QM NOTES (SUPPLEMENT TO GILLESPIE) BY JIM & BILL (FEBRUARY 2015)

Review — using Dirac's notation

The standard QM was created in the mid-1920s by a handful of people whose names are now household words: Heisenberg, Born, Pauli, Schrödinger, and Dirac. Around 1929-30, the theory was put into its current systematic form by John von Neumann.

In Dirac's notation the state ψ is represented by a "ket" vector $|\psi\rangle$. The inner product of "bra" and ket vectors, a bracket, is $\langle\phi|\psi\rangle$. This corresponds to (ϕ, ψ) of the standard theory, used by Gillespie. If $\alpha_1, \alpha_2, \alpha_3, ...$ are eigenvectors, they are denoted $|\alpha_1\rangle$, $|\alpha_2\rangle$, and so on in Dirac's notation. The corresponding eigenvalues are a_1, a_2, a_3,... Sometime, instead of Greek letters, Roman letters are used for both eigenstates $|a_1\rangle$, $|a_2\rangle$, ... and eigenvalues, $a_1, a_2, ...$

We will repeat some of the postulates and theorems using Dirac's notation. We won't stick to the same numbering, but will (when we can) use the common name for a principle (e.g. the Born Rule). Since much QM literature (both physics and philosophy) is in the notation, it's important to be familiar with it. You can master it quickly.

Representation of States*: A physical system is represented by a vector $|\psi\rangle$ in a Hilbert space. All possible information is contained in $|\psi\rangle$.*

Hilbert Space*: $\mathcal{H}$, is a vector (or linear) space, possibly infinite dimensional. Like any vector space over the field of complex numbers, it satisfies the following rules (where $|\psi\rangle, |\psi_1\rangle, |\psi_2\rangle, ... \in \mathcal{H}$ and*

$c, c_1, c_2, ... \in \mathbb{C}$, the set of complex numbers).

$$\begin{aligned} |\psi_1\rangle + |\psi_2\rangle &= |\psi_2\rangle + |\psi_1\rangle \\ (|\psi_1\rangle + |\psi_2\rangle) + |\psi_3\rangle &= |\psi_1\rangle + (|\psi_2\rangle + |\psi_3\rangle) \\ c(|\psi_1\rangle + |\psi_2\rangle) &= c|\psi_2\rangle + c|\psi_1\rangle \\ (c_1 + c_2)|\psi\rangle &= c_1|\psi\rangle + c_2|\psi\rangle \\ (c_1 c_2)|\psi\rangle &= c_1(c_2|\psi\rangle) \end{aligned}$$

$\mathcal{H}$ is also an inner product space. The operation $\langle\,|\,\rangle$ on pairs of vectors defines a complex number. Thinking of this geometrically, the inner product is a measure of the "overlap" of different vectors; in the case of a single vector, the inner product gives the square of the norm or length of the vector. Here are the standard rules for inner products:

$$\begin{aligned} \langle\psi_1|\psi_2\rangle &= \langle\psi_2|\psi_1\rangle^* \text{ (called the complex conjugate)} \\ \langle\psi_1|\psi_2 + \psi_3\rangle &= \langle\psi_1|\psi_2\rangle + \langle\psi_1|\psi_3\rangle \\ \langle\psi_1|c\psi_2\rangle &= c\langle\psi_1|\psi_2\rangle \\ ||\psi||^2 &= \langle\psi|\psi\rangle \end{aligned}$$

A Hilbert space has two more important properties which we shall only mention. It is *separable* (meaning it has a countable basis) and it is *complete* (meaning if any sequence of vectors in the space converges to a vector ψ, then ψ is also in the space).

A quick review of complex numbers, here are the relevant facts.

$$\begin{aligned} c &= a + ib \;\; (\text{where } i = \sqrt{-1} \text{ and } a, b \in \mathbb{R}) \\ c^* &= a - ib \text{ (the complex conjugate of } c) \\ |c|^2 &= cc^* = a^2 + b^2 \;\; (\text{always a real number}) \\ e^{ix} &= \cos x + i \sin x \end{aligned}$$

A basis for a vector space is a set of vectors that are sufficient to express every vector in the space as a linear combination of basis vectors. We're particularly interested in a basis consisting of unit vectors that are orthogonal (i.e., at right angles) to one another.

