Linear and circular photon spin states in the Mach-Zehnder interference experiment

Jean Louis Van Belle, Drs, MA*Ec*, BA*Ec*, B*Phil*

Email: jeanlouisvanbelle@outlook.com

5 November 2018

Abstract: This paper continues to explore a possible physical interpretation of the wavefunction but with a focus on the wavefunction(s) of a single photon in the Mach-Zehnder experiment. It focuses, in particular, on how one might visualize linear and circular polarization states for photon waves, and how beam splitters may or may not split a circular polarization state into two independent linear polarization states or – vice versa – recombine two linear polarization states into one circular state.

As such, it attempts to provide a more refined approach to the rather crude hidden-variable theory for explaining quantum-mechanical interference that was presented in a previous paper (http://vixra.org/pdf/1811.0005v1.pdf). The outcome is the same, however: the theory does not work. Hence, this paper again shows the limit of such physical interpretations, thereby confirming the intuition behind Bell's Theorem.

Contents

Linear and circular photon spin states in the Mach-Zehnder interference experiment

Introduction

Duns Scotus wrote: *pluralitas non est ponenda sine necessitate*. Plurality is not to be posited without necessity.[1] William of Ockham turned this idea into an intuitive *lex parsimoniae*: the simplest solution tends to be the correct one.[2] However, redundancy in the description does not seem to bother physicists. When explaining the basic axioms of quantum physics in his famous *Lectures* on quantum mechanics, Richard Feynman writes:

> "We are not particularly interested in the mathematical problem of finding the minimum set of independent axioms that will give all the laws as consequences. Redundant *truth* does not bother us. We are satisfied if we have a set that is complete and not apparently inconsistent."[3]

Some *ambiguity* in the description is apparently not eschewed either. For example, most introductory courses on quantum mechanics will show that both $\psi = \exp(i\theta) = \exp[i(kx-\omega t)]$ and $\psi^* = \exp(-i\theta) = \exp[-i(kx-\omega t)] = \exp[i(\omega t-kx)] = -\psi$ are acceptable waveforms to describe a particle that is propagating in the *x*-direction. Both have the required mathematical properties—as opposed to, say, some real-valued sinusoid.[4] We would then think some proof should follow of why one would be better than the other or, preferably, one would expect as a discussion on what these two mathematical possibilities might represent—but, no. That does not happen. The physicists conclude that "the choice is a matter of convention and, happily, most physicists use the same convention."[5]

Instead of making a choice here, we could, perhaps, use the various mathematical possibilities to incorporate spin in the description, as real-life particles – think of electrons and photons here – have two spin states[6] (*up* or *down*), as shown below.

[1] Duns Scotus, *Commentaria*.

[2] https://en.wikipedia.org/wiki/Occam%27s_razor.

[3] Feynman's *Lectures* on Quantum Mechanics, Vol. III, Chapter 5, Section 5.

[4] The argument is based on whether or not the superposition of similar waveforms gives us a sensible composite waveform.

[5] See, for example, the MIT's edX Course 8.04.1*x* (Quantum Physics), Lecture Notes, Chapter 4, Section 3.

[6] Photons are spin-one particles but they do *not* have a spin-zero state.

Table 1: Matching mathematical possibilities with physical realities?[7]

Spin and direction of travel	**Spin up**	**Spin down**
Positive *x*-direction	$\psi = \exp[i(kx-\omega t)]$	$\psi^* = \exp[-i(kx-\omega t)] = \exp[i(\omega t-kx)]$
Negative *x*-direction	$\chi = \exp[-i(kx+\omega t)] = \exp[i(\omega t-kx)]$	$\chi^* = \exp[i(kx+\omega t)]$

This seems to make sense. Theoretical spin-zero particles do not exist and we should therefore, perhaps, use the extra degree of freedom in the mathematical description to describe the spin state of our particle. An important added benefit here is that the weird 720-degree symmetry of spin-1/2 particles collapses into an ordinary 360-degree symmetry and that we would, therefore, have no need, perhaps, to describe them using spinors and other complicated mathematical objects. We have written about this at length elsewhere[8] and so we won't repeat ourselves here.

Let us return to the topic of ambiguity in the description. It seems to apply to the concept of spin or polarization states for photons. Indeed, when discussing electrons, we can think of a spinning charge and, therefore, we can easily relate it to classical concepts.[9] In contrast, the same discussion for photons becomes complicated. This paper explores some elements in this discussion that may or may not be useful for a better understanding.

