter C_1 and C_2, the resulting detections will be given by $\Psi_1(x, t)^2$ or $\Psi_2(x, t)^2$, where $\Psi_1(x, t)^2$ will be concentrated at C_1 and $\Psi_2(x, t)^2$ will be concentrated at C_2 (see Fig. 2.9). On the other hand, with the detection plate P inserted, one gets the usual interference pattern of $(\Psi_1(x, t') + \Psi_2(x, t'))^2$, with t' now being the time when the waves reach the detection plate (see Fig. 2.10).

This finishes our explanation of how the quantum mechanical rules concerning the wave function predict the observed behavior in the double-slit experiment. But we do not claim to have *explained* what happens in the experiment. In fact, ordinary quantum mechanics usually emphasizes the fact that the theory does not explains what happens but only predicts it.

4.3 Einstein's Early Worries

At the Solvay Congress of 1927, which was a historical landmark in the discussions about the meaning of quantum mechanics, Einstein considered a particle going through a hole, as shown in Fig. 4.11. In the situation described in the picture, the wave function spreads itself over the half circle,[9] would but one always detects the particle in one piece at a given point, somewhere on the detection surface denoted by P in Fig. 4.11.[10]

Einstein raised the following objection:

> But the interpretation, according to which $(\Psi)^2$ expresses the probability that *this* particle is found at a given point,[11] assumes an entirely peculiar mechanism

---

[9] We speak of a half circle because the picture is two dimensional, but of course in three dimensions the spreading would be over a hemisphere.

[10] The situation is not very different than what happens in the two slits experiment when only one slit is open (see parts (a) and (b) of Fig. 2.6).

[11] Here Einstein writes $(\Psi)^2$ of what we write $\Psi(x)^2$, x being the point to which he refers (Note by J.B.).

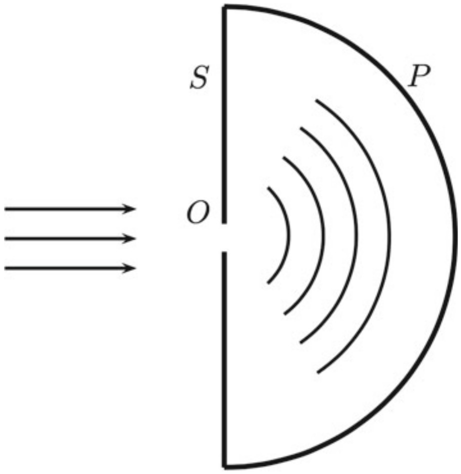

Fig. 4.11 A version of Einstein's objection at the 1927 Solvay Conference. Drawing by Travis Norsen. See [7, p. 440], or [182, p. 254] for the "original" (published in French translation at the time of the Solvay Conference)

> of action at a distance, which prevents the wave continuously distributed in space from producing an action in *two* places on the screen.
>
> Albert Einstein [7, p. 441]

Indeed, if the wave function is spread out in space before the particle is detected, then the fact that the latter is always detected at a given point implies that the wave function collapses on that point (or on a small neighborhood of that point), i.e. that it vanishes everywhere else instantaneously. Thus some sort of action at a distance (detecting the particle at one place makes the wave function vanish everywhere else) must be taking place. Einstein adds:

> In my opinion, one can remove this objection [action at a distance] only in the following way, that one does not describe the process solely by the Schrödinger wave, but that at the same time one localises the particle during the propagation.
>
> Albert Einstein [7, p. 441]

Here Einstein stresses the essential ambiguity concerning the meaning of $\Psi(x)^2$: does it describe the probability of the particle *being* somewhere or of the particle *being detected* somewhere if we detect its position, but not being anywhere before that? And Einstein also stresses that the second meaning, which is the orthodox view of quantum mechanics, implies this "peculiar mechanism of action at a distance", namely the collapse of the wave function at a point.

Einstein could not accept this "action at a distance" (nor did anybody else at that time), for reasons that will be discussed in Chaps. 7 and 10. In Chap. 8 we will see that the situation described by Einstein does *not* imply any action at a distance, but that this will be true only within a more complete theory than ordinary quantum mechanics, a theory in which "one localises the particle during the propagation".

The essence of Einstein's objection to the orthodox view of quantum mechanics, which is often misunderstood, was always based on this nonlocal aspect of the orthodox view.

4.4 Heisenberg's Inequality or "Uncertainty Principle"

Heisenberg's inequality is one of the mathematical consequences of the quantum formalism that is often presented as being one of the main "mysteries" of quantum mechanics (although we do not see it that way).

