On Unitary Evolution and Collapse in Quantum Mechanics

Francesco Giacosa

Institute of Physics, Jan Kochanowski University, Kielce, Poland
Institute for Theoretical Physics, Johann Wolfgang Goethe University, Frankfurt am Main, Germany
E-mail: giacosa@th.physik.uni-frankfurt.de

Editors: Chariton A. Dreismann & Danko Georgiev

Article history: Submitted on June 11, 2014; Accepted on August 21, 2014; Published on November 27, 2014.

In the framework of an interference setup in which only two outcomes are possible (such as in the case of a Mach–Zehnder interferometer), we discuss in a simple and pedagogical way the difference between a standard, unitary quantum mechanical evolution and the existence of a real collapse of the wavefunction. This is a central and not-yet resolved question of quantum mechanics and indeed of quantum field theory as well. Moreover, we also present the Elitzur–Vaidman bomb, the delayed choice experiment, and the effect of decoherence. In the end, we propose two simple experiments to visualize decoherence and to test the role of an entangled particle.


1 Introduction

Quantum mechanics is a well-established theoretical construct, which passed countless and ingenious experimental tests [1]. Still, it is renowned that quantum mechanics has some puzzling features [2–8]. Are macroscopic distinguishable superpositions (Schrödinger-cat states) possible or there is a limit of validity of quantum mechanics? Do measurements imply a non-unitary (collapse-like) time evolution or are they also part of a unitary evolution? In the latter case, should we simply accept that the wavefunction splits in many branches (i.e., parallel worlds), which decohere very fast and are thus independent from each other? It is important to stress that these issues are not only central in nonrelativistic quantum mechanics but apply also in relativistic quantum field theory. Namely, the generalization to quantized fields does not modify the role of measurements. In this work we discuss in a introductory way some of the questions mentioned above. We study the quantum interference in an idealized two-slit experiment and we analyze the effect that a detector measuring “which path has been taken” has on the system. In particular, we shall concentrate on the collapse of the wavefunction, such as the one advocated by collapse models [7–14] and show which are the implications of it.

Variants of our setup also lead us to the presentation of the famous Elitzur–Vaidman bomb [15] and to delayed choice experiments [16–17]. Thus, we can describe in a unified framework and with simple mathematical steps (typical of a quantum mechanical course) concepts related to modern issues and experiments of quantum mechanics.

Besides the pedagogical purposes of this work, we also aim to propose two experiments (i) to see decoherence at work in an interference setup with only two possible outcomes and (ii) to test the dependence of the interference on an idler entangled particle.
2 Collapse vs no-collapse: no difference?

2.1 Interference setup

We consider an interference setup as the one depicted in Fig. 1. A particle $P$ flies toward a barrier which contains two ‘slits’ and then flies further to a screen $S$. Usually in such a situation there is a superposition of waves which generates on the screen $S$ many maxima and minima. We would like to avoid this unnecessary complication here but still use the language of a double-slit experiment in which a sum over paths is present. To this end, we assume that the particle can hit the screen in two points only, denoted as $A$ and $B$. All the issues of quantum mechanics can be studied in this simplified framework. We assume also to ‘sit on’ the screen $S$: when the particle hits $A$ or $B$ we ‘see’ it.

First, we consider the case in which only the left slit is open, see Fig. 1a. In order to achieve our goal, the slit is actually not a simple hole in the barrier (out of which a spherical wave would emerge) but a more complicated filter which projects the particle either to a straight trajectory ending in $A$ or to a straight trajectory ending in $B$. In the language of quantum mechanics, this situation amounts to a wavefunction $|L\rangle$ associated to the particle which has gone through the left slit, which is assumed to be

$$|L\rangle = \frac{1}{\sqrt{2}}(|A\rangle - |B\rangle). \quad (1)$$

Then, by simply using the Born rule (i.e., by squaring the coefficient multiplying $|A\rangle$ or $|B\rangle$), we predict that the particle ends up either in the endpoint $A$ with probability 50% or in the endpoint $B$ with probability 50%. This is indeed what we measure by repeating the experiment many times. As we see, the probability is a fundamental ingredient of quantum mechanics, which however enters only in the very last step, i.e. when the measurement comes into the game. The state $|L\rangle$ is an equal (antisymmetric) superposition of $|A\rangle$ and $|B\rangle$, but in a single experiment we do not find a pale spot on $A$ and a pale spot on $B$: we always find the particle either fully in $A$ or in $B$. It is only after many repetitions of the experiment that we realize that the outcome $A$ and the outcome $B$ are equally probable.

If only the right slit is open, see Fig. 1b, we have a similar situation in which only two trajectories ending in $A$ and in $B$ are present. The wavefunction of the particle after having gone through the right slit is denoted by $|R\rangle$ and is described by the orthogonal combination to $|L\rangle$:

$$|R\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle). \quad (2)$$

Figure 1: Hypothetical experiment with only two possible outcomes ($A$ and $B$). Panel (a) only the left slit is open. Panel (b) only the right slit is open. Note, each slit is not a simple hole but acts as a filter which projects the particle either to a trajectory with endpoint $A$ or to a trajectory with endpoint $B$. Panel (c) shows the experiment with both slits open: interference takes place and all particles hit the screen in $A$. 

Quanta | DOI: 10.12743/quanta.v3i1.26

November 2014 | Volume 3 | Issue 1 | Page 157
In this case one also finds the particle in 50% of cases in A and 50% in B.

We now turn to the case in which both slits are open, see Fig. 16. The wave function of the particle is assumed to be the sum of the contributions of the two slits:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|L\rangle + |R\rangle),$$

(3)
i.e. the contributions of both slits add coherently. A simple calculation shows that

$$|\Psi\rangle = |A\rangle,$$

(4)
which means that the particle $P$ always hits the screen in A and never in $B$. Namely, in A we have a constructive interference, while in $B$ we have a destructive interference. (Notice that the points $A$ and $B$ are not equidistant from the two slits. However, we take the two slits as being close to each other and the points $A$ and $B$ as being far from each other: the difference between the segments $LA$ and $RA$ (and so between $LB$ and $RB$) is assumed to be negligible such that the two contributions of the wave packet of the particle $P$ from the left and right slit arrive almost simultaneously and the depicted interference effect takes place).