Linear polarization states

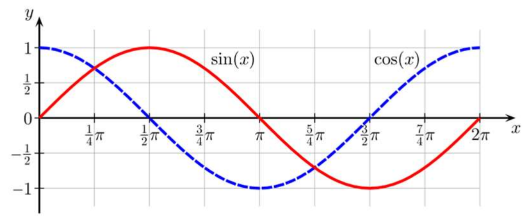
Electrons have a wavefunction and, as mentioned above, we can come up with easy geometric or physical interpretations. Unfortunately, such easy geometric or physical interpretations seem to break down when tryring to explain the weird results we get from interference experiments. When analyzing interference, the wavefunction concept gives way to the concept of a probability amplitude which we associate with a possible *path* rather than a particle. The math looks somewhat similar but models very different ideas and concepts. Before the electron goes through the two slits, we had one wavefunction. When it goes through, we have two probability amplitudes that – somehow – recombine to give us a diffraction pattern.

Do we have a wavefunction for the photon when it hits the first beam splitter in the Mach-Zehnder interferometer? If we do, how do we capture the spin states? We can easily distinguish between left- and right-hand *circular* polarization, but if we have *linearly* polarized light, can we distinguish between a plus and a minus direction? Maybe. Maybe not. Suppose the light is polarized along the *x*-direction. We know the component of the electric field vector along the y-axis[10] will then be equal to $E_y = 0$, and the magnitude of the *x*-component of $\boldsymbol{E}$ will be given by a sinusoid. However, here we have two distinct

[7] Of course, the formulas only give the *elementary* wavefunction. The wave *packet* will be a Fourier sum of such functions.
[8] Jean Louis Van Belle, *Euler's wavefunction* (http://vixra.org/abs/1810.0339, accessed on 30 October 2018)
[9] Jean Louis Van Belle, *In Search of Schrödinger's Electron* (http://vixra.org/abs/1809.0350, accessed on 4 November 2018)
[10] The *z*-direction is the direction of wave propagation in this example. In quantum mechanics, we often define the direction of wave propagation as the *x*-direction. This will, hopefully, not confuse the reader. The choice of axes is usually clear from the context.

possibilities: $E_x = cos(\omega \cdot t)$ or, alternatively, $E_x = sin(\omega \cdot t)$. These are the same functions but – crucially important – with a phase difference of 90°: $sin(\omega \cdot t) = cos(\omega \cdot t + \pi/2)$.

Figure 1: Two varieties of linearly polarized light?[11]

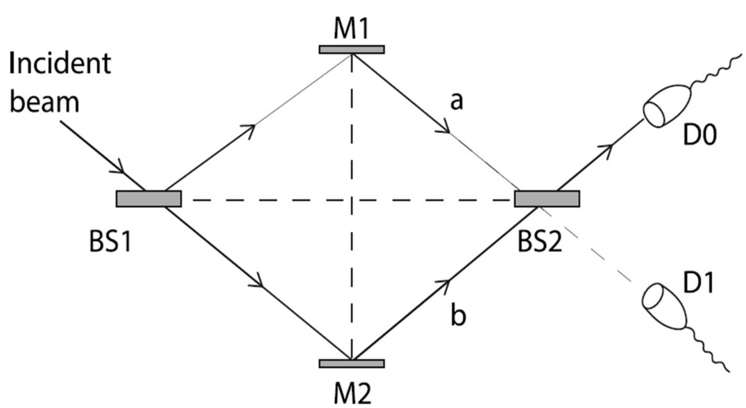


Would this matter? Sure. We can easily come up with some classical explanations of why this would matter. Think, for example, of birefringent material being defined in terms of quarter-wave plates. In fact, the more obvious question is: why would this *not* make a difference, and why would we have *two* possibilities only? What if we add an additional 90° shift to the phase? We know that $cos(\omega \cdot t + \pi) = -cos(\omega \cdot t)$. We cannot reduce this to $cos(\omega \cdot t)$ or $sin(\omega \cdot t)$. Hence, if we think in terms of 90° phase differences, then $-cos(\omega \cdot t) = cos(\omega \cdot t + \pi)$ and $-sin(\omega \cdot t) = sin(\omega \cdot t + \pi)$ are different waveforms too. And why should we think in terms of 90° phase shifts only? Why shouldn't we think of a continuum of linear polarization states? The answer is: we probably should.

