Understanding the quantum pigeonhole principle of Aharonov *et al*.

W.M. Stuckey
Department of Physics, Elizabethtown College, Elizabethtown, PA 17022

Herein I will explain what is wrong with the prediction of the so-called "quantum pigeonhole principle" of Aharonov *et al*. as given in theory in the first section of their paper[(1)] and by a proposed experiment later therein. According to their claim, quantum mechanics says it is possible for three particles to simultaneously occupy two boxes without two or more particles occupying the same box. The pigeonhole principle (and intuition) says there must be at least two particles in any box, thus they write, "[The pigeonhole principle] seems therefore to be an abstract and immutable truth, beyond any doubt. Yet, as we show here, for quantum particles the principle does not hold."

We can ignore particle 3 to understand their conclusion (as they point out), and likewise to understand my analysis, so start with each of particle 1 and 2 in the state

$$|+> = \frac{1}{\sqrt{2}}(|L> + |R>) \tag{1a}$$

which is orthogonal to

$$|-> = \frac{1}{\sqrt{2}}(|L> - |R>) \tag{1b}$$

so that collectively

$$|\Psi> = |+>_1 |+>_2 \tag{2}$$

Now we do a measurement on Eq (1a) changing it to (upon re-normalization)

$$|+i> = \frac{1}{\sqrt{2}}(|L> + i|R>) \tag{3a}$$

which is orthogonal to

$$|-i> = \frac{1}{\sqrt{2}}(|L> - i|R>) \tag{3b}$$

so the new state is

$$|\Phi> = |+i>_1 |+i>_2 \tag{4}$$

The measurement resulting in Eq (4) is called the "final measurement." They talk about a counterfactual "intermediate measurement" taking $|\Psi >$ to $|\Phi >$ via the following basis (I'll call the "numerical basis")

$$|1>=\frac{1}{\sqrt{2}}(|L_1>|L_2>+|R_1>|R_2>)$$

$$|2>=\frac{1}{\sqrt{2}}(|L_1>|L_2>-|R_1>|R_2>)$$

$$|3>=\frac{1}{\sqrt{2}}(|L_1>|R_2>+|R_1>|L_2>)$$

$$|4>=\frac{1}{\sqrt{2}}(|L_1>|R_2>-|R_1>|L_2>)$$

Noting that $|\langle 1|\Psi\rangle|^2=\frac{1}{2}$ and $|\langle 3|\Psi\rangle|^2=\frac{1}{2}$ ("On the initial state alone, the probabilities to find particles 1 and 2 in the same box and in different boxes are both 50%."), while $|\langle 1|\Phi\rangle|^2=0$, $|\langle 2|\Phi\rangle|^2=\frac{1}{2}$ and $|\langle 3|\Phi\rangle|^2=\frac{1}{2}$, and assuming "that at the intermediate time we find the particles in the same box," they conclude:

> Hence in this case the final measurements cannot find the particles in the state $|\Phi >$. Therefore, the only cases in which the final measurement can find the particles in the state $|\Phi >$ are those in which the intermediate measurement found that particles 1 and 2 are in different boxes.
>
> Crucially, as noted before, the state is symmetric under permutation, hence what is true for particles 1 and 2 is true for all pairs. In other words, given the above pre- and post-selection, we have three particles in two boxes, yet no two particles can be found in the same box – our quantum pigeonhole principle.

If I understand their argument, then the claim that "the only cases in which the final measurement can find the particles in the state $|\Phi >$ are those in which the intermediate measurement found that particles 1 and 2 are in different boxes," means that $|\Psi >$ and $|\Phi >$ do not share a mutual non-zero projection onto a state representing particles 1 and 2 being in the same box. But, both $|\Psi >$ and $|\Phi >$ have non-zero projections onto every element of this basis (I'll call the "alphabetical basis")

$$|A>=|L_1>|L_2>$$

$$|B>=|R_1>|R_2>$$

$$|C>=|L_1>|R_2>$$

$$|D>=|R_1>|L_2>$$

which includes two states where particles 1 and 2 are in the same box. Indeed, the probability of particles 1 and 2 both being found in box *L* or *R* for state $|\Phi>$ is

$$|\langle L_1L_2|\Phi\rangle|^2 = |\langle L_1L_2|{+i_1}{+i_2}\rangle|^2 = \frac{1}{4}$$

or

$$|\langle R_1R_2|\Phi\rangle|^2 = |\langle R_1R_2|{+i_1}{+i_2}\rangle|^2 = \frac{1}{4}$$

which is exactly the same result we have for $|\langle L_1L_2|\Psi\rangle|^2$ or $|\langle R_1R_2|\Psi\rangle|^2$. Thus, introducing the factor $+i$ to the *R* term for both particles has no effect on this probability. So, the lack of a route from $|\Psi>$ to $|\Phi>$ through $|1>$ of the first basis does not justify their conclusion. In fact, we can understand the lack of that route very simply, and this understanding proves important for understanding their mistaken prediction for their proposed experiment.