In conclusion, we have chosen the language of a two-slit experiment because it is the most intuitive. The price to pay is a slit acting as a filter and not as a simple hole. However, one can easily build analogous setups as the one here described by using photon polarizations, electron spins or equivalent quantum objects, or by using a Mach–Zehnder interferometer, see details in §2.3.3.

2.2 Detector measuring the path

As a next step we put a detector $D$ right after the two open slits. $D$ measures through which hole the particle has passed, without destroying it. We analyze the situation in two ways: first, by assuming the collapse of the wavefunction as induced by $D$ and, second, by studying the entanglement of the particle with the detector. Note, we still assume that we sit on (or watch) the screen $S$ only, but we are not directly connected to the detector $D$.

2.2.1 Collapse

In this case we assume that the detector $D$ generates a collapse of the wavefunction. Suddenly after the interaction with $D$, the state of the particle $P$ collapses into $|L\rangle$ with a probability of 50% or into $|R\rangle$ with a probability of 50%. Then, the state is described by either $|L\rangle$ or $|R\rangle$, but not any longer by the superposition of them. As a consequence, we have in half of the cases a situation analogous to having only the left slit open and in the other half to having only the right slit open.

What we will then see on the screen $S$? The probability to find the particle in $A$ is given by

$$P[A] = P[L, A] + P[R, A] = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2},$$

(5)
where $P[L, A] = 1/4$ is the probability that the detector $D$ has measured the particle going through the left slit and then the particle has hit the screen in $A$. Similarly, $P[R, A] = 1/4$ is the probability that the detector $D$ has measured the particle going through the right slit before the latter hits $A$. For $P[B]$ holds a similar description

$$P[B] = P[L, B] + P[R, B] = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}.$$  

(6)
The collapse is obviously part of the standard interpretation of quantum mechanics, in which a detector is treated as a classical object which induces the collapse of the quantum state. As a result, there is no interference on the screen $S$. As renowned, the standard interpretation does not put any border between what is a classical system and what is a quantum system. Nevertheless, one can interpret the collapse postulate as an effective description of a physical process. Namely, in theories with the collapse of the wavefunction, the collapse is a real physical phenomenon which takes place when one has a macroscopic displacement of the position wavefunction of the detector (or, more generally, of the environment). In this framework, somewhere in between the quantum world and the classical macroscopic world, a new physical process takes place which realizes the collapse: this could be, for instance, the stochastic hit in the Ghirardi-Rimini-Weber model [7] or the instability due to gravitation in the Penrose-Diosi approach [8,12,13]. Neglecting details, the main point is that such collapse theories realize physically the collapse which is postulated in the standard interpretation and liberates it from inconsistencies. Still, it is an open and well posed physical question if (at least one of) such collapse theories are (is) correct.

2.2.2 No-collapse

In this case we do not assume that the detector $D$ generates a collapse of the wavefunction, but we enlarge the whole wave function of the system by including also the wavefunction of the detector. We assume that, prior to measurement, the detector is in the state $|D_0\rangle$ (we can, for definiteness, think of a old-fashion indicator which points to 0). Then, when both slits are open, the state of the whole system just after having passed through them but not yet in contact with the detector $D$, is given by

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|L\rangle + |R\rangle)|D_0\rangle.$$  

(7)
Then, the particle-detector interaction induces a (we assume very fast) time evolution which generates the following state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left[(|L\rangle |D_L\rangle + |R\rangle |D_R\rangle)\right], \quad (8)$$

where $|D_L\rangle$ ($|D_R\rangle$) describes the pointer of the detector pointing to the left (right). Thus, no collapse is here taken into account, because the whole wavefunction still includes a superposition of $|L\rangle$ and $|R\rangle$, which, however, are now entangled with the detector states $|D_L\rangle$ and $|D_R\rangle$, respectively.

An important point is that the overlap of $|D_L\rangle$ and $|D_R\rangle$ is small

$$\langle D_L|D_R\rangle \approx 0, \quad (9)$$

to a very good degree of accuracy. To show it, let us ignore the rest of the detector and the environment and concentrate on the pointer only, which is assumed to be made of $N$ atoms, where $N$ is of the order of the Avogadro constant. The atom $\alpha$ of the pointer is in a superposition of the type $\left(\psi^a_L(\vec{x}) + \psi^a_R(\vec{x})\right)/\sqrt{2}$, where $\psi^a_L(\vec{x})$ ($\psi^a_R(\vec{x})$) is the wavefunction of the atom when the pointer points to the left (right). We have

$$\langle D_L|D_R\rangle = \prod_{\alpha=1}^{N} \int d^3x \left(\psi^a_L(\vec{x})\right)^* \psi^a_R(\vec{x}). \quad (10)$$

The quantity $\int d^3x \left(\psi^a_L(\vec{x})\right)^* \psi^a_R(\vec{x}) = \lambda_\alpha$ is such that $|\lambda_\alpha| < 1$. For a large displacement, $\lambda_\alpha$ is itself a very small number (small overlap), but the crucial point is to observe that $\langle D_L|D_R\rangle$ is the product of many numbers with modulus smaller than 1. Assuming that $\lambda_\alpha = \lambda$ for each $\alpha$ (each atom gets a similar displacement: this assumption is crude but surely sufficient for an estimate), we get

$$\langle D_L|D_R\rangle \approx \lambda^N, \quad (11)$$

which is extremely small for large $N$. Even if we take $\lambda = 0.99$ (which is indeed quite large and actually overestimates the overlap of the wave functions of an atom belonging to macroscopic distinguishable configuration), we obtain

$$\langle D_L|D_R\rangle \approx 0.99^N \approx 10^{-10^{31}} \quad (12)$$

which is tremendously small.

After having clarified the de facto orthogonality of $|D_L\rangle$ and $|D_R\rangle$, we rewrite the full wavefunction of the system $|S\rangle$ as

$$|\Psi\rangle = \frac{1}{2} \left[|A\rangle \langle D_R| + |L\rangle \langle D_L| \right]. \quad (13)$$

Then, the probability to find the particle $P$ in $A$ is obtained (now by using the Born rule, because we are observing the screen $S$):

$$P[A] = P[L, A] + P[R, A] = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2} \quad (14)$$

where $P[L, A] = 1/4$ is the probability that the system is described by $|A\rangle |D_L\rangle$ and $P[R, A] = 1/4$ the probability that it is described by $|A\rangle |D_R\rangle$. A similar situation holds for $P[B] = 1/2$. Thus, also in this case the presence of $D$ causes the disappearance of interference.