Linear polarization in the Mach-Zehnder experiment

Let us look at the Mach-Zehnder interferometer once again. We have two beam splitters (BS1 and BS2) and two perfect mirrors (M1 and M2). An incident beam coming from the left is split at BS1 and recombines at BS2, which sends two outgoing beams to the photon detectors D0 and D1. More importantly, the interferometer can be set up to produce a precise interference effect which ensures all the light goes into D0, as shown below. Alternatively, the setup may be altered to ensure all the light goes into D1.

Figure 2: The Mach-Zehnder interferometer[12]



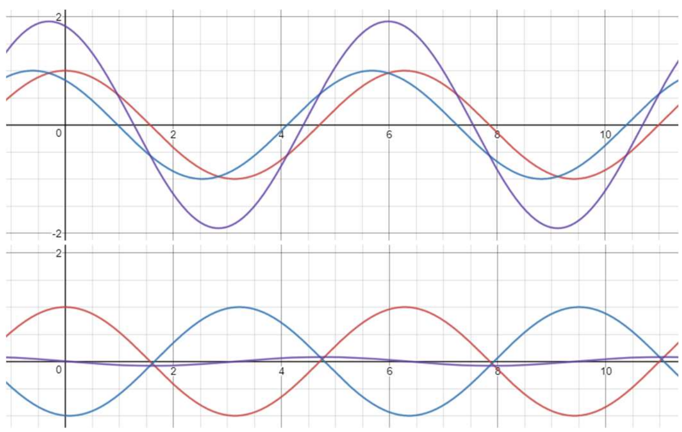
[11] Source of the illustration: https://upload.wikimedia.org/wikipedia/commons/7/71/Sine_cosine_one_period.svg..
[12] Source of the illustration: MIT edX Course 8.04.1*x* (Quantum Physics), Lecture Notes, Chapter 1, Section 4 (*Quantum Superpositions*).

What is the classical explanation? The classical explanation is something like this: the first beam splitter (BS1) splits the beam into *two* beams. These two beams arrive in phase or, alternatively, out of phase and we, therefore, have constructive or destructive interference that recombines the original beam and makes it go towards D0 or, alternatively, towards D1. That explanation does not make any sense when thinking of a continuum of linear polarization states. If beam *a* and *b* are linearly polarized, then we can describe them by $cos(\omega \cdot t - k \cdot x) = cos(\theta)$ and $cos(\theta + \Delta)$ respectively. In the classical analysis, the difference in phase (Δ) will be there because of a difference of the path lengths[13] and the recombined wavefunction will be equal to the same cosine function, but with argument $\theta + \Delta/2$, multiplied by an *envelope* equal to $2 \cdot \cos(\Delta/2)$. We write[14]:

$$\cos(\theta) + \cos(\theta + \Delta) = 2 \cdot \cos(\theta + \Delta/2) \cdot \cos(\Delta/2)$$

We always get a recombined beam with the same frequency, but when the phase difference between the two incoming beams is small, its amplitude is going to be much larger. To be precise, it is going to be twice the amplitude of the incoming beams for $\Delta = 0$. In contrast, if the two beams are out of phase, the amplitude is going to be much smaller, and it's going to be zero if the two waves are 180 degrees out of phase ($\Delta = \pi$), as shown below. That doesn't make sense because twice the amplitude means *four* times the energy, and zero amplitude means zero energy. The energy conservation law is being violated: photons are being multiplied or, conversely, are being destroyed.

Figure 3: Constructive and destructive interference for linearly polarized beams



Can we solve the problem by assuming that, when the beam splits at BS1, the energy of the *a* and *b* beam must be split in half too? We know the energy is given by (or, to be precise, proportional to) the square of the amplitude (let us denote this amplitude by A).[15] Hence, if we want the energy of the two individual beams to add up to $A^2 = 1^2 = 1$, then the (maximum) amplitude of the *a* and *b* beams must be $1/\sqrt{2}$ of the amplitude of the original beam, and our formula becomes:

[13] Feynman's path integral approach to quantum mechanics allows photons (or probability amplitudes, we should say) to travel somewhat slower or faster than *c*, but that should not bother us here.

[14] We are just applying the formula for the sum of two cosines here. If we would add sines, we would get $\sin(\theta) + \sin(\theta + \Delta) = 2 \cdot \sin(\theta + \Delta/2) \cdot \cos(\Delta/2)$. Hence, we get the same envelope: $2 \cdot \cos(\Delta/2)$.