$\langle 1|\Phi\rangle = 0$ doesn't mean you can't find particles 1 and 2 in the same box (as I showed above). It's referring to the fact that you can't get $(|L_1>|L_2>+|R_1>|R_2>)$ from $(|L>+i|R>)$ because to get $|R_1>|R_2>$ you need $(\pm i|R_1>)(\mp i|R_2>)$, but both particle 1 and particle 2 have $+i|R>$ and the distinction between $+i|R>$ and $-i|R>$ is what makes Eq (3a) orthogonal to Eq (3b). However, $\langle 2|\Phi\rangle \neq 0$ because you can get $-|R_1>|R_2>$ in $(|L_1>|L_2>-|R_1>|R_2>)$ from $(i|R_1>)(i|R_2>)$. These computations are telling you whether or not you can get *both* $|L_1>|L_2>$ *and* $\pm|R_1>|R_2>$ when particles 1 and 2 are in the state given by Eq (3a). A similar argument obtains for the results of $\langle 3|\Phi\rangle$ and $\langle 4|\Phi\rangle$. And, the same reasoning applies to why $\langle 1|\Psi\rangle \neq 0$ and $\langle 2|\Psi\rangle = 0$ (as well as the results of $\langle 3|\Psi\rangle$ and $\langle 4|\Psi\rangle$) from the state given by Eq (1a). Clearly, the question you are answering with this calculation is not, "Can I find particles 1 and 2 in the same box?" Now let's look at the experiment.

They consider measuring the interaction between electrons in an interferometer as a means of empirically determining whether or not two electrons occupy the same path (box) at the same time. [See the Appendix for a pedagogical introduction to the interferometer instantiations of the states used here.] Again, let's consider a pair of electrons and that will make it clear as to how to analyze three electrons.

First, there's no reason to believe that electrons will not interact simply because they're in an interferometer. And, per the notation of Aharonov *et al.*, the result of that interaction depends on their spatial separation (as given by V) and interaction duration (as given by T). In order to simplify the calculation, let's consider the case where the electrons on each path are emitted at the same time and each such pair is isolated from other such pairs. [It will be obvious how to calculate a deviation in the results predicted by this assumption if you relax the assumption. All we're trying to do here is find the flaw in the logic of the proposed quantum pigeonhole experiment and, as will be clear, this suffices.] Now, either the electrons of each pair go through the interferometer on the same path, as given by $|L_1>|L_2>$ and $|R_1>|R_2>$, in which case they interact, or they take different paths, as given by $|L_1>|R_2>$ and $|R_1>|L_2>$, and they don't interact. Therefore, we expect to see four spots on the detector face, two spots for the pair of interacting beams and two spots for the pair of non-interacting beams. The amount of transverse separation of the pairs of interacting and non-interacting beams is given by V and T, and has nothing to do the interferometer states (in this simplified situation). The interferometer states become relevant when one wants to know the relative intensities of the spots. To answer that question, all we need is the fraction of outcomes (intensity) for the interacting cases $|L_1>|L_2>$ and $|R_1>|R_2>$, and the fraction of outcomes for the non-interacting cases $|L_1>|R_2>$ and $|R_1>|L_2>$. This is where the analysis above is relevant.

If you naively assume the intensity for, say, the interacting cases is given by $|\langle 1|\Psi\rangle|^2$, then you would predict that the separation *per se* (not just the intensity of the separated beams) between interacting and non-interacting cases depends on the value of the phase shifter (PS, see Appendix if necessary). For example, as shown above, we have $|\langle 1|\Psi\rangle|^2 = 0$ for $|\Psi> = |+i>_1 |+i>_2$ and $|\langle 1|\Psi\rangle|^2 \neq 0$ for $|\Psi> = |+>_1 |+>_2$. Of course, unless the electrons are in the PS for a relatively long time and the PS greatly affects V, we don't expect to see the PS affect the spread of the beams. Ignoring that effect (which doesn't bear on the issue here[1]), we don't expect any difference in the intensities of the interacting versus non-interacting beams between the cases for $|\Psi> = |+i>_1 |+i>_2$ and $|\Psi> = |+>_1 |+>_2$, since $|\langle L_1 L_2|\Psi\rangle|^2$ and $|\langle R_1 R_2|\Psi\rangle|^2$ are the same in both of these cases. The problem is obviously with the naïve assumption. Clearly the

[1] Indeed, $|\Psi> = |+>_1 |+>_2$ gives the same result for $|\langle 1|\Psi\rangle|^2$ as $|\Psi> = |->_1 |->_2$.

relative intensity between interacting and non-interacting beams in these cases is ½, so as I showed above, if you want to use the numerical basis in lieu of the alphabetical basis to write the intensity for these two cases, you have to use $|\langle 1|\Psi\rangle|^2$ for $|\Psi> = \ |+>_1\ |+>_2$ and $|\langle 2|\Psi\rangle|^2$ for $|\Psi> = \ |+i>_1\ |+i>_2$ for the reasons given above. Aharonov *et al*. used the numerical basis, but did not account for the different possible values of PS when constructing their interaction Hamiltonian. This amounts to taking the inner product with $|1>$ regardless of the value of PS, so (as I showed above) their probabilities will sometimes erroneously tell them that the particles are not in the same path. And, in those cases where their calculation erroneously tells them the particles are not in the same path, they will predict no separation of the beams when in fact there will be separation.