The same result is obtained if we use the formalism of the statistical operator, which is defined by $\hat{\rho} = |\Psi\rangle \langle \Psi|$ (see, for instance, Refs. [1,7]). Upon tracing over the detector states (environment states) the reduced statistical operator reads (we use here $\langle D_L|D_R\rangle = 0$):

$$\hat{\rho}_{\text{red}} = \langle D_L|\hat{\rho}|D_L\rangle + \langle D_R|\hat{\rho}|D_R\rangle = \left(|A\rangle \langle B| \right) \left( \frac{1}{2} \ 0 \ \frac{1}{2} \right) \left( |A\rangle \langle B| \right), \quad (15)$$

where the diagonal elements represent $p[A] = p[B] = 1/2$ respectively, while the off-diagonal elements vanish in virtue of the (for all practical purposes) orthogonality of $|D_L\rangle$ and $|D_R\rangle$.

### 2.2.3 Summary

We find that, for us sitting on the screen $S$, the very same outcome, i.e. the absence of interference, is obtained by applying the collapse postulate as an intermediate step due to the detector $D$ or by considering the whole quantum state (including the detector $D$) and by applying the Born rule only in the very end. This equivalence holds as long as the (anyhow very small) overlap of the detector states of Eq. (12) is neglected (see also the related discussion in §3). The question is then: do we need the collapse? The second calculation (no-collapse) seems to answer us: ‘no, we don’t’. In this respect, one has a superposition of macroscopic distinct states, which coexist and are nothing else but the branches of the Everett’s many worlds interpretation of quantum mechanics [18]. Thus, assuming that no collapse takes place brings us quite naturally to the many worlds interpretation (possibly [19][21]. (Originally, Everett [18] introduced the concept of ‘relative state formulation’, which was reinterpreted as the many worlds interpretation by Wheeler and Dewitt [19][20]. The many worlds interpretation is the most natural interpretation when no collapse is present, but the definition of what is a ‘world’ is not trivial. Intuitively, it is a piece of the wavefunction which is a pointer-state, i.e. it does not contain spacial superpositions of macroscopic objects. Other points of view, such as ‘many histories’ and ‘many minds’ were also considered.)
However, care is needed: in fact, the ‘no collapse’ assumption is a general statement and means also that there is no collapse when the particle $P$ hits the screen $S$ (where our own wavefunction is part of the game). Let us clarify better this point by going back to the very first case we have studied, in which only the left slit was open and no detector $D$ was present (Fig. 1a). The wavefunction of the particle just before hitting the screen is given by Eq. (1). But then, after the hit and assuming no collapse, the whole wavefunction (including us, who are the observers) reads:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} |A\rangle \langle\text{Screen recording A and we observing A}| - \frac{1}{\sqrt{2}} |B\rangle \langle\text{Screen recording B and we observing B}|.$$  

(16)

The question is why the coefficient in front of the vector $|A\rangle \langle\text{Screen recording A and we observing A}|$ tells us which is the subjective probability of observing $A$ for the observer (us) sitting on the screen. In other words, how does the many worlds interpretation explain the probabilities according to the Born rule? The Born rule seems to be an additional postulate, which has to be put $ad$ $hoc$ into it. This situation is however not satisfactory, because the main idea of the many worlds interpretation is to eliminate the collapse from the description of the quantum mechanics and consequently to derive the standard Born probabilities. Although there are attempts to show that there is no need of postulating the Born rule in this context [22,24] (see also Ref. [25]), no agreement has been reached up to now [7,26,27]. (Notice that in the case of Eq. (16) one could understand the many worlds interpretation by noticing that there are two worlds, ergo the subjective probability to be in one of those is 50% in agreement with the Born rule. However, this is a particular case with equal coefficients. When the coefficients in front of the kets are not $1/\sqrt{2}$ (but say $a$ and $b$ with $|a|^2 + |b|^2 = 1$) one still has two worlds but the subjective probability to be in one of those is not 1/2, but the one given by the Born rule ($|a|^2$ and $|b|^2$ respectively). This is exactly the point discussed in Refs. [22,24,26,27] with, however, different conclusions.) This is indeed an argumentation in favor of the possibility that a collapse really takes place. Surely, ‘real collapse’ scenarios deserve to be studied theoretically and experimentally [7,29,30].

Note, up to now we did not mention the decoherence, see e.g. Refs. [2,28,32] and references therein. This is possible because we have put a detector that makes a measurement by evolving from the state $|D_0\rangle$ into two (almost) orthogonal states $|D_L\rangle$ and $|D_R\rangle$, but actually one can interpret this fast change of the detector state as the result of a decoherence phenomenon. This is however a rather peculiar decoherence, because we have prepared the detector in a particular (low entropic) $|D_0\rangle$ state, which is ‘ready to’ evolve into $|D_L\rangle$ and $|D_R\rangle$ as soon as it interacts with the particle $P$. In § 3 we will describe what changes when the environment, instead of the detector, is taken into account.

### 2.3 Variants of the setup

#### 2.3.1 The bomb

A simple change of the setup allows us to present the famous Elitzur–Vaidman bomb, first described in Ref. [15] and then experimentally verified in Ref. [33]. We substitute the detector with a ‘bomb’, which can be activated by the particle $P$. We place the bomb only in front of the left slit, see Fig. 2. This means that, if only the left slit is open, the bomb explodes soon after the particle has gone through the slit. If, instead, only the right slit is open, it doesn’t explode. For definiteness and simplicity we assume that the particle is not destroyed nor absorbed by the bomb.

Just as previously, we can interpret the experiment applying either the collapse or by studying the whole wavefunction. In the collapse approach, the bomb simply makes a measurement. When both slits are open the wavefunction, before the interaction with the bomb, is given by Eq. (3) we will have an explosion in 50% of cases and no explosion in the remaining 50%. Notice that in the second case the bomb is doing a null measurement. The very fact that the bomb does not explode means that the particle went to the right slit (we assume 100% efficiency in our ideal experiment). When the bomb explodes there is a collapse into $|\Psi\rangle = |L\rangle$, when it does not explode there is a collapse into $|\Psi\rangle = |R\rangle$. Then, we have a situation which is very similar to the case of the detector $D$ which we have studied previously: no interference on the screen $S$ is observed, but we observe the particle in the endpoint $A$ and $B$ with probability 1/2 each.