[15] If we would to reason in terms of average energies, we would have to apply a 1/2 factor because the average of the $\sin^2\theta$ and $\cos^2\theta$ over a cycle is equal to 1/2.

$$(1/\sqrt{2})\cdot\cos(\theta) + (1/\sqrt{2})\cdot\cos(\theta + \Delta) = (2/\sqrt{2})\cdot\cos(\theta + \Delta/2)\cdot\cos(\Delta/2)$$

This reduces to $(2/\sqrt{2})\cdot\cos(\theta)$ for $\Delta = 0$. Hence, we still get *twice* the energy – $(2/\sqrt{2})^2$ equals 2 – when the beams are in phase and zero energy when the two beams are 180 degrees out of phase. This doesn't make sense.

Of course, the mistake in the argument is obvious: we cannot just add the amplitudes of the *a* and *b* beams because they have different directions. If the *a* and *b* beams – after being split from the original beam – are linearly polarized, then the angle between the axes of polarization will be 90 degrees. Hence, we cannot just add the two amplitudes. In fact, we should ask ourselves: can we add them at all? The incident angle between the two beams in a Mach-Zehnder apparatus is 90 degrees and the two oscillations are, therefore, independent. Hence, we need to add them like we would add the two parts of a complex number. Remembering the geometric interpretation of the imaginary unit as a counterclockwise rotation, we can – perhaps – try writing the sum of our a and b beams as:

$$(1/\sqrt{2})\cdot\cos(\theta) + i\cdot(1/\sqrt{2})\cdot\cos(\theta + \Delta) = (1/\sqrt{2})\cdot[\cos(\theta) + i\cdot\cos(\theta + \Delta)]$$

What can we do with this? Not all that much, except noting that we can write the *cos*$(\theta + \Delta)$ as a sine for $\Delta = \pm \pi/2$. To be precise, we get:

$$(1/\sqrt{2})\cdot\cos(\theta) + i\cdot(1/\sqrt{2})\cdot\cos(\theta + \pi/2) = (1/\sqrt{2})\cdot(\cos\theta - i\cdot\sin\theta) = (1/\sqrt{2})\cdot e^{-i\cdot\theta}$$

$$(1/\sqrt{2})\cdot\cos(\theta) + i\cdot(1/\sqrt{2})\cdot\cos(\theta - \pi/2) = (1/\sqrt{2})\cdot(\cos\theta + i\cdot\cos\theta) = (1/\sqrt{2})\cdot e^{i\cdot\theta}$$

Combining circular and linear polarization states

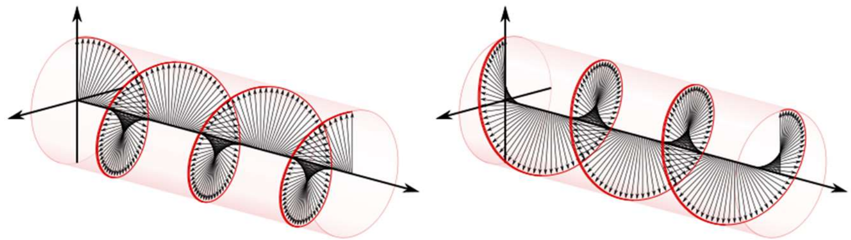
The result above is interesting, because we can now build an alternative theory of what might happen in the Mach-Zehnder interferometer:

1. If we would assume the incoming light has *either* a left- *or* a right-handed circular polarization; and
2. If the first beam splitter would effectively split the beam into two linearly polarized waves; and
3. If the second beam splitter would combine those two beams back into a circularly polarized wave; and
4. Then we can, effectively, explain the binary outcome of the Mach-Zehnder experiment – at the level of a photon – in terms of an alternative theory.

What about the $1/\sqrt{2}$ factor? If the $e^{-i\cdot\theta}$ and $e^{i\cdot\theta}$ wavefunctions can, effectively, be interpreted geometrically as a *physical* oscillation in *two* dimensions, as illustrated below[16], then each of the two (independent) oscillations will pack one half of the energy of the wave. Hence, if such *circularly* polarized wave splits into two *linearly* polarized waves, then the two linearly polarized waves will effectively, pack half of the energy without any need for us to think their (maximum) amplitude should be adjusted.

[16] Such physical interpretation is very easy in the case of light. We only need to assign the physical dimension of the electric field (force per unit charge, N/C) to the two perpendicular oscillations.