The bottom line is the authors did not use the correct functional form for the projections in their interaction Hamiltonian because they did not include the (trivial) calculational effect the PS introduces to the computation of probability. Due to the flaws in their analysis pointed out here, their paper provides no reason to believe that "for quantum particles the [pigeonhole] principle does not hold." Of course, the correct functional form for the interaction Hamiltonian is amendable to empirical investigation. So, do the experiment and we'll see who's right!

Appendix

To construct the amplitude at D1 or D2, simply add contributions from beam splitters, mirrors and the PS where applicable along paths *L* and *R* and combine them to D1 or D2 (see their figure below). The amplitude at D2 with no PS is (path *R* contribution + path *L* contribution)

$$\left(\frac{1}{\sqrt{2}}\right)(i)\left(\frac{i}{\sqrt{2}}\right) + \left(\frac{i}{\sqrt{2}}\right)(i)\left(\frac{1}{\sqrt{2}}\right) = -1$$

The amplitude at D1 with no PS is (path *R* contribution + path *L* contribution)

$$\left(\frac{1}{\sqrt{2}}\right)(i)\left(\frac{1}{\sqrt{2}}\right) + \left(\frac{i}{\sqrt{2}}\right)(i)\left(\frac{i}{\sqrt{2}}\right) = 0$$

They let this configuration and outcome (D2 intensity = 1, D1 intensity = 0) constitute a measurement outcome associated with the state $|+>$. Now if we add a PS of π to path *R*, we've added a factor of –1 to the path *R* contribution in the above computations. Thus, the amplitude at D2 is now

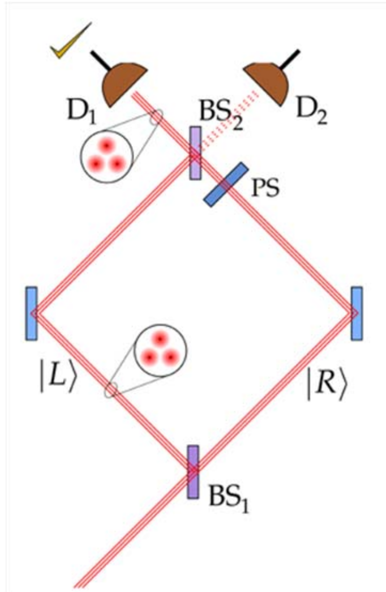
$$(-1)\left(\frac{1}{\sqrt{2}}\right)(i)\left(\frac{i}{\sqrt{2}}\right)+\left(\frac{i}{\sqrt{2}}\right)(i)\left(\frac{1}{\sqrt{2}}\right)=0$$

And the amplitude at D1 is now

$$(-1)\left(\frac{1}{\sqrt{2}}\right)(i)\left(\frac{1}{\sqrt{2}}\right)+\left(\frac{i}{\sqrt{2}}\right)(i)\left(\frac{i}{\sqrt{2}}\right)=-i$$

They let this configuration and outcome (D2 intensity = 0, D1 intensity = 1) constitute a measurement outcome associated with the state $|->$.

If you add PS = π/2, you add a factor of *i* to the contribution from path *R* giving equal amplitudes at D1 and D2 of $-\frac{i}{2}-\frac{1}{2}$ corresponding to an intensity of ½. If you add PS = 3π/2, you add a factor of – *i* to the contribution from path *R* giving amplitudes at D1 and D2 that differ by a factor of –1, i.e., $\pm\frac{i}{2}\mp\frac{1}{2}$ at D2/D1. This still corresponds to an intensity of ½ at both D1 and D2, so haven't created the orthogonal states $|\pm i>$. You can modify the interferometer to make the $|\pm i>$ measurement by replacing D1 and D2 with mirrors, putting a PS2 = π/2 after the mirror on the right (say), then having the beams pass through another beam splitter in route to the new D1 (upper left) and D2 (upper right). This gives the amplitude at D2 = 0 and D1 = $\sqrt{2}\left(\frac{i}{2}+\frac{1}{2}\right)$ for PS = π/2 and D2 = $i\sqrt{2}\left(-\frac{i}{2}+\frac{1}{2}\right)$ and D1 = 0 for PS = 3π/2.



References

(1) Aharonov, Y., Colombo, F., Popescu, S., Sabadini, I., Struppa, D.C., & Tollaksen, J.: The quantum pigeonhole principle and the nature of quantum correlations. (2014) http://arxiv.org/abs/1407.3194v1 .