To explain this inequality, we have to mention that, besides measurements of positions, one can also measure velocities and that, given a wave function, one can compute the probability distribution of results of "measurements of the velocity".[12] Of course, since we stressed that, in ordinary quantum mechanics, particles do not have trajectories, the reader may wonder how they can possibly have a velocity, since the velocity simply quantifies the way the position of a particle changes along its trajectory. But this has again to do with the fact that we speak of "results of measurements", not of what exists independently of measurements and measurements may not simply measure some intrinsic property of the particle being measured, but may perturb the particle or interact with it in an uncontrollable way.

In 1927, right at the time when the quantum theory was being developed, Heisenberg discovered an inequality relating the probability distribution of the position measurements to the probability distribution of the velocity measurements.[13]

To be more precise, Heisenberg wrote an inequality relating how "spread out" the probability distribution of the position measurements is compared to how "spread out" the probability distribution of the velocity measurements

---

[12] We will come back to what measurements of velocity really are in Sect. 8.3.

[13] Heisenberg, and physicists generally, speak of momentum measurements rather than velocity measurements, but momentum is simply defined as the product of the mass times the velocity.

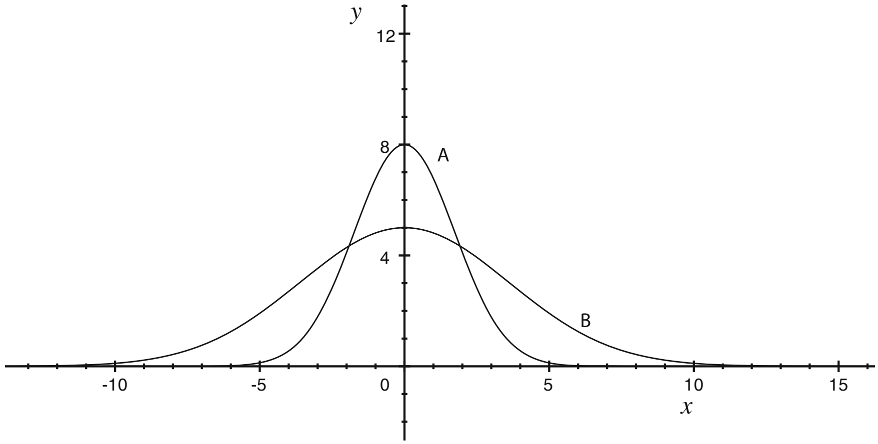

Fig. 4.12 The *curve B* is more "spread out" or more "flat" than the *curve A*

is. The notion of a probability distribution being "spread out" or "flat" is illustrated in Fig. 4.12.

In words, Heisenberg's inequality says that the less "spread out" the probability distribution of the position measurements is, the more "spread out" the probability distribution of the velocity measurements must be, and vice versa.[14]

But what does this inequality mean? Sometimes it is interpreted as showing that particles *do not have* well-defined positions and velocities before the latter are measured (which, of course, again puts back measurements on center stage).

But that is not what the inequality means, strictly speaking; it only relates how spread out the probability distributions of certain *measurements* (of positions and velocities) are and says nothing whatsoever about what happens outside of measurements or before them. Since a priori "measurements" might affect or disturb the system being measured and not simply reveal some property of it, we cannot conclude anything, at this stage, about what Heisenberg's inequality implies concerning what goes on outside of measurements.

---

[14] For the more mathematical reader: there is a standard way to measure how spread out a probability distribution is and it is given by the *variance* of that probability distribution. Let $\mathrm{Var}(x)$ denote the variance of the distribution of the measurements of the position x and $\mathrm{Var}(p)$ the variance of the distribution of the product of the mass m times the measurements of the velocity v: $p = mv$. What Heisenberg showed is that their product cannot be made arbitrarily small and satisfies a lower bound:

$$\mathrm{Var}(x)\mathrm{Var}(p) \geq \frac{1}{4}, \tag{4.6}$$

where the value $\frac{1}{4}$ depends on a choice of physical units that we shall not discuss. But, independently of this value, what this lower bound implies is that, if $\mathrm{Var}(x)$is very small, then $\mathrm{Var}(p)$ must be very large, and vice-versa.

Let us also mention that calling Heisenberg's inequality a "principle" (like in "uncertainty principle" or sometimes "indeterminacy principle") is misleading because it suggests that this is a principle independent of other principles of the quantum theory, while it is in fact a mathematical consequence of ordinary quantum mechanics (the purpose of this remark is only to clarify the status of this inequality, and obviously not to minimize the value of Heisenberg's discovery, which was quite extraordinary).

