Wave-particle duality - The Mach-Zehnder interferometer

M. Betz, I. de Lima and G. Mussatto

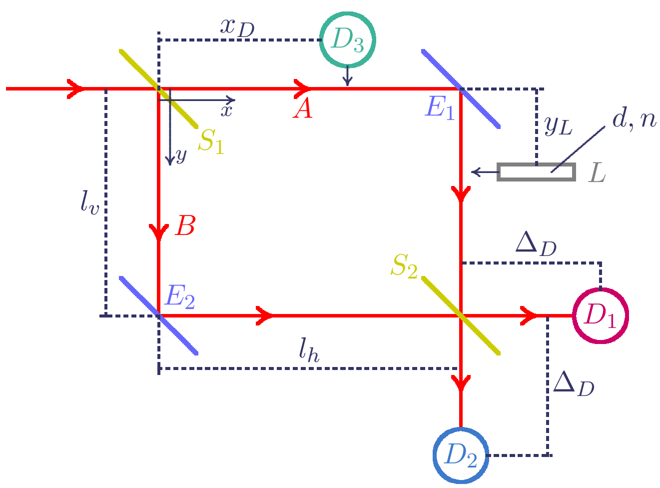
Formalism

In this part of the material, the mathematical description of the phenomena observed with the interferometer is presented in some details. The underlying theory, known as wave mechanics or quantum mechanics, makes use of a rather abstract formalism. However, a reader having some familiarity with linear algebra, complex numbers, and elementary (exponential and trigonometric) functions should be able to comprehend the development. A reader not fluent in these subjects, or not interested in diving in somewhat technical details, should be able anyway to grasp the essential conceptual aspects, merely by running the animations and reading the short texts that comment them.

I. Specification of the experimental device

The conventions indicated in the figure below will be used to refer to the various parts of the interferometer. The latter consists of two semireflective mirrors S_1 and S_2 and of two totally reflective mirrors E_1 and E_2. At the apparatus exit are placed two detectors D_1 and D_2. Optionally, a third detector D_3 and/or a transparent blade L[1] may be inserted in one of the beam's paths.

The coordinate system has its origin at the center of the beam splitter S_1, with the vertical[2] axis pointing downward.



The following notations shall also be employed:

- l_h for the length of the horizontal arms of the apparatus or, in other words, for the distances between S_1 and E_1 and between E_2 and S_2. These distances must be equal.
- l_v for the length of the vertical arms of the apparatus or, in other words, for the distances between S_1 and E_2 and between E_1 and S_2. These distances must also be equal, but l_v does not need to be equal to l_h.

[1] This text was originally written in Portuguese, in which *lâmina* is the word for blade. A reader of this English translation may perhaps think of the word *leaf* to recall the motivation for the notation L.

[2] Here the terms “horizontal” and “vertical” are used in reference to the figure. In the laboratory, the four arms are usually in the same horizontal plane.

- d for the thickness of the transparent blade, and n for the refraction index of the material of which it is made.
- y_L for the vertical position of the transparent blade.
- Δ_D for the distance between the beam splitter S_2 and the detectors D_1 and D_2. It is assumed that this distance is the same for both detectors.
- x_D for the distance between the beam splitter S_1 and the detector D_3, in case the latter is in use.

It shall be assumed that the semireflective mirrors are symmetric and do not absorb light. In this case, it can be shown[3] that the reflected component of light undergoes a phase shift of $\pi/2$ or $\lambda/4$, where λ is the wave length. In the case of a totally reflective mirror, the phase shift depends on the type of mirror. However, since in the Mach-Zehnder interferometer each beam component experiences necessarily one (and only one) such reflection, it suffices that both mirrors be of the same type for the phase shift they bring about to be in fact irrelevant. For the sake of simplicity, it will be assumed that each totally reflective mirror also introduces a phase shift equal to $\pi/2$.

II. Continuous flux - Plane waves

A. Specification and interpretation of the wave

The propagation through the interferometer of a plane wave, that is, a wave of well-defined frequency f and wave length λ shall be studied first. In quantum mechanics, such a wave corresponds to a "stationary state", an uninterrupted flux of particles that are not observed individually. From the perspective of classical wave theory, this wave describes a radiation flux of constant intensity.