If we do not assume the collapse of the wavefunction, the whole wavefunction is given by (after interaction with the bomb)

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|L\rangle |B_E\rangle + |R\rangle |B_0\rangle)$$

$$= \frac{1}{2} (|A\rangle (|B_0\rangle + |B_E\rangle) + |B\rangle (|B_0\rangle - |B_E\rangle))$$

(17)

where $|B_0\rangle$ is the state describing the unexploded bomb and $|B_E\rangle$ the exploded one. Obviously, as in Eq. (12) we have $\langle B_E|B_0\rangle \approx 0$. Again and just as before no interference is seen on $S$ but the two outcomes $A$ and $B$ are...
Figure 2: Variant of the Elitzur–Vaidman experiment: a bomb is placed just after the left slit.

Another interesting configuration is obtained by assuming that a second entangled particle, denoted as $I$ (for idler), is emitted when $P$ goes through the slit(s). The system is built in the following way: if the particle $P$ goes through the left slit, the particle $I$ is described by the state $|I_L\rangle$. Similarly, when the particle $P$ goes to the right slit, the particle $I$ is described by the state $|I_R\rangle$. We assume that the two idler states are orthogonal: $\langle I_L|I_R\rangle = 0$. This situation resembles closely that of delayed choice experiments [16][17].

When both slits are open the whole wavefunction of the system is given by

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|L\rangle |L\rangle + |R\rangle |I_R\rangle)$$

$$= \frac{1}{\sqrt{2}} (|A\rangle |I_+\rangle + |B\rangle |L\rangle) ,$$

(19)

where

$$|I_+\rangle = \frac{1}{\sqrt{2}} (|I_R\rangle + |I_L\rangle)$$

(20)

$$|L\rangle = \frac{1}{\sqrt{2}} (|I_R\rangle - |I_L\rangle) .$$

(21)

The idler particle $I$ is entangled with the particle $P$, but being a microscopic object, we surely cannot apply the collapse hypothesis because the particle $I$ is not a measuring apparatus.

Do we have interference on the screen $S$ in this case? The answer is clear: no. The states $|A\rangle |I_L\rangle$, $|A\rangle |I_R\rangle$, $|B\rangle |I_L\rangle$, $|B\rangle |I_R\rangle$ represent a basis of this system, thus the probability to obtain $|A\rangle$ (that is, the probability of $P$ hitting $S$ in $A$) is $1/4 + 1/4 = 1/2$. So for $B$. The presence of the entangled idler state destroys the interference on $S$.

It is sometime stated that this result is a consequence of the fact that the state of the idler particle $I$ carries the information of which way $P$ has followed. For this reason, the interference has disappeared (this is a modern reformulation of the complementarity principle). However, such expressions, although appealing, are often too vague and need to be taken with care.

As a next step we study what happens if we perform a measurement on the idler particle $I$. We study separately two distinct types of measurements.

Measuring the idler particle in the $|I_L\rangle,|I_R\rangle$ basis

First, we perform a measurement which tells us if the state of the idler particle is $|I_L\rangle$ or $|I_R\rangle$. For simplicity, we apply the collapse hypothesis (as usual, the results would not change by keeping track of the whole unitary quantum evolution). But first, we have to clarify the following

\[ |L\rangle = \frac{1}{\sqrt{2}} (|I_R\rangle - |I_L\rangle) . \]
issue: when do we perform the measurement on $I$? We have two possibilities:

- If we measure the state of $I$ before the particle $S$ hits the screen, the wavefunction reduces to $|L\rangle |I_L\rangle$ or to $|R\rangle |I_R\rangle$ with $50\%$ probability, respectively. Then, the screen $S$ performs a second measurement: we find as usual $50\%$ of times $A$ ($25\% |A\rangle |I_L\rangle$ and $25\% |A\rangle |I_R\rangle$) and $50\%$ of times $B$ ($25\% |B\rangle |I_L\rangle$ and $25\% |B\rangle |I_R\rangle$).

- If, instead, the particle $P$ arrives first on the screen $S$, the quantum state collapses into $|A\rangle |I_A\rangle$ in $50\%$ of cases ($A$ has clicked), or into $|B\rangle |I_B\rangle$ in the other $50\%$ of cases ($B$ has clicked). The subsequent measurement of the $I$ particle will then give $|I_L\rangle$ or $|I_R\rangle$ ($50\%$ each).

In conclusion, we realize that it is absolutely irrelevant which experiment is done before the other. In particular, for us sitting on the screen $S$, it does not matter at all when and if the measurement of the idler state is performed. We simply see no interference.

**Measuring the idler particle in the $|I_+,I_-\rangle$ basis**

Being the particle $P$ entangled with another particle and not with a macroscopic state, we can also decide to perform a different kind of measurement on $I$. For instance, we can put a detector measuring $I$ by projecting onto the basis $|I_+\rangle$ and $|I_-\rangle$. If we do this measurement before the particle $P$ has hit the screen $S$, we have the following outcome as a consequence of the collapse induced by the $I$-detector:

$$|\Psi\rangle = |A\rangle |I_+\rangle \text{ with probability } 50\%; \quad (22)$$

$$|\Psi\rangle = |B\rangle |I_-\rangle \text{ with probability } 50\%. \quad (23)$$

In the former case, the particle $P$ will surely hit $S$ in $A$, in the latter case the particle $P$ will surely hit $S$ in $B$.

One sometimes interpret the experiment in the following way: the detector measuring the state of $I$ as being either $|I_+\rangle$ or $|I_-\rangle$ ‘erases the which-way information’. When the detector measures $|I_+\rangle$ we still have interference and we see the particle $P$ in the position $A$, just as the case with two open slits (Fig. 1). In the other case, when the detector measures $|I_-\rangle$, we also have a kind of interference in which the final position $B$ is the only outcome. In the language of Ref. [16], one speaks of ‘fringes’ in the former case, and of ‘anti-fringes’ in the latter.