$\{|\phi_i\rangle\}$ is an **orthonormal** set if and only if

$$\langle\phi_i|\phi_j\rangle = \delta_{ij} = \begin{cases} 1, \text{ if } i = j \\ 0, \text{ if } i \neq j \end{cases}$$

Any vector $|\psi\rangle$ can be expressed as a linear combination of basis vectors, that is, $|\psi\rangle = \sum_i c_i|\phi_i\rangle$. When the basis is an orthonormal set, we have the following important result.

$$\begin{aligned} \langle\phi_i|\psi\rangle &= \langle\phi_i| \sum\nolimits_j c_j|\phi_j\rangle \\ &= \sum\nolimits_i c_j\langle\phi_i|\phi_j\rangle \\ &= \sum\nolimits_j c_j\delta_{ij} \\ &= c_i \end{aligned}$$

Theorem: $\langle\psi|\psi\rangle = \sum_i |c_i|^2$

We've been talking rather abstractly about the vectors in Hilbert space. A simple example (in one dimension) might be useful here. Let $|\psi\rangle = \cos kx + i \sin kx = e^{ikx}$. The inner product is defined

$$\langle\psi_1|\psi_2\rangle = \int_{-\infty}^{\infty} \psi_2^* \psi_1 \, dx$$

Representation of properties (observables)*: Observables,* $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$*,... (e.g., position, momentum, energy, or more precisely position in the x-direction, momentum in the y-direction, spin in the z-direction, etc.) are represented by Hermitian operators,* $\mathbf{A}, \mathbf{B}, \mathbf{C}$*, ... each with a complete set of eigenvectors,* $|a_1\rangle, |a_2\rangle, ..., |b_1\rangle, |b_2\rangle, ...$*, and corresponding eigenvalues,* $a_1, a_2, ..., b_1, b_2, ...$

Like the first principle, this one also tells us how to represent the physical world mathematically. The term "observable" is, unfortunately, standard in QM. All it means is "property," but Heisenberg, who introduced the term, was perhaps unduly influenced by a kind of positivism.

The state of a quantum system such as an electron is represented by a vector in a Hilbert space. The electron has properties such as:

position in the z-direction, momentum in the y-direction, and so on. And a particular property will have a specific magnitude. That is, the property of position in the x-direction will have the magnitude q_1, or q_2, or $q_3, ...$; and the property spin in the z-direction will have the magnitude $\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$, often called $+$ or $-$, or simply *up* or *down*. Once again, it's important to distinguish properties, $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$, which are part of the physical world, from their mathematical representation by operators, **A, B, C,** and their magnitudes by eigenvalues. The term "Hermitian" means that the eigenvalues are real numbers. We'll now explain these notions.

A **linear operator**, $\mathbf{O}$, is a function on a Hilbert space, $\mathbf{O} : \mathcal{H} \rightarrow \mathcal{H}$,which satisfies the following conditions.

$$\begin{aligned}
(c\mathbf{O})|\psi\rangle &= c(\mathbf{O}|\psi\rangle) \\
\mathbf{O}(|\psi_1\rangle + |\psi_2\rangle) &= \mathbf{O}|\psi_1\rangle + \mathbf{O}|\psi_2\rangle \\
(\mathbf{O}_1 + \mathbf{O}_2)|\psi\rangle &= \mathbf{O}_1|\psi\rangle + \mathbf{O}_2|\psi\rangle \\
(\mathbf{O}_1\mathbf{O}_2)|\psi\rangle &= \mathbf{O}_1(\mathbf{O}_2)|\psi\rangle)
\end{aligned}$$