In quantum mechanics, the wave function is a complex function. For a well-defined frequency f, it possesses the form[4]

$$\Psi(x,y,t) = \psi(x,y)\, e^{-i\omega t} \ , \tag{1}$$

where

$$\omega = 2\pi f \ . \tag{2}$$

For a plane wave propagating in the x direction, as is the case before the beam enters the apparatus, one has

$$\psi(x,y) \equiv \psi_0 = e^{ikx} \ , \tag{3}$$

with

$$k = \frac{2\pi}{\lambda} \ . \tag{4}$$

Quantum theory interprets the squared norm $|\Psi|^2$ of the wave function as a measure of the intensity of the beam's particle flux. With (**??**) and (**??**), the initial intensity is[5]

$$I_0 = |\psi_0|^2 = 1 \ . \tag{5}$$

[3] See the accompanying material entitled *Additional Informations*.

[4] In the interferometer, propagation occurs in the x or in the y direction, depending on the path segment considered. Therefore, a bidimensional description is required. The third dimension, perpendicular to the apparatus plane, may be ignored.

[5] It follows that intensities will be obtained as fractions (or percentages, after multiplication by 100) of the initial intensity. One should note at this point that the classical description, in which the wave function is real, can be obtained by taking the real part (or, equivalently, the imaginary part) of all expressions presented here. Such an artifice is in fact frequently used for convenience reasons. The deduced intensities will then be interpreted as fractions of the initial energy flux (averaged over one period).

B. Equal propagation through both arms

After encountering the semireflective mirror S_1, the wave is divided in two components ψ_A and ψ_B, of equal intensity $(1/2)$. The modulus of each component is therefore equal to $1/\sqrt{2}$.

The component ψ_A was transmitted by S_1, which did not occasion any phase shift. Therefore, between S_1 and E_1,

$$\psi_A(S_1 \to E_1) = \frac{1}{\sqrt{2}}\, e^{ikx} \,. \tag{6}$$

This component is then reflected by E_1, which brings about a phase shift of $\pi/2$, besides changing by 90° the propagation direction. Hence, after this reflection

$$\psi_A(E_1 \to S_2) = \frac{1}{\sqrt{2}}\, e^{i(kl_h + \frac{\pi}{2} + ky)} \,, \tag{7}$$

were l_h refers to the horizontal distance traveled by the beam along path A (obviously equal to the distance between S_1 and E_1). When it reaches S_2, ψ_A is divided into a transmitted component $\psi_A(D_2)$ which proceeds toward detector D_2, and a reflected component $\psi_A(D_1)$ which heads for detector D_1. This component undergoes a second shift by $\pi/2$ upon reflection. Each one of these components has modulus equal to that of ψ_A divided by $\sqrt{2}$, hence equal to $1/2$. Thus one has

$$\begin{aligned} \psi_A(D_1) &= \frac{1}{2}\, e^{i(kl_h + \frac{\pi}{2} + kl_v + \frac{\pi}{2} + kx')} = \frac{1}{2}\, e^{i(kl_v + \pi + kx)} \,, && (8) \\ \psi_A(D_2) &= \frac{1}{2}\, e^{i(kl_h + \frac{\pi}{2} + kl_v + ky')} = \frac{1}{2}\, e^{i(kl_h + \frac{\pi}{2} + ky)} \,, && (9) \end{aligned}$$

where a coordinate system (x', y') whose origin coincides with S_2 has been used temporarily, and where l_v refers to the vertical distance traveled by the beam along path A (obviously equal to the distance between E_1 and S_2).

The ψ_B component was reflected by S_1, which resulted in a phase shift of $\pi/2$, besides rotating by 90° the direction of propagation. Therefore, between S_1 and E_2,

$$\psi_B(S_1 \to E_2) = \frac{1}{\sqrt{2}}\, e^{i(ky + \frac{\pi}{2})} \,. \tag{10}$$

This component is then reflected by E_2, which brings about another phase shift by $\pi/2$, and also rotates again by 90° the direction of propagation. Therefore, after this reflection

$$\psi_B(E_2 \to S_2) = \frac{1}{\sqrt{2}}\, e^{i(kl_v + \frac{\pi}{2} + \frac{\pi}{2} + kx)} = \frac{1}{\sqrt{2}}\, e^{i(kl_v + \pi + kx)} \,, \tag{11}$$

where use has been made of the fact that the vertical distance traveled by the beam on path B (obviously equal to the distance between S_1 and E_2) is also equal to l_v. When it reaches S_2, ψ_B is divided into a transmitted component $\psi_B(D_1)$ which proceeds to detector D_1, and a reflected component $\psi_B(D_2)$ which heads for detector D_2. This component suffers a third shift by $\pi/2$ as it is reflected. Each one of these components has modulus equal to the modulus of ψ_B divided by $\sqrt{2}$, hence equal to $1/2$. Thus, one has[6]

$$\begin{aligned} \psi_B(D_1) &= \frac{1}{2}\, e^{i(kl_v + \pi + kl_h + kx')} = \frac{1}{2}\, e^{i(kl_v + \pi + kx)} \,, && (12) \\ \psi_B(D_2) &= \frac{1}{2}\, e^{i(kl_v + \pi + kl_h + \frac{\pi}{2} + ky')} = \frac{1}{2}\, e^{i(kl_h + \frac{3\pi}{2} + ky)} \,. && (13) \end{aligned}$$

[6]Here is used the fact that the distance between E_2 and S_2 is l_h.