However, care is needed: for us sitting on $S$, if we do not know which measurement is performed on $I$, we simply see that no interference occurs ($50\%$-$A$ and $50\%$-$B$). But, if we could then speak with a colleague working with the $I$-detector, we would realize that, each time we have measured $A$ he has found the state $|I_+\rangle$, while each time we have measured $B$ he has found $|I_-\rangle$. Thus, we have a correlation of our results (measurement of the screen $S$) with those of the $I$-detector. This is actually no surprise if we look at the quantum state of Eq. [19] This statement is indeed more precise than the statement of having interference because we have erased the which-way information. Namely, we do not have interference.

Indeed, we can perform the measurement of $I$ even after (in principle long time after) the screen $S$ has measured $P$ in either $A$ or $B$. Here the name ‘delayed choice’ comes from: we choose if we retain the which-way information or not. Still, the result is the same because there is no influence on the time-ordering of the measurements. If the measurement of the screen $S$ occurs first, we have a collapse onto the very same Eqs. (22) and (23). Then, a measurement of the idler particle $I$ would simply find either $|I_+\rangle$ correlated with $A$ or $|I_-\rangle$ correlated with $B$. For sure, there is no change of the past by a measurement of the idler state, but simply a correlation of states. Still, such a very interesting setup visualizes many of the peculiarities of quantum mechanics and can be used for quantum cryptography.

### 2.3.3 Realizations of the setup

In a two-slit experiment all the peculiarities of quantum mechanics are evident due to the fact that the particle $P$ follows (at least) two paths at the same time. This is extremely fascinating as well as counterintuitive for our imagination based on a childhood with rolling ‘classical’ marbles. However, as already mentioned in §2.1, a simple implementation of the two-slit experiment does not produce only two possible outcomes, but gives rise to a superposition of waves with many maxima and minima. In the following we present two possible realizations of our Gedankenexperiment which do not make use of slits.

An interference experiment in which only two outcomes are possible can be realized by using particles with spin $1/2$ (such as electrons in a Stern-Gerlach-type experiment) or photons (spin $1$, but due to gauge invariance only two polarizations are realized). Clearly, all the quantum mechanics features do not depend on which particle or on which quantum number are implemented, but solely on the presence of superpositions and on the effect of measurements. In the case of photon polarizations we can use the fact that a photon can be horizontally or vertically polarized (corresponding to the kets $|\uparrow\rangle$ and $|\downarrow\rangle$ respectively). In our analogy, the state $|\uparrow\rangle$ corresponds to the state of our particle $P$ coming out from the left slit, $|\uparrow\rangle \equiv |L\rangle$, and similarly $|\downarrow\rangle$ from the right slit, $|\downarrow\rangle \equiv |R\rangle$.

Then, we place a detector which acts as the screen $S$ by
making a measurement in the basis $|A\rangle = (|v\rangle + |h\rangle)/\sqrt{2}$ and $|B\rangle = (|v\rangle − |h\rangle)/\sqrt{2}$. In addition, we can place a second detector which plays the role of the detector $D$ by measuring the polarization in the $|h\rangle,|v\rangle$ basis. Indeed, in this case we do not need to send the photons along two different paths, because the polarization degree of freedom is enough for our purposes.

Another possible realization of our setup is the Mach–Zehnder interferometer [34, 35], see Fig. 3 which makes use of beam splitters. When a photon is sent to path $|1\rangle$ of Fig. 3a, both photon counters $A$ and $B$ can detect the photon with a probability of 50%. For our analogy, we have $|1\rangle \equiv |L\rangle$. Similarly, when the photon is sent to path $|2\rangle$ of Fig. 3b, we hear a click in detector $A$ and destructive interference occurs at detector $B$. In addition, we can place a beam splitter in the beginning of the setup as shown in Fig. 3c after the photon passes through, we get a superposition $(|1\rangle + |2\rangle)/\sqrt{2}$. The inclusion of the detector $D$, the bomb, entangled particle(s) as well as the environment evolves as function of time $t$ as

$$|\Psi(t)\rangle = |L\rangle |E_L(t)\rangle,$$

(24)

where by construction $|E_L(0)\rangle = |E_0\rangle$. Similarly, if only the right slit is open, at the time $t$ the system is described by $|\Psi(t)\rangle = |R\rangle |E_R(t)\rangle$ with $|E_R(0)\rangle = |E_0\rangle$.

We now turn to the case in which both slits are open. It is important to stress that, by assuming a weak interaction of the particle $P$ with the environment, we surely do not have (at first) a collapse of the wavefunction, but an evolution of the whole quantum state given by

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} \left( |L\rangle |E_L(t)\rangle + |R\rangle |E_R(t)\rangle \right) \nonumber$$

$$= \frac{1}{\sqrt{2}} \left[ |L\rangle (|E_L(t)\rangle + |E_R(t)\rangle) + |R\rangle (|E_R(t)\rangle − |E_L(t)\rangle) \right].$$

(25)

This is indeed very similar to the detector case, but there is a crucial aspect that we now take into consideration. The state $|E_L(t)\rangle$ and $|E_R(t)\rangle$ coincide at $t = 0$ and then smoothly depart from each other. At the time $t$ we assume to have

$$c(t) = \langle E_L(t)|E_R(t)\rangle = e^{-\lambda t}.$$

(26)

In this section we show that there is a difference between the collapse and no-collapse scenarios. To this end, instead of having a detector, a bomb, or an idler entangled state, we assume that the space between the slits and the screen is not the vacuum. Then, we study the time evolution of the environment which interacts with the particle $P$. This interaction is assumed to be soft enough not to absorb or kick away the particle in such a way that the final outcomes on the screen $S$ are still the endpoints $A$ or $B$.

Before the particle $P$ goes through the slit(s), the environment is described by the state $|E_0\rangle$. First, we study the case in which only the left slit is open. Denoting as $t = 0$ the time at which $P$ passes through the left slit, the wavefunction of the environment evolves as function of time $t$ as

$$|\Psi(t)\rangle = |L\rangle |E_L(t)\rangle,$$

(24)

where by construction $|E_L(0)\rangle = |E_0\rangle$. Similarly, if only the right slit is open, at the time $t$ the system is described by $|\Psi(t)\rangle = |R\rangle |E_R(t)\rangle$ with $|E_R(0)\rangle = |E_0\rangle$.