4.5 Conclusions

Coming back to our three fundamental questions, what does the duality of rules described here imply?

1. It suggests some sort of "reality created by the observer", since the rules of physics are different when one performs a measurement and when one does not.
2. Moreover, the rules are fundamentally statistical: one predicts the probabilities of events, but nothing is said about what might determine the behavior of each individual particle. Hence, quantum mechanics looks "intrinsically random".
3. There seems to be something nonlocal going on, which is what bothered Einstein already in 1927.

As we saw, the mysterious behavior described in Sects. 2.1 and 2.2 can easily be predicted by the quantum formalism. But we then face a basic puzzle: how can one understand all this talk about wave functions, and this duality in their rules of evolution, that singles out the role of measurements?

In the next chapter, we shall examine some natural sounding answers to those questions.

4.6 Summary

In this chapter, we introduced, without real formulas, the mathematical formalism used by physicists to predict the results described in Chap. 2. We emphasized that this formalism has to be viewed, at this point, only as an efficient recipe, but nothing more. All questions concerning its meaning will be discussed later.

The central concept is the one of the wave function Ψ, which, in quantum mechanics, is associated to any physical system.

The wave function $\Psi = \Psi(x)$ is a function of one variable for a single particle in one dimension of space, and its sole meaning is that $\Psi(x)^2$ determines the probability of finding the particle somewhere, when one detects its position. This is illustrated in Fig. 4.3.

It is important to understand that $\Psi(x)^2$ does *not* represent a density of "stuff" (mass or electric charge) nor does it determine the probability of the particle *being* somewhere, when its position is not measured. Indeed a wave function can in principle be as in Figs. 4.5 and 4.6. Since the regions A and B can be far apart, the density of stuff picture does not make sense. In that situations, one says that the wave function is a superposition (or a sum) of two wave functions $\Psi_1(x)$ and $\Psi_2(x)$.

One could of course say that the particle is either in region A or in region B *before* being observed. But that is not what the theory says. It is agnostic as to what happens before observations.

The reason for that attitude, which sounds paradoxical, is that, in order to account for the double-slit experiment, one has to assume that the wave function, when both slits are open, is a superposition (meaning a sum as in (4.1)) of two wave functions, one behind each slit, at the time when the "particle" (or the "wave") reaches the first wall, with the two slits in it. And, in that situation, quantum mechanics does not commit itself to saying that the particle goes through one slit or the other.

The evolution of the wave functions puts observations on center stage: when no observations are made, they propagate smoothly, and deterministically, like waves and their evolution is linear (see (4.2)). But when observations are made, they suddenly jump or get reduced: for example, if we measure whether the particle is in region A or B in Fig. 4.5, the wave function becomes $\Psi_1(x)$ or $\Psi_2(x)$, depending on the result. This is a sudden jump, which moreover is "random" in the sense that one can only attribute probabilities to its outcome (here one-half for each possibility, $\Psi_1(x)$ or $\Psi_2(x)$).

In the double-slit experiment, one has two wave functions, one behind each slit, when both slits are open. The interference pattern that we see in part (c) of Fig. 2.6 is obtained by letting these two wave functions evolve as waves between the two walls. Then, at some places, they interfere constructively, at other places, destructively.

But, if one observes through which slit the particle goes, for example by putting a detector behind one of the slits (see Fig. 2.7), then the wave function collapses and is reduced to the wave function behind the slit through which the particle goes: either the upper one if the particle is detected there or the lower one if it is not. Then, that remaining wave function evolves between the walls, but, since the other wave function is now absent, it no longer contributes to the production of an interference pattern.

We then mentioned an early objection of Einstein, that he repeated throughout his life: consider Fig. 4.11; here the wave function spreads itself on the half circle P, yet the particle is always detected at a single point. For Einstein, either the wave function is not the complete story and the particle has, besides its wave function, also a well defined position or the wave function is the complete story but then the fact that the particle is detected at a single point means that the wave function suddenly collapses at that point and thus vanishes everywhere else. But that is a nonlocal effect or "action at a distance" that Einstein could not accept.

We added a remark on Heisenberg's inequality, which is a mathematical relation saying that the less spread out the results of measurements of the position of the particles with a given wave function are, the more spread out will be the results of measurements of their velocities, and vice versa. Although this inequality is often interpreted as showing that particles do not have definite positions or velocities before being measured, we emphasized that precisely because the status of the wave function and the role of the measurements is unclear in ordinary quantum mechanics, no such definite conclusion can be drawn.