Combining (**??**) and (**??**), one obtains the resultant wave that reaches D_1:

$$\psi(D_1) = \psi_A(D_1) + \psi_B(D_1) = \frac{1}{2}\, e^{i(kl_v+\pi+kx)} + \frac{1}{2}\, e^{i(kl_v+\pi+kx)} = e^{i(kl_v+\pi+kx)} \; . \tag{14}$$

Similarly, combining (**??**) and (**??**), the resultant wave reaching D_2 is:

$$\begin{aligned} \psi(D_2) = \psi_A(D_2) + \psi_B(D_2) &= \frac{1}{2}\, e^{i(kl_h+\frac{\pi}{2}+ky)} + \frac{1}{2}\, e^{i(kl_h+\frac{3\pi}{2}+ky)} \\ &= \frac{1}{2}\, e^{i(kl_h+\frac{\pi}{2}+ky)}(1+e^{i\pi}) = 0 \; , \end{aligned} \tag{15}$$

where the well-known formula $e^{i\pi} = -1$ has been used. Thus, it turns out that the wave reaching D_1 has amplitude (and therefore intensity) equal to that of the incident wave, whereas nothing is recorded by D_2. Actually, this result is quite easy to understand. The waves that reach D_1 following paths A and B both undergo one reflection by a semireflecting mirror. Therefore, they arrive with the same phase and interfere constructively. As for the waves reaching D_2, the one that follows path A is not subjected to any reflection by a semireflecting mirror. In contrast that which travels along path B is reflected successively by both semireflecting mirrors and therefore its phase is shifted by half a wave length with respect to the first one. The interference is completely destructive. As has already been made clear, the totally reflecting mirrors may be ignored in this argument.

C. Transparent blade in one of the arms

Consider now what happens when a transparent blade L of thickness d, made of a material of refraction index n, is inserted on path A, for example on the way between E_1 and S_2.

It is well known that in a material medium of refraction index n, the speed of light is c/n, being c the velocity of light in vacuum.[7] Therefore, since the wave frequency is not affected by the material, the latter induces a modification of the wave length λ resulting in a value λ' given by

$$\lambda' = \frac{c}{nf} = \frac{\lambda}{n} \; . \tag{16}$$

Correspondingly, the wave number k becomes

$$k' = \frac{2\pi}{\lambda'} = nk \; . \tag{17}$$

It follows that the presence of the material occasions a phase shift in the component ψ_A, given by

$$\phi = (k'-k)d = (n-1)kd \; . \tag{18}$$

Since a phase shift of 2π corresponds to one wave length, this result may be interpreted as a spatial shift of the wave fronts by

$$\varphi = (n-1)d \; . \tag{19}$$

Thus, with the introduction of the blade, expression (**??**) gets modified in the following manner:

$$\psi_A(E_1 \to S_2) = \begin{cases} \frac{1}{\sqrt{2}}\, e^{i(kl_h+\frac{\pi}{2}+ky)} & \text{for} \quad y < y_L \\ \frac{1}{\sqrt{2}}\, e^{i(kl_h+\frac{\pi}{2}+\phi+ky)} & \text{for} \quad y > y_L \; , \end{cases} \tag{20}$$

[7] The velocity of light in air is practically equal to c.

where y_L is the blade's vertical position. The corresponding modifications of expressions (**??**) and (**??**) are

$$\begin{aligned}\psi_A(D_1) &= \frac{1}{2}\,e^{i(kl_v+\pi+\phi+kx)}\ , && (21)\\ \psi_A(D_2) &= \frac{1}{2}\,e^{i(kl_h+\frac{\pi}{2}+\phi+ky)}\ . && (22)\end{aligned}$$