Measurement: *The measurement of a physical system results in eigenvalues only; i.e., a_1, a_2, a_3, ... will be the measurement outcomes; they correspond to the eigenvectors $|a_1\rangle, |a_2\rangle, |a_3\rangle, ...$*

Expectation Value: *The expectation value for some observable is $\langle A\rangle = \langle\psi|A|\psi\rangle$.*

Born Rule*: If a system is in state $|\psi\rangle$ and $\mathfrak{A}$ is measured, then the probability of getting the result $a_i = |\langle a_i|\psi\rangle|^2 = |a_i|^2$.*

Heisenberg's Uncertainty/Indeterminacy Principle*: When* $\mathbf{A}$ *and* $\mathbf{B}$ *do not commute (i.e.,* $\mathbf{AB} - \mathbf{BA} = c \neq 0$*), then* $\Delta\mathbf{A}\Delta\mathbf{B} \geq c/2$*. (*$\Delta\mathbf{A}$ *is the root-mean-square deviation, roughly the spread of eigenvalues of measurements on identically prepared systems.) A useful piece of notation is the* ***commutator*** $[\mathbf{A}, \mathbf{B}]$*, defined to equal* $\mathbf{AB} - \mathbf{BA}$*.*

Temporal evolution*: The state of a system evolves in time according to the Schrödinger equation:* $\mathbf{H}|\psi\rangle = i\hbar\, d\,|\psi\rangle/dt$. $\mathbf{H}$ *is the Hamiltonian (or energy) operator (in one dimension):* $\mathbf{H} = -\frac{\hbar}{2m}\frac{d^2}{dx^2} + V(x)$.

Projection Postulate*: If a measurement of observable* $\mathfrak{A}$ *results in eigenvalue* a_i *then immediately after the measurement, the system is in the eigenstate* $|a_i\rangle$.

Spin

An electron (and other particles) can have a position, momentum, energy, angular momentum, and so on. These properties or observables have corresponding operators with eigenvectors and eigenvalues. One of the most important properties of an electron is called spin. It is tempting to think of electron spin as like the rotation of the earth around its axis. This is a mistake. For instance, it does not depend on any coordinate system in which it rotates. There is no way to conceptualize spin that makes it analogous to anything we comprehend. All we can do is specify an operator, eigenvectors and eigenvalues.

For a moment, set aside what we just said about imagining spin and pretend that you can. An electron would then be a spinning charged particle. This gives rise to a magnetic moment, which could be detected. A so-called Stern-Gerlach apparatus produces an irregular magnetic field. If an electron passes through, it will deviate from a straight path either upward or downward, depending on the direction of its spin. Since this is a quantum phenomenon, it will be no surprise that the result is not spread over a large range of outcomes, but is rather confined to two quite specific results. They are called "up" and "down" for short; they are the measured eigenvalues. We will now briefly explain the spin states, and the eigenvectors and eigenvalues of the spin operator.

We begin with what are known as the Pauli spin matrices.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The linear operators that represent spin (in the x-direction, the y-direction, and the z-direction, respectively), are defined as follows:

$$S_x = \frac{1}{2}\hbar\sigma_x, S_y = \frac{1}{2}\hbar\sigma_y, S_z = \frac{1}{2}\hbar\sigma_z,$$

The eigenvalue equation allows us to find the eigenvectors and eigenvalues. There are two eigenvectors for each operator, an α and a β.

$$\begin{array}{cc} \alpha_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} & \beta_x = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ \alpha_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} & \beta_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} \\ \alpha_z = \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \beta_z = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{array}$$

