Notes on David Mermin's article "Is the moon there when nobody looks?"

It all comes down to the question: *Where does the information come from that tells the detectors what to do (whether to flash red or green)?* The information has to come from somewhere. The detectors can't be randomly flipping coins independently to decide how to flash; if they were, then when the two detectors had the same switch settings, sometimes their flashes would disagree because their coin tosses came out differently. That never happens. Whenever the switches are set the same, the detectors *always* match.

Maybe the information is carried by the particles. If that is true, since a particle cannot know in advance which switch setting it will encounter when it reaches a detector (the switches could be set randomly *after* the particle leaves the source), the particle has to be prepared for any contingency: it has to carry *three* pieces of information with it, telling the detector what to do for each possible switch setting. Said another way, the particle has to have three definite properties, which we can call its "1-color", "2-color", and "3-color". We'll abbreviate them as RRR, RGR, GGR, etc. The experimental design ensures that whenever the source creates a new pair of particles, they will both share the same set of color properties.

When we see the detectors flash, we learn what some of the color properties are, but there will always be at least one color property that remains unknown to us, since there are only two switches. For example, if the switches are set to 12, we'll learn the left particle's 1-color and the right particle's 2-color, but we still won't know their 3-color. If the switches are set to 33, we won't learn their 1-color or 2-color. But Einstein would say that all three particle properties nevertheless exist, independently of our knowledge of them (or lack thereof).

If that is indeed the case, then we can analyze what must happen for all nine possible switch combinations: 11, 12, 13, 21, 22, 23, 31, 32, 33. If the particle colors are RRR or GGG, we'll get a match in 9 out of the 9 cases, no matter how the switches are set. If the particle colors are mixed (RRG, GRG, RGG, etc.), we'll get a match in 5 of the 9 cases. So no matter what, we would expect to get a match *at least* $\frac{5}{9}$ of the time, since the switches are set randomly on each trial. That means that in an experiment with a million trials, we would expect to see at least 555,555 matches.

But that is *not* what happens when we do the actual experiment. Instead, we get a match approximately 50% of the time. Out of a million trials, we observe only about 500,000 matches, which is significantly less than 555,555. The $\geq \frac{5}{9}$ inequality is violated. So the particles *cannot* have had three definite color properties before arriving at the detector. The information telling the detectors how to flash must be coming from somewhere else.

According to quantum theory, the information comes essentially from a random coin flip that occurs as soon as the first particle of a pair arrives at a detector. For example, if the left particle arrives first and the detector's switch is set to 1, the particle acquires a 1-color of red or green, chosen randomly by the coin flip, and the detector flashes that color. But because the particles were created together in an entangled state, the right particle somehow instantly "hears" this coin flip as well, no matter how far away from the left particle it is, and instantly acquires the *same* 1-color property as the left particle. From that point on—even if the right particle hasn't yet arrived at its detector (and even if, for that matter, the switch on the right detector hasn't been set yet)—the particle has a definite 1-color, which is identical to the left particle's 1-color. However, both particles still lack a 2-color or a 3-color.

What happens when the right particle eventually arrives at its detector? If the right detector is set to 1, it will flash the same color that the left detector flashed, with 100% certainty, because the right particle has the same 1-color as the left particle. No coin flip is needed for the right particle in this case. As another example, the same thing would happen if the detectors were both set to 2 and the right particle were the first to arrive: it would randomly acquire a definite 2-color (red or green), which would also be instantly acquired by the left particle, so that when the left particle arrived, its detector would be guaranteed to flash the same color as the right detector. In general, if the switch settings of the two detectors match, they are guaranteed to flash the same color, no matter which particle is the first to arrive, and regardless of which color it randomly acquires upon its arrival. This accounts for Feature I of the data, as illustrated in Figure 4.

But what if the right particle, upon arriving at the right detector, finds the switch set to 2, instead of 1? According to quantum theory, the particle would *lose* its definite 1-color at that moment, a new coin flip would ensue, and the particle would acquire a definite 2-color (red or green) based on the coin flip, which the detector would then flash.

The new 2-color might, purely by chance, be the same as the old 1-color, thus producing identical flashes, but it might be different. What is the probability that the new 2-color will be the same as the old 1-color? You might think the probability would be $\frac{1}{2}$, but it turns out that the coin is biased. According to the mathematics of quantum theory, the probability depends on the angle θ between the switch settings, and is equal to $\cos^2(\theta/2)$. (To be more precise, in a real experiment this angle refers to the orientations of the *detector magnets* controlled by the switches, not to the switches themselves.) Switch 1 corresponds to $0°$, switch 2 to $120°$, and switch 3 to $240°$, so the probability that the new 2-color is the same as the old 1-color is $\cos^2((120° - 0°)/2) = \frac{1}{4}$. If the switch were set to 3 instead of 2, the probability that the new 3-color would be the same as the old 1-color would be $\cos^2((240° - 0°)/2)$, which is also $\frac{1}{4}$. In fact, for any two *different* switch settings, the probability always works out to $\frac{1}{4}$. Notice, however, when the two switch settings are the same, we get $\cos^2(0) = 1$, meaning that the new color is guaranteed to be the same as the old color, so in effect the coin flip doesn't occur.