From this, the modifications of expressions (**??**) and (**??**) are as follows:

$$\begin{aligned}\psi(D_1)=\psi_A(D_1)+\psi_B(D_1) &= \frac{1}{2}\,e^{i(kl_v+\pi+\phi+kx)}+\frac{1}{2}\,e^{i(kl_v+\pi+kx)}\\ &= e^{i(kl_v+\pi+\frac{\phi}{2}+kx)}\cos\frac{\phi}{2}\ , && (23)\end{aligned}$$

$$\begin{aligned}\psi(D_2)=\psi_A(D_2)+\psi_B(D_2) &= \frac{1}{2}\,e^{i(kl_h+\frac{\pi}{2}+\phi+ky)}+\frac{1}{2}\,e^{i(kl_h+\frac{3\pi}{2}+ky)}\\ &= e^{i(kl_h+\frac{\pi}{2}+\frac{\phi}{2}+ky)}\,i\,\mathrm{sen}\frac{\phi}{2}\ . && (24)\end{aligned}$$

Altogether, for a unit intensity of the incident beam, the intensities observed in the two detectors are

$$\begin{aligned}I(D_1) &= \cos^2\frac{\phi}{2}\ , && (25)\\ I(D_2) &= \mathrm{sen}^2\frac{\phi}{2}\ . && (26)\end{aligned}$$

By modifying the thickness d of the blade and/or the refraction index n of its material, it is possible to modify at will these relative intensities. In the visualization, the blade's thickness may be varied between 0 and 2λ, and the refraction index between 1 and 2. These intervals are sufficient for it to be possible to obtain any given percentages for the relative intensities measured.

III. Monophotonic states - Wave packets

A. Single-photon wave packet

As is well known, although it is formulated as a wave theory, quantum mechanics needs to be interpreted in terms of corpuscles, since the individual events that are observed - clicks in detectors or impacts on a screen - are localized in space and time. The relation between the wave formalism and the experimental facts is statistical. Specifically, in the case of the interferometer, the corpuscles are photons. It is possible to arrange the apparatus in such a manner as to observe the photon impacts on a screen. The undulatory aspect manifests itself in the interference pattern formed be accumulating a large number of impacts. Another procedure, which will be employed here, consists simply in using photon counters as detectors. As will be seen below, the undulatory aspect then manifests itself in the numbers of photons counted.

In order to analyze in details the propagation and detection of photons, one needs to describe mathematically the state of a single photon. For this purpose, a *wave packet*[8] must be constructed, that is, a superposition of plane waves, with different values of the wave number k. Mathematics

[8]The association of a wave packet to a photon is not without limitations. See the accompanying material entitled *Additional Informations* for references.

demonstrates that if one wishes to form a wave packet localized in a region of size Δx, one needs to superpose plane waves whose wave numbers differ at least by values of order[9] Δk, such that

$$\Delta k \simeq \frac{1}{\Delta x} \; . \tag{27}$$

Quantum mechanics stipulates relationships between the physical quantities associated to the wave and those associated to the particle. In particular, it postulates the following relation between the momentum p of the particle and the wave length λ of the wave:

$$p = \frac{h}{\lambda} \; , \tag{28}$$

where h is the famous *Planck constant*, a new fundamental constant that is ubiquitous in quantum physics. Using (**??**), this relation may be rewritten in the form

$$p = \hbar k \; , \tag{29}$$

with $\hbar = h/2\pi$. Thus, the mathematical condition (**??**) leads to the physical condition

$$\Delta p \, \Delta x \simeq \hbar \; , \tag{30}$$

where Δp is the dispersion or indeterminacy in the value of the momentum. It is worth emphasizing that minimal values of the dispersions are concerned here. This condition is known as *Heisenberg's indeterminacy (or uncertainty) relation*.

In order for the wave packet to describe a photon of approximately defined momentum, one must have

$$\Delta p << \overline{p} \;\; \text{or} \;\; \Delta k << \overline{k} \;\; , \tag{31}$$

where $\overline{p}$ is the average momentum and $\overline{k}$ the average wave number. From relation (**??**), it is easy to deduce

$$\frac{\Delta\lambda}{\overline{\lambda}} = \frac{\Delta k}{\overline{k}} \; , \tag{32}$$

where $\overline{\lambda}$ is the average wave length. From (**??**) and (**??**), it follows that

$$\Delta x >> \frac{1}{\overline{k}} \; , \tag{33}$$

or equivalently, using (**??**) again,

$$\Delta x >> \overline{\lambda} \; . \tag{34}$$

In words, the size of the wave packet must be much bigger than the average wave length.

The conclusion that has just been reached has some implications for the visualization, on the screen, of a monophotonic state. Associating to the photon a packet which is visible but not very large, the wave length ends up being too small to be shown on the screen. If, as was done above, one introduces on the beam path a blade whose maximum thickness is two wave length, it becomes also impossible to discern this thickness at the screen's scale. For this reason, in the animation, a lens is provided to the user when he needs to see the blade in order to set its thickness.