The corresponding eigenvalues are, $\frac{1}{2}$ and $-\frac{1}{2}$, associated with each pair of eigenvectors. The $\frac{1}{2}$ term is why it is called spin-half. In any given direction, the spin can have the value $\frac{1}{2}$ or $-\frac{1}{2}$. No other value is possible. (For convenience, the two eigenvalues are sometimes called 1 and 0, or + and –, or $\uparrow$ and $\downarrow$, or simply *up* and *down*.) Using this formalism, we can make predictions about how electrons behave in a Stern-Gerlach apparatus. If a beam of electrons passes through this device, then, generally, two beams will come out, corresponding to the two eigenvalues, $\frac{1}{2}$ and $-\frac{1}{2}$. We can call the two beams the upper and the lower, respectively (spin up and spin down). A single electron will have a one half chance of coming out in either beam. When we know the initial state, we can make specific predictions. For instance, if the electrons in the beam are all in state $|\psi\rangle = |\alpha_z\rangle$, and we decide to measure the spin in the z-direction, then we will certainly find each electron to have eigenvalue $\frac{1}{2}$, i.e., each will be in the upper beam. How do we know this? We apply the operator $S_z = \frac{1}{2}\hbar\sigma_z$ to the state $|\psi\rangle = |\alpha_z\rangle$. The calculation is:

$$S_z\alpha_z = \frac{1}{2}\hbar\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\hbar\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

The result of the operator acting on the vector is multiplying that vector by a number, namely, $\frac{1}{2}\hbar$, so that number is the eigenvalue. If we wanted to measure the spin component in, say, the $x-$ or $y-$direction, or more generally, along any angle θ, then we would use the S_θ operator. If the Stern-Gerlach apparatus is oriented at angle θ in the $z-x$ plane, then the operator S_θ is defined:

$$S_\theta = S_z cos\theta + sin\theta = \frac{1}{2}\hbar \begin{pmatrix} cos\theta & sin\theta \\ sin\theta & -cos\theta \end{pmatrix}$$

The eigenvalues of this operator are exactly the same as the others, $\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$. This matrix will also allow us to calculate the probabilities of getting one or the other of these eigenvalues.

Fermions (electrons, protons, quarks) have spin $\frac{1}{2}$; more generally, fermions have half-integer spin: $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, ...$ Bosons (photons, mesons, gravitons, the Higgs particle) have whole number spins: $0, 1, 2, 3, ...$; thus, the Higgs particle has spin 0, photons have spin 1, and gravitons are spin 2. The thing that makes either electrons and photons ideal in foundational discussions is that they are two-state systems: they have two eigenstates, up and down, so their Hilbert spaces are only two-dimensional, making them relatively easy to use.

Exercise 1: Show that $\alpha_x = \frac{1}{\sqrt{2}}\alpha_z + \frac{1}{\sqrt{2}}\beta_z$.

Tensor Products

Suppose $\mathcal{H}_1$ is Hilbert space with dimension n and $\mathcal{H}_2$, is another space of dimension m. Then we can form a vector space $\mathcal{H}_1 \otimes \mathcal{H}_2$, called the *tensor product*, which is also a Hilbert space. It has dimension $n \times m$.

If $|\alpha\rangle \in \mathcal{H}_1$ and $|\beta\rangle \in \mathcal{H}_2$ then $|\alpha\rangle \otimes |\beta\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$, that is, it is also a vector in a Hilbert space. Sometimes this notation is abbreviated to $|\alpha\rangle\,|\beta\rangle$ or, more simply, to $|\alpha\beta\rangle$.

Similarly, if $\mathbf{A}$ and $\mathbf{B}$ are operators, then we can form their tensor product, $\mathbf{A} \otimes \mathbf{B}$. Such an operator works on vectors as follows:

$$(\mathbf{A} \otimes \mathbf{B})\,|\alpha\rangle \otimes |\beta\rangle = \mathbf{A}\,|\alpha\rangle \otimes \mathbf{B}\,|\beta\rangle$$

Composite Systems*: After an interaction, two systems,* 1 *and* 2*, with their respective states in Hilbert spaces* $\mathcal{H}_1$ *and* $\mathcal{H}_2$*, are represented by the state* $|\psi_{12}\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ *which is in the tensor product space* $\mathcal{H}_1 \otimes \mathcal{H}_2$*.*

The Inner Product. The definition of the inner product in the tensor product space is pretty much what you would expect:

$$\langle |\psi_1\rangle \otimes |\phi_1\rangle \,|\, |\psi_2\rangle \otimes |\phi_2\rangle \rangle = \langle \psi_1|\psi_2\rangle \times \langle \phi_1|\phi_2\rangle$$

Exercise 2: (a) The two slit experiment. Let the state of a particle be the superposition of passing through the Left slit and passing through the Right slit:

$$\psi = \frac{1}{\sqrt{2}}(|\psi_L\rangle + |\psi_R\rangle)$$

If Q is an operator of position observable with eigenvalue q_k show that the probability of finding the particle at position q_k contains cross (or interference) terms in addition to terms corresponding to the particle travelling through each slit.