We now turn to the case in which both slits are open. It is important to stress that, by assuming a weak interaction of the particle $P$ with the environment, we surely do not have (at first) a collapse of the wavefunction, but an evolution of the whole quantum state given by

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} \left( |L\rangle |E_L(t)\rangle + |R\rangle |E_R(t)\rangle \right) \nonumber$$

$$= \frac{1}{\sqrt{2}} \left[ |L\rangle (|E_L(t)\rangle + |E_R(t)\rangle) + |R\rangle (|E_R(t)\rangle − |E_L(t)\rangle) \right].$$

(25)

This is indeed very similar to the detector case, but there is a crucial aspect that we now take into consideration. The state $|E_L(t)\rangle$ and $|E_R(t)\rangle$ coincide at $t = 0$ and then smoothly depart from each other. At the time $t$ we assume to have

$$c(t) = \langle E_L(t)|E_R(t)\rangle = e^{-\lambda t}.$$

(26)

In this section we show that there is a difference between the collapse and no-collapse scenarios. To this end, instead of having a detector, a bomb, or an idler entangled state, we assume that the space between the slits and the screen is not the vacuum. Then, we study the time evolution of the environment which interacts with the particle $P$. This interaction is assumed to be soft enough not to absorb or kick away the particle in such a way that the final outcomes on the screen $S$ are still the endpoints $A$ or $B$.

Before the particle $P$ goes through the slit(s), the environment is described by the state $|E_0\rangle$. First, we study the case in which only the left slit is open. Denoting as $t = 0$ the time at which $P$ passes through the left slit, the wavefunction of the environment evolves as function of time $t$ as

$$|\Psi(t)\rangle = |L\rangle |E_L(t)\rangle,$$

(24)

where by construction $|E_L(0)\rangle = |E_0\rangle$. Similarly, if only the right slit is open, at the time $t$ the system is described by $|\Psi(t)\rangle = |R\rangle |E_R(t)\rangle$ with $|E_R(0)\rangle = |E_0\rangle$.

We now turn to the case in which both slits are open. It is important to stress that, by assuming a weak interaction of the particle $P$ with the environment, we surely do not have (at first) a collapse of the wavefunction, but an evolution of the whole quantum state given by

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} \left( |L\rangle |E_L(t)\rangle + |R\rangle |E_R(t)\rangle \right) \nonumber$$

$$= \frac{1}{\sqrt{2}} \left[ |L\rangle (|E_L(t)\rangle + |E_R(t)\rangle) + |R\rangle (|E_R(t)\rangle − |E_L(t)\rangle) \right].$$

(25)

This is indeed very similar to the detector case, but there is a crucial aspect that we now take into consideration. The state $|E_L(t)\rangle$ and $|E_R(t)\rangle$ coincide at $t = 0$ and then smoothly depart from each other. At the time $t$ we assume to have

$$c(t) = \langle E_L(t)|E_R(t)\rangle = e^{-\lambda t}.$$

(26)
less and less by the time passing. The constant $\lambda$ describes the speed of the decoherence and depends on the number of particles involved and the intensity of the interaction. Note, strictly speaking, this non-orthogonality is also present in the case of the detector (if no collapse is assumed), but the overlap is amazingly small, see the estimate in Eq. [12] (In the case of the detector $D$ of § 2.2 $\lambda$ is very large and consequently $\lambda^{-1}$ is a very short time scale, shorter than any other time scale in that setup. For that reason we assumed that the detector state evolved for all practical purposes instantaneously from the ready-state (pointer at 0) to pointing either to the left or to the right.)

Now we ask the following question: what is the probability that the particle $P$ hits the screen in $A$? We assume that the particle $P$ hits the screen at the time $\tau$. At this instant, the state is given by $|\Psi(\tau)\rangle$ with $\langle E_L(\tau)|E_R(\tau)\rangle = c(\tau)$.

We now present the mathematical steps leading to $p[A, \tau]$, which, although still simple, are a bit more difficult than the previous ones. The reader who is only interested in the result can go directly to Eq. [31].

At the time $\tau$ we express the state $|E_L(\tau)\rangle$ as

$$|E_L(\tau)\rangle = c(\tau)|E_R(\tau)\rangle + \sum_\alpha b_\alpha(\tau)|E^\alpha_{R,\perp}(\tau)\rangle \tag{27}$$

where the summation over $\alpha$ includes all states of the environment which are orthogonal to $|E_R(\tau)\rangle$: $\langle E^\alpha_{R,\perp}(\tau)|E_R(\tau)\rangle = 0$. This expression is possible because the set $\{E_R(\tau), E^\alpha_{R,\perp}(\tau)\}$ represents an orthonormal basis for the environment state. Its explicit expression will be extremely complicated, but we do not need to specify it. The normalization of the state $|E_L(\tau)\rangle$ implies that

$$|c(\tau)|^2 + \sum_\alpha |b_\alpha(\tau)|^2 = 1. \tag{28}$$

Then, the state of the system at the instant $\tau$ is given by the superposition

$$|\Psi(\tau)\rangle = \frac{1}{2} |1 + c(\tau)\rangle |A\rangle |E_R(\tau)\rangle + \frac{1}{2} |1 - c(\tau)\rangle |B\rangle |E_R(\tau)\rangle$$

$$+ \frac{1}{2} |A\rangle \sum_\alpha b_\alpha(\tau)|E^\alpha_{R,\perp}(\tau)\rangle \tag{29}$$

At the time $\tau$ the probability of the particle $P$ hitting $A$ is given by

$$p[A, \tau] = \frac{1}{4} |1 + c(\tau)|^2 + \frac{1}{4} \sum_\alpha |b_\alpha(\tau)|^2$$

$$= \frac{1}{4} |1 + c(\tau)|^2 + \frac{1}{4} |1 - c(\tau)|^2. \tag{30}$$

where in the last step we have used Eq. [28]. A simple calculation leads to

$$p[A, \tau] = \frac{1}{2} + c(\tau) = \frac{1}{2} + \frac{1}{2} e^{-\lambda \tau}. \tag{31}$$

A similar calculation leads to the probability of the particle $P$ hitting $S$ in $B$ as

$$p[B, \tau] = \frac{1}{2} - c(\tau) = \frac{1}{2} - \frac{1}{2} e^{-\lambda \tau}. \tag{32}$$

We see that ‘a bit’ of interference is left (no matter how large the time interval $\tau$ is):

$$p[A, \tau] - p[B, \tau] = e^{-\lambda \tau}, \tag{33}$$

showing that there is always an (eventually very slightly) enhanced probability to see the particle in $A$ rather than in $B$.