With supposition (**??**), the phase shift introduced by the blade is approximately the same for all waves that compose the packet and is given, in terms of the mean wave number, by an expression analog to (**??**):

$$\phi = (n-1)\, \overline{k} \, d \; . \tag{35}$$

[9]Inequalities denoted by the symbol $\simeq$, similarly to those denoted by $<<$ and $>>$, express relationships between *orders of magnitude*. Factors of "order one", such as 2 or π, are ignored.

The spacial shift of all wave fronts, and therefore of the packet, is given by (**??**). Being of the same order of magnitude as the blade's thickness, it is equally imperceptible at the scale of the screen.

In order to describe the photon's propagation in the plane of the interferometer, it is necessary to construct a bidimensional packet, whose general form is

$$\Psi(x,y,t) = \int \frac{dk_1\, dk_2}{2\pi}\, g(k_1,k_2)\, e^{i[k_1 x + k_2 y - \omega(k_1,k_2)t]}\ , \tag{36}$$

where

$$\omega(k_1,k_2) = \sqrt{k_1^2 + k_2^2}\, c\ . \tag{37}$$

Relation (**??**) ensures that the packet propagates with velocity c. The motion desired for the photon determines the characteristics of the function $g(k_1,k_2)$. For example, in order for the packet to describe a photon propagating in the direction of the x axis with average momentum $\overline{p} = \hbar\overline{k}$, $g(k_1,k_2)$ must be a function centered at $k_1 = \overline{k}$ and $k_2 = 0$, with dispersions $\Delta k_1 << \overline{k}$ and $\Delta k_2 << \overline{k}$. Writing $k_1 = \overline{k} + k_1'$, one may then approximate (**??**) by

$$\omega(k_1,k_2) = \sqrt{(\overline{k}+k_1')^2 + k_2^2}\, c \simeq \overline{k}\sqrt{1 + 2\frac{k_1'}{\overline{k}}}\, c \simeq \overline{k}(1 + \frac{k_1'}{\overline{k}})\, c = (\overline{k} + k_1')\, c\ , \tag{38}$$

where terms of second order in $\Delta k_1/\overline{k}$ and $\Delta k_2/\overline{k}$ have been neglected. Thus, expression (**??**) takes the form

$$\Psi(x,y,t) = e^{i\overline{k}(x-ct)} \int \frac{dk_1'\, dk_2}{2\pi}\, \overline{g}(k_1',k_2)\, e^{i[(k_1'(x-ct)+k_2 y]} = e^{i\overline{k}(x-ct)}\Phi(x-ct,y)\ , \tag{39}$$

where the notations

$$\overline{g}(k_1',k_2) = g(\overline{k}+k_1',k_2) \tag{40}$$

and

$$\Phi(x,y) = \int \frac{dk_1'\, dk_2}{2\pi}\, \overline{g}(k_1',k_2)\, e^{i(k_1' x + k_2 y)} \tag{41}$$

have been introduced. Since function $\overline{g}(k_1',k_2)$ is centered at $k_1' = k_2 = 0$, function Φ is centered at $x = y = 0$. In order to allow the probabilistic interpretation of the theory, the latter function should be normalized:

$$\int |\Phi(x,y)|^2\, dx\, dy = 1\ . \tag{42}$$

The wave function (**??**) corresponds to a packet whose center propagates with positive velocity c in the xS direction along the line $y = 0$, passing through $x = 0$ at $t = 0$. For the sake of simplicity in further developments, the compact notation[10]

$$|ct, 0\rangle \equiv e^{i\overline{k}(x-ct)}\Phi(x-ct,y) \tag{43}$$

will be used for this packet. It is easy to generalize this to the case of a packet whose center propagates with positive velocity c along the line $y = y_0$ and goes through $x = x_0$ at $t = 0$:

$$|ct + x_0, y_0\rangle \equiv e^{i\overline{k}(x-x_0-ct)}\Phi(x-x_0-ct, y-y_0)\ . \tag{44}$$

Analogously, a packet whose center propagates with positive velocity c in the y direction along the line $x = x_0$ and passes by $y = y_0$ at $t = 0$ may be written

$$|x_0, ct + y_0\rangle \equiv e^{i\overline{k}(y-y_0-ct)}\Phi(x-x_0, y-y_0-ct)\ . \tag{45}$$

[10] This notation is inspired by that introduced by P. A. M. Dirac to systematize the formalism of quantum mechanics. However, it is employed here in an informal manner only.