This shows how physicists deal with the phenomenon of interference, but it does not really explain anything, because the physical meaning of the wave function, and particularly of superpositions is unclear: it is an extremely accurate tool for predicting results of measurements, but, so far, nothing else.

Appendix

4.A The Wave Function

The first precision to be made about the wave function is that $\Psi(x)$ is in general a complex number and, to be correct, one should have written everywhere $|\Psi(x)|^2$ instead of $\Psi(x)^2$ in Sects. 4.1 and 4.2, where, for a complex number $z = a + ib$, $|z|^2 = a^2 + b^2$.

The fact that the total area under the curve in Fig. 4.3 is equal to one means that $\int_{\mathbb{R}} |\Psi(x)|^2 dx = 1$. This ensures that the probability of finding the particle somewhere is equal to one, as it should!

The probability of finding the particle in a region A is therefore $\int_A |\Psi(x)|^2 dx$, see Fig. 4.3.

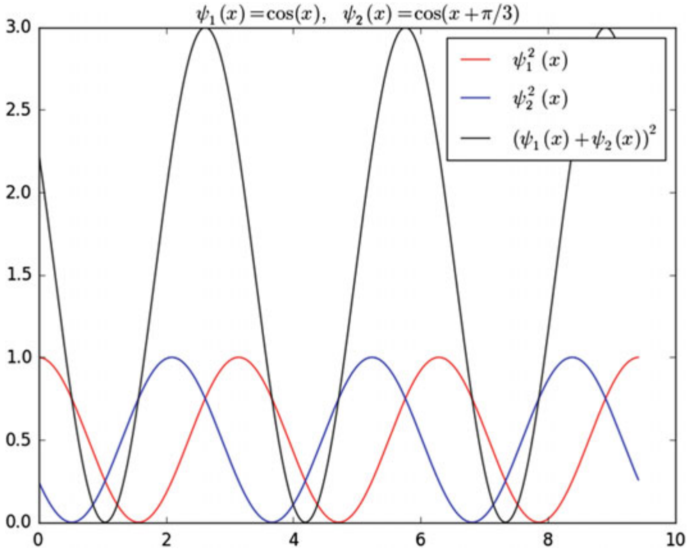

Fig. 4.13 The graphs of $\Psi_1(x)^2$, $\Psi_2(x)^2$ and $(\Psi_1(x) + \Psi_2(x))^2$, for two different wave functions $\Psi_1(x)$ and $\Psi_2(x)$ (for those who are familiar with trigonometric functions, we chose $\Psi_1(x) = \cos x$ and $\Psi_2(x) = \cos(x + \frac{\pi}{3})$, but this is not important) (A. Gondran cc by-sa 4.0)

Finally, in order to keep that constraint, each of the collapsed wave functions, after a measurement, must also satisfy $\int_{\mathbb{R}} |\Psi(x)|^2 dx = 1$. In Fig. 4.5, the situation is symmetric and, since the regions where $\Psi_1(x)$ and $\Psi_1(x)$ are non-zero do not overlap, one has $\int_{\mathbb{R}} |\Psi(x)|^2 dx = \int_{\mathbb{R}} |\Psi_1(x)|^2 dx + \int_{\mathbb{R}} |\Psi_2(x)|^2 dx = 1$ and thus $\int_{\mathbb{R}} |\Psi_1(x)|^2 dx = \int_{\mathbb{R}} |\Psi_2(x)|^2 dx = \frac{1}{2}$. So, the collapsed wave function is not $\Psi_1(x)$ or $\Psi_2(x)$, as we said in Sect. 4.2, but rather $\sqrt{2}\Psi_1(x)$ or $\sqrt{2}\Psi_2(x)$, that satisfy $\int_{\mathbb{R}} |\sqrt{2}\Psi_1(x)|^2 dx = \int_{\mathbb{R}} |\sqrt{2}\Psi_2(x)|^2 dx = 1$.

To illustrate the phenomenon of constructive and destructive interferences, consider Fig. 4.13, where the three curves represent the functions $\Psi_1(x)^2$, $\Psi_2(x)^2$ and $(\Psi_1(x) + \Psi_2(x))^2$ (we suppress the variable t here). We chose those functions (with the x axis drawn horizontally), so that they resemble (qualitatively) the wiggly blue curve on the right of Fig. 4.10. We note that the function $(\Psi_1(x) + \Psi_2(x))^2$ may vanish at points x where neither $\Psi_1(x)^2$ nor $\Psi_2(x)^2$ vanish. It can also be larger than the sum $\Psi_1(x)^2 + \Psi_2(x)^2$ for other x's. In the latter case, one says that the waves interfere constructively and in the former one that they interfere destructively.