Notice that the very same result is found by using the reduced statistical operator

$$\hat{\rho}_{red}(\tau) = \langle E_R(\tau)|\hat{v}(\tau)|E_R(\tau)\rangle$$

$$+ \sum_\alpha \langle E^\alpha_{R,\perp}(\tau)|\hat{v}(\tau)|E^\alpha_{R,\perp}(\tau)\rangle = \left( \begin{array}{c|c} |A\rangle & |B\rangle \end{array} \right) \left( \begin{array}{c} p[A, \tau] \\ c(\tau) \end{array} \right) \left( \begin{array}{c} \langle A \rangle \\ \langle B \rangle \end{array} \right) \tag{34}$$

where $\hat{v}(\tau) = |\Psi(\tau)\rangle \langle \Psi(\tau)|$. The diagonal elements are the usual Born probabilities, while the non-diagonal elements quantify the overlap of the two branches and become very small for increasing time. (A related subject to the quantum evolution described here is that of the weak measurement, in which the ‘measurement’ is performed by a weak interaction and thus a unitary evolution of the whole system is taken into account, see the recent review [37] and references therein.)

All these considerations do not require any collapse of the wavefunction due to the environment (see also Ref. [38]). Indeed, if we replace the environment with the detector $D$ of § 2.2 (which was nothing else than a particular environment), the whole discussion is still valid (but see the comments on time scale after Eq. [26]). The only point when the Born rule enters is when we see the particle being either in $A$ or in $B$, but as we commented previously in this no-collapse many worlds scenario, we do not know why the Born rule applies [26,27]. In this sense, decoherence alone is not a solution of the measurement problem [39]. The wavefunction is still a superposition of different and distinguishable macroscopic states. Still, because of decoherence, these states (branches) become almost orthogonal, thus decoherence is an important element of the many worlds interpretation although it does
not explain the emergence of probabilities. What do theories with the collapse of the wavefunction predict? As long as few particles of the environment are involved (i.e., at small times), for sure we do not have any collapse and the entanglement in Eq. 25 is the correct description of the system. Namely, we know that interference effects occur for systems which contains about 1000 (and even more) particles [40]. But, if we wait long enough we can reach a critical number of particles at which the collapse takes place. Thus, simplifying the discussion as much as possible, according to collapse models there should be a critical time-interval \( \tau^* \) at which the probability \( p[A, \tau] \) suddenly jumps to 1/2:

\[
p[A, \tau] = \begin{cases} 
\frac{1}{2} + \frac{1}{2} e^{-i\lambda} & \text{for } \tau < \tau^*; \\
\frac{1}{2} & \text{for } \tau \geq \tau^*. 
\end{cases}
\] (35)

(In the presented example we vary the time of flight \( \tau \) by keeping all the rest unchanged, but the crucial point is the number of particles involved. Alternatively, one could change the density of the particles of the environment, which induces a change of the parameter \( \lambda \). In that case, one would have a critical \( \lambda_c \).) Indeed, such a sudden jump is an oversimplification, but is enough for our purposes: it shows that a new phenomenon, the collapse, takes place.

In Fig. 4 we show schematically the difference between the ‘no-collapse’ and the ‘collapse’ cases. Obviously, if \( \tau^* \) is very large, it becomes experimentally very difficult to distinguish the two curves, but the qualitative difference between them is clear.

In Ref. [41] the gradual appearance of decoherence due to interaction of electrons with image charges has been experimentally observed. This is analogous to our Eq. 31 (For other decoherence experiment see Ref. [32] and references therein.) Indeed, it would be very interesting to study decoherence in a setup with only two outcomes, for instance with the help of a Mach–Zehnder interferometer or by using neutron interferometers. Namely, even if the distinction between collapse/non-collapse is not yet reachable [9], a clear demonstration of decoherence and the experimental verification of Eq. 31 would be useful on its own.

As a last step, we show that the behavior \( p[A, t] = 1/2 \) for all \( t \geq \tau^* \) is a peculiarity of the collapse approach which is impossible if only a unitary evolution is taken into account. The proof makes use of the Hamiltonian \( H \) of the whole system (particle+slits+environment), for which we assume that \( \langle R | H | L \rangle = \langle L | H | R \rangle = 0 \), i.e. the full Hamiltonian does not mix the states \( |L\rangle \) and \( |R\rangle \). (This is indeed a quite general assumption for the type of problems that we study: once the particle has gone through the left slit, its wavefunction is \( |L\rangle \) and stays such (and vice versa for \( |R\rangle \)). Similarly, in the example of a (photon or neutron) Mach–Zehnder interferometer, after the first beam-splitter the path is either the lower or the upper and the whole Hamiltonian does not mix them.) It then follows that:

\[
|\Psi(t)\rangle = e^{-iHt} \frac{1}{\sqrt{2}} (|L\rangle |E_0\rangle + |R\rangle |E_0\rangle) = \frac{1}{\sqrt{2}} (|L\rangle e^{-iH_L t} |E_0\rangle + |R\rangle e^{-iH_R t} |E_0\rangle) \] (36)

where we have expressed \( |E_L(t)\rangle = e^{-iH_L t} |E_0\rangle \) and \( |E_R(t)\rangle = e^{-iH_R t} |E_0\rangle \) by introducing the Hamiltonians \( H_L = (L | H | L) \) and \( H_R = (R | H | R) \) which act in the subspace of the environment. (These expressions hold because \( H^n |L\rangle |E_0\rangle = |L\rangle H^n_L |E_0\rangle \) for each \( n \).) The overlap \( c(t) \) defined in Eq. 26 can be formally expressed as

\[
c(t) = \langle E_L(t) | E_R(t) \rangle = \langle E_0 | e^{-i(H_R - H_L) t} | E_0 \rangle. \] (37)
4 Entanglement with a non-orthogonal idler state