B. Inclusion of the detectors

In order to be able to discuss the measurement process in quantum mechanics, it is necessary to include the detectors, which are photon counters, in the description. Denoting by χ_1 and χ_2 the states of counters D_1 and D_2, one may assume that these are in the state $\chi(0)$ before a photon arrives and that the passage of the photon through a counter induces the transition $\chi(0) \rightarrow \chi(1)$ in the state of the latter. It will be assumed that, after registering this event and therefore increasing by one unit the number of counts it indicates, a counter returns to the state $\chi(0)$, waiting for the next photon.

C. Propagation of the packet through the apparatus

Using arguments identical to those developed for a plane wave, one may now write the expression of the wave function associated to the photon and the counters in the various parts of the interferometer. In order to be able to specify the time ordering of the reflections by the mirrors, it will be assumed that $l_h > l_v$. For convenience, the relation

$$e^{i\pi/2} = i \; , \tag{46}$$

will be invoked to write the additional phase introduced by each reflection. Since the system of two detectors remains in the state $\chi_1(0)\,\chi_2(0)$ until the last step in the process (the counting of the photon), it will be omitted for simplicity in the description of the steps anterior to the last one.

Suppose that the wave packet reaches the beam splitter S_1 in $t = 0$. Before that, the photon's state is

$$\Psi = |ct, 0\rangle \;\; \text{for} \;\; t < 0 \; . \tag{47}$$

After passing through the splitter S_1, and before reaching the mirror E_2, the quantum state is

$$\Psi = \frac{1}{\sqrt{2}}\,|ct, 0\rangle + i\frac{1}{\sqrt{2}}\,|0, ct\rangle \;\; \text{for} \;\; 0 < ct < l_v \; . \tag{48}$$

After reflection by mirror E_2, but before reflection by mirror E_1, the state is (using $i^2 = -1$)

$$\Psi = \frac{1}{\sqrt{2}}\,|ct, 0\rangle - \frac{1}{\sqrt{2}}\,|ct - l_v, l_v\rangle \;\; \text{for} \;\; l_v < ct < l_h \; . \tag{49}$$

After reflection by mirror E_1, but before reaching the blade L, the state is

$$\Psi = i\frac{1}{\sqrt{2}}\,|l_h, ct - l_h\rangle - \frac{1}{\sqrt{2}}\,|ct - l_v, l_v\rangle \;\; \text{for} \;\; l_h < ct < l_h + y_L \; . \tag{50}$$

Within the assumptions already discussed, the blade L merely induces a phase shift ϕ in the component that goes through it. Hence, after the passage of the component through the blade, the state is

$$\Psi = i\frac{1}{\sqrt{2}}\,e^{i\phi}\,|l_h, ct - l_h\rangle - \frac{1}{\sqrt{2}}\,|ct - l_v, l_v\rangle \;\; \text{for} \;\; l_h + y_L < ct < l_h + l_v \; . \tag{51}$$

Since the next steps involve the detection process, the detectors will be included in the state's description from now on. After going through beam splitter S_2 but before detection, the total state of the system is

$$\begin{aligned} \Psi_{tot} \;\; = \;\; & \{i\frac{1}{\sqrt{2}}\,e^{i\phi}[\frac{1}{\sqrt{2}}|l_h, ct - l_h\rangle + i\frac{1}{\sqrt{2}}\,|ct - l_v, l_v\rangle] \\ & -\frac{1}{\sqrt{2}}[\frac{1}{\sqrt{2}}|ct - l_v, l_v\rangle + i\frac{1}{\sqrt{2}}|l_h, ct - l_h\rangle]\} \\ & \times\, \chi_1(0)\chi_2(0) \;\; \text{for} \;\; l_h + l_v < ct < l_h + l_v + \Delta_D \; , \end{aligned} \tag{52}$$

where Δ_D denotes the distance between the splitter S_2 and the detectors, which shall be assumed to be the same for both detectors. Rearranging terms, expression (**??**) may be simplified as:

$$\begin{aligned}\Psi_{tot} &= -e^{i\phi/2}\,[\mathrm{sen}\frac{\phi}{2}\,|l_h, ct-l_h\rangle + \cos\frac{\phi}{2}\,|ct-l_v, l_v\rangle] \\ &\quad \times \chi_1(0)\chi_2(0) \ \text{ for } \ l_h+l_v < ct < l_h+l_v+\Delta_D \ . \end{aligned} \tag{53}$$