To recap, we can summarize the situation as follows. If the two detectors have the same switch settings (11, 22, or 33), then they are guaranteed with probability 1 to flash identical colors. If they have different settings (12, 13, 21, 23, 31, or 32), then the probability that they will flash identical colors is only $\frac{1}{4}$, regardless of the particular switch combination. Since the switches are set randomly on each trial, they have a $\frac{1}{3}$ chance of being the same (3 out of 9 combinations), and a $\frac{2}{3}$ chance of being different (6 out of 9 combinations). Therefore, the overall probability that the detectors will flash identical colors is: $\frac{1}{3} \times 1 + \frac{2}{3} \times \frac{1}{4} = \frac{1}{2}$.

Quantum theory thus predicts identical flashes half of the time, instead of $\geq \frac{5}{9}$ of the time, and this is indeed exactly what we observe when we do the actual experiment. This accounts for Feature II of the data, as illustrated in Figure 5. So the particles must acquire their color properties only *by being measured* by the detectors. Since they are entangled particles, as soon as one is measured, the other acquires exactly the same property. Furthermore, only *one color property at a time* can be known with certainty: as soon as a new property is measured (say, 2-color), the particle loses its old property (1-color or 3-color). This is Heisenberg's uncertainty principle in action.

Suppose that the right particle, during its journey to the right detector, acquires a definite 1-color of red, as a result of the left particle being measured by the left detector. We will denote the current state of the right particle as $|R_1\rangle$. When the particle arrives at its detector, a measurement is made of its N-color, according to the detector's switch setting N. How do we calculate the probability of observing red or green as the particle's new N-color?

The properties of 1-color, 2-color, and 3-color correspond to three different *observables*, which are mutually exclusive in the sense that it is impossible to know with certainty the value of more than one of them at a time. (Not all observables are mutually exclusive, but these three are.) Mathematically, an observable is represented as a square matrix. For example, here are the matrices that represent 1-color, 2-color, and 3-color, respectively:

$$C_1 \text{ (switch 1)} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$C_2 \text{ (switch 2)} = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -0.5 & 0.866 \\ 0.866 & 0.5 \end{bmatrix}$$

$$C_3 \text{ (switch 3)} = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -0.5 & -0.866 \\ -0.866 & 0.5 \end{bmatrix}$$

These switches correspond to measurements made at different orientation angles from the vertical, with switch 1 at 0°, switch 2 at 120°, and switch 3 at 240°. But in principle, we could have any number of other switches with different orientations, each of which would measure a different color property of the particle. (We could measure, for example, "45°-color", or "180°-color", or "−17.2°-color", etc.) In general, the matrix for a switch oriented at angle θ from the vertical is:

$$C_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

No matter which color property we decide to measure, we will only observe one of two possible outcomes: red or green. We can write these states as $|R_1\rangle$ and $|G_1\rangle$ for 1-color, $|R_2\rangle$ and $|G_2\rangle$ for 2-color, and $|R_3\rangle$ and $|G_3\rangle$ for 3-color. Mathematically, they are *eigenvectors* of their associated observable matrices (and are also called *eigenstates*), and can be represented as shown below:

$$|R_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |G_1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad |R_2\rangle = \begin{bmatrix} 0.5 \\ 0.866 \end{bmatrix} \quad |G_2\rangle = \begin{bmatrix} -0.866 \\ 0.5 \end{bmatrix} \quad |R_3\rangle = \begin{bmatrix} -0.5 \\ 0.866 \end{bmatrix} \quad |G_3\rangle = \begin{bmatrix} 0.866 \\ 0.5 \end{bmatrix}$$

Suppose we measure the particle's 2-color when it is in state $|R_1\rangle$ (that is, when its 1-color is definitely red). The observed outcome will be either red or green, and the new state of the particle will be either $|R_2\rangle$, meaning that the particle's 2-color is definitely red, or $|G_2\rangle$, meaning that its 2-color is definitely green. To calculate the probability of each outcome, we first compute the inner product of the starting and ending states as shown below:

$$\langle R_2 \,|\, R_1\rangle = \begin{bmatrix} 0.5 & 0.866 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0.5 \qquad \langle G_2 \,|\, R_1\rangle = \begin{bmatrix} -0.866 & 0.5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -0.866$$

The values 0.5 and -0.866 are called *amplitudes*; they are *not* the probabilities of obtaining $|R_2\rangle$ and $|G_2\rangle$. After all, probabilities can't be negative. However, they are closely related to the probabilities. To find the actual probabilities, we just square the magnitudes of the amplitudes: in this case, $|0.5|^2 = 0.25$ for $|R_2\rangle$ and $|-0.866|^2 = 0.75$ for $|G_2\rangle$. This means that when the right detector receives a particle having a 1-color of red, and then measures its 2-color, there is a $\frac{1}{4}$ chance that the particle's new 2-color will be red, and a $\frac{3}{4}$ chance that it will be green. Either way, as a result of the measurement the particle loses its 1-color and acquires a definite 2-color. If the detector flashes green as a result of the measurement, the new state of the particle will be $|G_2\rangle$; if it flashes red, the new state will be $|R_2\rangle$.

Interpreted geometrically, the inner product $\langle A \mid B\rangle$ can be thought of as a measure of the amount of similarity or "overlap" between vector $|A\rangle$ and vector $|B\rangle$, assuming that both $|A\rangle$ and $|B\rangle$ are of length 1. In that case, if $|A\rangle$ and $|B\rangle$ are identical vectors pointing in the same direction, the value of their inner product will be 1. If they are aligned but point in opposite directions, their inner product will be -1. And if they are perpendicular (orthogonal) to each other, their inner product will be 0, meaning that they are as dissimilar as possible, with no overlap at all. If there is partial overlap, their inner product will be a non-zero value somewhere between -1 and 1.