As a last example, we design an ideal setup in which the environment is represented again by a single particle, the idler state (see § 2.3.2). However, we assume now that a time-evolution of the idler state takes place

\[ |\Psi(t)\rangle = \frac{1}{\sqrt{2}} (|L\rangle |E_L(t)\rangle + |R\rangle |E_R(t)\rangle), \]  

(38)

with the ‘environment’ states now expressed in terms of the orthonormal idler-basis \(|I_1\rangle, |I_2\rangle\).

\[ |E_L(t)\rangle = |I_1\rangle, \]  

(39)

\[ |E_R(t)\rangle = \cos(\omega t) |I_1\rangle + \sin(\omega t) |I_2\rangle. \]  

(40)

Thus, while \(|E_L(t)\rangle = |I_1\rangle\) is a constant over time, we assume that \(|E_R(t)\rangle\) rotates in the space spanned by \(|I_1\rangle\) and \(|I_2\rangle\). Then, we can rewrite \(|\Psi(t)\rangle\) as

\[ |\Psi(t)\rangle = \frac{1}{2} |A\rangle [(1 + \cos(\omega t)) |I_1\rangle + \sin(\omega t) |I_2\rangle] + \frac{1}{2} |B\rangle [(-1 + \cos(\omega t)) |I_1\rangle + \sin(\omega t) |I_2\rangle]. \]  

(41)

The probability \(p[A, \tau]\) is given by

\[ p[A, \tau] = \frac{1}{2} + \frac{1}{2} \cos(\omega \tau) \]  

(42)

where \(\tau\) is the time at which the particle \(P\) hits the screen.

In conclusion, in a real implementation of this simple idea, it would be interesting to see the appearance and the disappearance of interference (with both fringes and antifringes) as function of the time of flight \(\tau\), see Fig. 5

It should be however stressed that the full interaction Hamiltonian does not act on the idler state alone. Indeed, the corresponding Hamiltonian has the form

\[ H = \alpha (|I_1\rangle \langle I_1| |R\rangle \langle R| + \text{h.c.}). \]  

(43)

This is indeed a quite peculiar type of interaction because if the idler state rotates only if the particle \(P\) is in the state \(|R\rangle\) (in the language of § 4, it means: \(H_L = 0, H_R = \alpha (|I_1\rangle \langle I_1| |I_2\rangle + \text{h.c.})\)). This implies that the spatial trajectory of both states \(|I_1\rangle\) and \(|I_2\rangle\) must be the same, otherwise the overlap \(|E_L(t)|E_R(t)\rangle\) would be an extremely small number and the effect that we have described would not take place.

5 Conclusions

We have presented an ideal interference experiment in which we have compared the unitary evolution and the existence of a collapse of the wavefunction. We have analyzed the case in which a detector measures the which-way information and we have shown that the collapse postulate as well as the no-collapse unitary evolution lead to the same outcome: the disappearance of interference on the screen. In the unitary (no-collapse) evolution, this is true only if the states of the detector are orthogonal. This is surely a very good, but not exact, approximation.

It was then possible to describe within the very same Gedankenexperiment two astonishing quantum phenomena: the Elitzur–Vaidman bomb and the delayed-choice experiment.

We have then turned to a description of the entanglement with the environment. The phenomenon of decoherence ensures that the interference smoothly disappears. However, as long as the quantum evolution is unitary, it never disappears completely. Conversely, the real collapse of the wave function introduces a new kind of dynamics which is not part of the linear Schrödinger equation. While the details differ according to which model is chosen [9], the main features are similar: a quantum state in which one has a delocalized object (superposition of ‘here’ and ‘there’) is not a stable configuration, but is metastable and decays to a definite position (either ‘here’ or ‘there’). In conclusion, the collapse and the no-collapse views are intrinsically different, as Fig. 4 shows. At a fundamental level, the unitary (no-collapse) evolution leads quite naturally to the many worlds interpretation in which also detectors and observers are included in a superposition. (For a different view see the Bohm interpretation in which an equation describing the trajectories of the particles is added [42].) The positions are the hidden variables of this approach. The Born rule is put in from the very beginning. An extension of the Bohm interpretation to the relativistic framework and to quantum field theories is a difficult task, see Ref. [44] for a critical analysis.)
Even if the distinction between the collapse and the no-collapse alternatives is probably still too difficult to be detected at the moment, the demonstration of decoherence in an experiment with two final states would be an interesting outcome on its own (see the dashed curve in Fig. 4). Also a situation in which an entangled particle is emitted in such a way that an ‘oscillating interference’ takes place (see Fig. 5) might be an interesting possibility.

A further promising line of research to test the existence of the collapse of the wavefunction is the theoretical and experimental study of unstable quantum systems. The non-exponential behavior of the survival probability for short times renders the so-called Zeno and Anti-Zeno effects possible [45–54]: these are modifications of the survival probability due to the effect of the measurement, which have been experimentally observed [55, 56]. The measurement of an unstable system (for instance, the detection of the decay products) can be modelled as a series of ideal measurements in which the collapse of the wavefunction occurs, but can also be modelled through a unitary evolution in which the wave function of the detector is taken into account and no collapse takes place [57–60]. Then, if differences between these types of measurement appear, one can test how a detector is performing a certain measurement [61]. Quite remarkably, such effects are not restricted to nonrelativistic quantum mechanics, but hold practically unchanged also in the context of relativistic quantum field theory [62–65] and are therefore applicable in the realm of elementary particles.

In conclusion, quantum mechanics still awaits for better understanding in the future. It is surely of primary importance to test the validity of (unitary) standard quantum mechanics for larger and heavier bodies. In this way the new collapse dynamics, if existent, may be discovered.

**Acknowledgments**

These reflections arise from a series of seminars on “Interpretation and New Developments of Quantum Mechanics” and lectures “Decays in Quantum Mechanics and Quantum Field Theory”, which took place in Frankfurt over the last 4 years.

The author thanks Francesca Sauli, Stefano Lottini, Giuseppe Pagliara, and Giorgio Torrieri for useful discussions. Stefano Lottini is also acknowledged for a careful reading of the manuscript and for help in the preparation of the figures.

**References**