The interaction of the photon with the detectors, which takes place at time $t = (l_h + l_v + \Delta_D)/c$, transforms this state into

$$\Psi_{tot} = -e^{i\phi/2}\,[\mathrm{sen}\frac{\phi}{2}\,|l_h, ct-l_h\rangle \times \ \chi_1(0)\chi_2(1) + \cos\frac{\phi}{2}\,|ct-l_v, l_v\rangle \times \ \chi_1(1)\chi_2(0)] \ . \tag{54}$$

D. Interpretation of the measurement and collapse of the wave packet

Up to this point, the evolution of the state of the system, including the counters, unfolded without intervention of an observer, who would witness the click of one of the detectors, or of a computer that would register the increase by one unit of the number of photons already accumulated by one of the counters. Such an intervention is necessary for it to be possible to talk about the result of the experiment. If counter D_1 registers one more photon, its state changes to $\chi_1(1)$. According to the usual interpretation of quantum mechanics, the probability for this to occur is given by the modulus squared of the coefficient of the corresponding term (the second one) of expression (**??**) above, that is,

$$P(D_1) = |-e^{i\phi/2}\cos\frac{\phi}{2}|^2 = \cos^2\frac{\phi}{2} \ . \tag{55}$$

Similarly, if counter D_2 registers the photon, it ends up in state $\chi_2(1)$. The associated probability is deduced from the coefficient of the first term of the complete state (**??**):

$$P(D_2) = |-e^{i\phi/2}\mathrm{sen}\frac{\phi}{2}|^2 = \mathrm{sen}^2\frac{\phi}{2} \ . \tag{56}$$

As is immediately verified, expressions (**??**) and (**??**) for the probabilities are identical to expressions (**??**)-(**??**) of intensities in the classical theory. The theoretical computation of these probabilities is what the formalism of quantum mechanics permits. It is not possible to predict the result of an individual measurement. In order to verify experimentally the theory, it is necessary to repeat the measurement many times. If a large number of photons is sent through the apparatus and counters D_1 and D_2 register, respectively, N_1 and N_2 photons, then the ratios $N_1/(N_1+N_2)$ and $N_2/(N_1+N_2)$ should approximate closely the theoretical values for probabilities $P(D_1)$ and $P(D_2)$.

After detection by counter D_1, the photon state will be given by the part of the second term of expression (**??**) that relates to the photon, that is, by the wave packet

$$\Psi_1 = |ct-l_v, l_v\rangle \ \text{ for } \ ct > l_h+l_v+\Delta_D \ . \tag{57}$$

On the other hand, if it was counter D_2 that registered the passage of the photon, the latter subsequent state is easily extracted from the first term of (**??**):

$$\Psi_2 = |l_h, ct-l_h\rangle \ \text{ for } \ ct > l_h+l_v+\Delta_D \ . \tag{58}$$

It should be noted that the coefficients (functions of the phase shift ϕ) present in (**??**) have been omitted in (**??**) and (**??**). The reason for this is that these states should be *normalized*, that is, the total probabilities associated to them should be equal to unity. If counter D_1 went off, the photon must necessarily been observed exiting from it (or being absorbed by it); idem for counter D_2. The global phases of the states are irrelevant and were chosen equal to 1 in (**??**) and (**??**). The reduction of the photon state described by (**??**), to state (**??**) or to state (**??**), depends on the measurement result; this is the (in)famous *collapse of the wave packet*. Although this collapse is made explicit in the visualizations, it should be recognized that its conceptualization is quite problematic. Should it be thought of as a real physical phenomenon, or merely as a step in a logical argument? This question is discussed further in the next section.

E. What path did the photon follow?

The development above was based on the linear superposition of the wave-packet components that describe the march of the photon along the two alternative paths A and B. One may ask what happens if one tries to find out which of the two paths a given photon really took. In order to answer this, a detector D_3 may be introduced on path A, at the place indicated in the figure. It will be assumed that $x_D < l_v$, but the reader will easily convince himself that the conclusions reached do not depend on the position of D_3, as long as it is somewhere on path A. It will be assumed that this detector is initially in state $\chi_3(0)$ and switches to state $\chi_3(1)$ when it registers the passage of the photon.