(b) Now consider a detector able to tell which slit the particle passed through. The detector has two states, R and L. The detector states are exclusive (it is a good detector) which we represent by requiring that R and L be orthogonal. And the detector is accurate, so that if the particle is in ψ_L then the detector is in L, and similarly for ψ_R. The state of the joint system can then be represented as

$$\psi_D = \frac{1}{\sqrt{2}}(|\psi_L\rangle \otimes |L\rangle + |\psi_R\rangle \otimes |R\rangle)$$

The position operator, Q, must now act on the joint system but of course we are not interested in measuring the detectors. We thus represent the operator as $Q \otimes I$ where I is the identity operator (i.e. for any ψ, $I\psi = \psi$). The joint operator works on joint states as defined above. Given this setup, show that the cross terms vanish or, in other

words, that if we put a which-path detector into the experimental arrangement, the interference effects disappear.

Illustrations

The tensor product space way of representing joint systems allows us to illustrate some important properties of quantum systems.

(a) Entanglement. The singlet state will be discussed below. It is a joint state of two systems (e.g. electrons) which cannot be understood as composed of two independent systems (it cannot be 'factored'). It can be written in general as

$$\psi_s = \frac{1}{\sqrt{2}} \left(|\psi_1\rangle \otimes |\phi_2\rangle - |\psi_2\rangle \otimes |\phi_1\rangle \right)$$

In this state the components have lost their independence. For example, if we find the system's first component is in the state ψ_1 then we must find the second component to be in ϕ_2. This property of the system is independent of the size of the joint system. The correlations evidently exist no matter how far apart the components might be. By contrast, the joint state

$$\psi_u = \left(\alpha |\psi_1\rangle + \beta |\psi_2\rangle \right) \otimes \left(\gamma |\phi_1\rangle + \delta |\phi_2\rangle \right)$$

is not entangled. The ϕ states are not 'determined' by the ψ states.

(b) No-Cloning. It would be extremely useful if we could clone quantum states, that is, if we had a way to reproduce an arbitrary quantum state without destroying the original. It turns out that this is unfortunately impossible. A sketch of a proof is possible here. Recall that quantum states evolve according to the Schrödinger equation in a deterministic manner which can be expressed in terms of the time evolution operator U_t, where $U_t |\psi\rangle = e^{-iHt/\hbar} |\psi\rangle$ (H being the Hamiltonian operator for the system in question). The crucial property of the time evolution operator is that is a unitary operator, which is defined to be an operator which preserves inner products (plus meeting the condition that for every vector of the Hilbert Space, ϕ there is an vector, ψ, such that $U(\psi) = \phi$).

Exercise 3: Show that U_t preserves the inner product. That is, for any two vectors, ψ and ϕ, show that $\langle\psi|\phi\rangle = \langle U_t\psi|U_t\phi\rangle$ (hint: you could try to use Gillespie's Eqs 4-42 and 4-43).