Figure 4: Left- and right-handed polarization[17]



Let us think of the geometry here. If ***x*** is the direction of propagation of the wave, then the ***z***-direction will be pointing upwards, and we get the ***y***-direction from the righthand rule for a Cartesian reference frame.[18] We may now think of the oscillation along the ***y***-axis as the cosine, and the oscillation along the ***z***-axis as the sine. If we then think of the imaginary unit *i* as a 90-degree counterclockwise rotation in the ***yz***-plane (and remembering the convention that angles (including the phase angle θ) are measured counterclockwise), then the right- and left-handed waves can be represented by the following wavefunctions:

$$cos\theta + i\cdot sin\theta = e^{i\cdot\theta} \text{ (RHC)}$$

$$cos(-\theta) + i\cdot sin(-\theta) = cos\theta - i\cdot sin\theta = e^{-i\cdot\theta} \text{ (LHC)}$$

If we now think of the ***x***-direction as the direction of the incident beam in the Mach-Zehnder experiment, and we would want to also think of rotations in the ***xz***-plane, then we need to need to introduce some new convention here. Let us introduce *another* imaginary unit, which we'll denote by *j*, and which will represent a 90-degree counterclockwise rotation in the ***xz***-plane.[19]

Photon polarization	**At BS1**	**At mirror**	**At BS2**	**Final result**
RHC	Photon ($e^{i\cdot\theta} = cos\theta + i\cdot sin\theta$) is split into two linearly polarized beams: Upper beam (vertical oscillation) = $j\cdot\sin\theta$ Lower beam (horizontal oscillation) = $\cos\theta$	The vertical oscillation gets rotated clockwise and becomes $-j\cdot j\cdot\sin\theta = -j^2\cdot\sin\theta = \sin\theta$ The horizontal oscillation is not affected and is still represented by $\cos\theta$	Photon is recombined. The upper beam gets rotated counter-clockwise and becomes $j\cdot\sin\theta$. The lower beam is still represented by $\cos\theta$	The photon wavefunction is given by $\cos\theta + j\cdot\sin\theta = e^{+j\cdot\theta}$. This is an RHC photon travelling in the *xz*-plane but rotated over 90 degrees.
LHC	Photon ($e^{-i\cdot\theta} = cos\theta - i\cdot sin\theta$) is split into two linearly polarized beams:	The vertical oscillation gets rotated clockwise and becomes	Photon is recombined. The upper beam gets rotated counter-clockwise and	The photon wavefunction is given by $\cos\theta - j\cdot\sin\theta = e^{-j\cdot\theta}$.

[17] Credit: https://commons.wikimedia.org/wiki/User:Dave3457.

[18] Note the reference frame in the illustrations of the LHC and RHC wave is left-handed. We will want to think in terms of a regular right-handed reference frame.

[19] This convention may make the reader think of the quaternion theory but we are thinking more of simple Euler angles here: *i* is a (counterclockwise) rotation around the *x*-axis, and *j* is a rotation around the *y*-axis.

	Upper beam (vertical oscillation) = $-j\cdot\sin\theta$ Lower beam (horizontal oscillation) = $\cos\theta$	$(-j)\cdot(-j)\cdot\sin\theta = = j^2\cdot\sin\theta = -\sin\theta$ The horizontal oscillation is not affected and is still represented by $\cos\theta$	becomes $-j\cdot\sin\theta$. The lower beam is still represented by $\cos\theta$	This is an LHC photon travelling in the *xz*-plane but rotated over 90 degrees.

Of course, we may also set up the apparatus with different path lengths, in which case the two linearly polarized beams will be out of phase when arriving at BS1. Let us assume the phase shift is equal to $\Delta = 180° = \pi$. This amounts to putting a minus sign in front of *either* the sine *or* the cosine function. Why? Because of the $\cos(\theta \pm \pi) = -\cos\theta$ and $\sin(\theta \pm \pi) = -\sin\theta$ identities. Let us assume the distance along the upper path is longer and, hence, that the phase shift affects the sine function.[20] In that case, the sequence of events might be like this:

Photon polarization	At BS1	At mirror	At BS2	Final result
RHC	Photon ($e^{i\cdot\theta} = \cos\theta + i\cdot\sin\theta$) is split into two linearly polarized beams: Upper beam (vertical oscillation) = $j\cdot\sin\theta$ Lower beam (horizontal oscillation) = $\cos\theta$	The vertical oscillation gets rotated clockwise and becomes $-j\cdot j\cdot\sin\theta = -j^2\cdot\sin\theta = \sin\theta$ The horizontal oscillation is not affected and is still represented by $\cos\theta$	Photon is recombined. The upper beam gets rotated counter-clockwise and – because of the longer distance – becomes $j\cdot\sin(\theta + \pi) = -j\cdot\sin\theta$. The lower beam is still represented by $\cos\theta$	The photon wavefunction is given by $\cos\theta - j\cdot\sin\theta = e^{-j\cdot\theta}$. This is an LHC photon travelling in the *xz*-plane but rotated over 90 degrees.
LHC	Photon ($e^{-i\cdot\theta} = \cos\theta - i\cdot\sin\theta$) is split into two linearly polarized beams: Upper beam (vertical oscillation) = $-j\cdot\sin\theta$ Lower beam (horizontal oscillation) = $\cos\theta$	The vertical oscillation gets rotated clockwise and becomes $(-j)\cdot(-j)\cdot\sin\theta = = j^2\cdot\sin\theta = -\sin\theta$ The horizontal oscillation is not affected and is still represented by $\cos\theta$	Photon is recombined. The upper beam gets rotated counter-clockwise and – because of the longer distance – becomes $-j\cdot\sin(\theta + \pi) = +j\cdot\sin\theta$. The lower beam is still represented by $\cos\theta$	The photon wavefunction is given by $\cos\theta + j\cdot\sin\theta = e^{+j\cdot\theta}$. This is an RHC photon travelling in the *xz*-plane but rotated over 90 degrees.

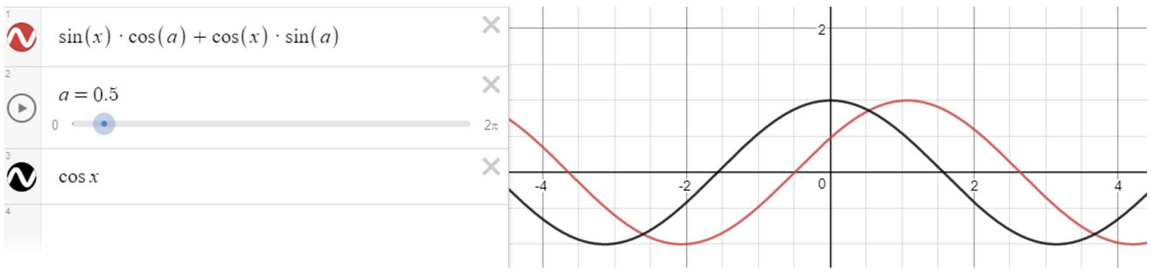
What happens when the difference between the phases of the two beams is not equal to 0 or 180 degrees? What if it is some random value in-between? Do we get an elliptically polarized wave or some other nice result? Denoting the phase shift as Δ, we can write:

$$\cos\theta + j\cdot\sin(\theta + \Delta) = \cos\theta + j\cdot(\sin\theta\cdot\cos\Delta + \cos\theta\cdot\sin\Delta)$$

[20] The reader can easily work out the math for the opposite case (longer length of the lower path).

However, this is also just a circularly polarized wave, but with a random phase shift between the horizontal and vertical component of the wave, as shown below. Of course, for the special values Δ = 0 and Δ = π, we get $cos\theta + j\cdot sin\theta$ and $cos\theta - j\cdot sin\theta$ once more.

Figure 5: Left- and right-handed polarization



Conclusion

Is this some sort of hidden-variable theory of how quantum-mechanical interference in Mach-Zehnder experiments *really* works? No.

First, the Mach-Zehnder interferometer does not work this way. We do *not* always get a nice (100%) output beam: we *do* have constructive and destructive interference and the simple model above does not explain that.[21]

Second, even if this model would work, we basically only diverted attention away from another problem: we may be able to explain the interference effect, but we are now not able to explain how these 50/50 beam splitters works. Indeed, why is it that, if we would measure the position of the photon immediately after it exits beam splitter BS1, the measurement would tell us the photon is *either* in the upper path or, *else*, in the lower path – and why is the probability of being in either equal to 50%?

In short, the mystery remains.

Jean Louis Van Belle, 5 November 2018

References

References are mentioned in footnotes.

[21] For a good *classical* explanation of the Mach-Zehnder interferometer, see: K.P. Zetie, S.F. Adams and R.M. Tocknell, January 2000, *How does a Mach–Zehnder interferometer work?* (https://www.cs.princeton.edu/courses/archive/fall06/cos576/papers/zetie_et_al_mach_zehnder00.pdf, accessed on 5 November 2018).
For a good quantum-mechanical explanation (interference of single photons), see – for example – the Mach-Zehnder tutorial from the PhysPort website (https://www.physport.org/curricula/QuILTs/, accessed on 5 November 2018).