Until the wave packet reaches detector D_3, nothing new happens and one may use (**??**) to describe the state of the photon, merely inserting also the state of the detector:[11]

$$\Psi = [\frac{1}{\sqrt{2}}\,|ct,0\rangle + i\frac{1}{\sqrt{2}}\,|0,ct\rangle] \times \chi_3(0) \;\; \text{for} \;\; 0 < ct < x_D \; . \tag{59}$$

Only the first term of this expression corresponds to the propagation of the packet along path A, therefore only for this term will there be a change in the state of detector D_3. Hence, in $t = x_D/c$, the state above becomes

$$\Psi = \frac{1}{\sqrt{2}}\,|ct,0\rangle \times \chi_3(1) + i\frac{1}{\sqrt{2}}\,|0,ct\rangle \times \chi_3(0) \; . \tag{60}$$

According to the already presented interpretation of what happens when a counter actually finds out if a photon is there or not, one may deduce from the above expression that the probability of counter D_3 clicking is $1/2$ and that, if this occurs, the photon's state becomes

$$\Psi_A = |ct,0\rangle \; . \tag{61}$$

The probability of the counter *not* clicking is also $1/2$ and, in that case, the state of the photon "collapses" to[12]

$$\Psi_B = i|0,ct\rangle \; , \tag{62}$$

which corresponds to a packet propagating exclusively along path B.

From then on, since the photon is in state (**??**) or in state (**??**), but not in a superposition of these two states, one needs to trace separately the evolution of each one of the states through the remainder of the apparatus. The results for $(l_h + l_v)/c < t < (l_h + l_v + \Delta_D)/c$, that is before the detection by D_1 or D_2 can be easily extracted from expression (**??**):[13]

$$\begin{aligned} \Psi_A &= i\, e^{i\phi}[\frac{1}{\sqrt{2}}|l_h, ct - l_h\rangle + i\frac{1}{\sqrt{2}}\,|ct - l_v, l_v\rangle] \times \chi_1(0)\chi_2(0) \; , & (63) \\ \Psi_B &= -[\frac{1}{\sqrt{2}}|ct - l_v, l_v\rangle + i\frac{1}{\sqrt{2}}|l_h, ct - l_h\rangle] \times \chi_1(0)\chi_2(0) \; . & (64) \end{aligned}$$

[11] Detectors D_1 and D_2 do not need to be considered at this stage of the argument.

[12] The factor i in (**??**) merely introduces a global phase which is unobservable and may be dropped.

[13] The states of detectors D_1 and D_2 are now included, but the state of detector D_3, which already played its role, is omitted.

At $t = (l_h + l_v + \Delta_D)/c$, the detectors are reached and these states are changed into

$$\Psi_A = i\, e^{i\phi}[\frac{1}{\sqrt{2}}|l_h, ct - l_h\rangle \times \chi_1(0)\chi_2(1) + i\frac{1}{\sqrt{2}}\,|ct - l_v, l_v\rangle \times \chi_1(1)\chi_2(0)]\,, \tag{65}$$

$$\Psi_B = -[\frac{1}{\sqrt{2}}|ct - l_v, l_v\rangle \times \chi_1(1)\chi_2(0) + i\frac{1}{\sqrt{2}}|l_h, ct - l_h\rangle \times \chi_1(0)\chi_2(1)]\,. \tag{66}$$

It can be seen that the transparent blade introduced on path A only modifies the global phase of the packet, which does not lead to any observable effect. In both cases - path A or path B - the probability of a click is $1/2$ for D_1 as well as for D_2. In effect, the act of finding out which path the photon followed destroyed the interference between the two paths. As before, if D_1 goes off, the photon's state after the counting will be described by (**??**); in case D_2 goes off, the state will be given by (**??**).

It is interesting to contrast the effect, on the state of the photon, of the two devices that were introduced in one of the arms of the interferometer: the blade and the detector. The blade introduces a phase shift in the component of the packet that goes through it, without affecting the coherence between the two components. In contrast, the detector "wipes out" one of the components. Why such a difference? Because the blade is merely an additional element added to the system, and not a device that provides information about the latter. On the other hand, the detector registers a new information. The modification of the state of the system - the collapse - corresponds to the incorporation in the description of the said new information. It is particularly noteworthy that, even if nothing happens to the detector (no click or increase in the registered number), the mere fact of it being present permits to conclude that the photon did not pass through arm A and hence, passed through arm B, thus provoking the collapse of the packet in this arm.[14] In other words, the fact that a possible event did no happen somewhere affects the state of the system somewhere else. However, one should clearly avoid the temptation to attribute this process to a "spooky" action at a distance.

It is worth emphasizing again that the theoretical analysis presented above is based on a specific interpretation of quantum mechanics. Interpretations that do not invoke the collapse of the state of the system in the act of observation have been proposed.[15]

[14] Evidently, this argument considers an ideal detector, to whose vigilance no photon escapes.

[15] See the additional material entitled *Additional Informations* for a brief review of the chief alternative interpretations.