What would cloning look like. We would begin with our source state and some system, call it the target, which we will transform into a copy of the source using some time evolution operator (a measurement operator obviously won't work in general because of the uncontrollable change it might induce in the source and/or the target). We have no idea what the special operator might look like, but it must meet these conditions

$$\psi \otimes \tau \Longrightarrow \psi \otimes \psi$$

where τ is the target and $\Longrightarrow$ represents our unitary cloning device. Our cloning machine is supposed to work on any source vector. So consider that this pair will be true

$$\psi \otimes \tau \Longrightarrow \psi \otimes \psi$$

$$\phi \otimes \tau \Longrightarrow \phi \otimes \phi$$

Now think about this inner product equation (being loose with the notation for simplicity):

$$\langle\psi|\phi\rangle = \langle\psi \otimes \tau|\phi \otimes \tau\rangle$$

It is true because of the rule for inner products of tensor product states along with the fact that our states are normalized ($\langle\psi|\psi\rangle = 1$ for all states). But then

$$\langle\psi \otimes \tau|\phi \otimes \tau\rangle = \langle\psi \otimes \psi|\phi \otimes \phi\rangle$$

This is true because of the action of the cloning machine along with the fact that unitary time evolution preserves the inner product. But then

$$\langle\psi \otimes \psi|\phi \otimes \phi\rangle = \langle\psi|\phi\rangle\langle\psi|\phi\rangle = (\langle\psi|\phi\rangle)^2$$

So we have

$$\langle\psi|\phi\rangle = (\langle\psi|\phi\rangle)^2$$

This is possible only if $\langle\psi|\phi\rangle = 1$ or $\langle\psi|\phi\rangle = 0$. That is, it is only possible in two special cases, the one where $\psi = c\phi$ (so they are essentially the same state) or the two states are orthogonal. So it is impossible *in general* to clone a quantum state. If one could clone states in general, then it would be possible to use entanglement to send signals faster than the speed of light.

Exercise 4: show how to use a cloning machine and entanglement to send superluminal signals (hint: use Exercise 1 above).

Interpretations

The phrase "interpretation of the QM formalism" is somewhat vague. Logicians have a precise notion of "interpretation" or "model of a formal system," but that won't do here. To start with, the formalism is already partially interpreted; it is hooked to observational input and output in a fairly clear and unambiguous way. Some philosophers and physicists insist that this is enough and that we should not seek more. Feynman, for example, held that no one understands QM or possibly could understand it.[1] Instead of getting ourselves in a hopeless knot trying to do so, we should just learn the quantum rules and apply them in various situations — don't look for anything deeper. This partial interpretation is often called the minimal statistical interpretation. What it can do is handle everything observable. It is often favoured by those who advocate an instrumentalist outlook for scientific theories in general. But our interest is with how the world really works, not just with making successful observable predictions. Only those lacking a soul are content with the minimal statistical interpretation.[2] What's needed is something over and above this instrumentally adequate, but otherwise incomplete account.

[1] Feynman, *Lectures*, vol III

[2] Of course this is quite unfair. Consider the remark merely a bit of gentle polemics. Many who hold the minimal view feel forced into it by the repeated failures of realistic approaches to QM. This is not an unreasonable view, though I do think it mistaken.

The spectrum of possible interpretations is exceedingly broad. On the one side it may be as trivial as so-called hidden variables. (That is, trivial in a philosophical or conceptual sense, since a hidden variable interpretation tries to make the quantum world out to be as much like the classical world as possible. Such interpretations are not trivial in the technical sense — indeed, they might even be impossible.) On the other hand, the range of possible interpretations may be limited only by our imaginations. The empirical consequences of rival views are largely unknown. It was long thought that the realism involved in the EPR thought experiment made no observable difference — hence it was often branded "idle metaphysics." But J.S. Bell, to the surprise of everyone, derived an empirically testable consequence. One could even imagine consequences of the right interpretation of QM being as significant as special relativity, for it is quite plausible to see Einstein's theory as an *interpretation* of Maxwell's electrodynamics.

We'll broadly consider two types. *Realist* interpretations of QM typically hold that

- The quantum world exists independently of us; we do not create it in any way.
- Quantum systems have all their properties all the time.
- Measurements discover those properties, they do not create them.

On the other hand, *anti-realist* views, such as the Copenhagen interpretation, hold that

- The quantum world is not independent of us.
- In some important sense observers make reality.
- Measurements create their results.

The minimal statistical interpretation is a kind of anti-realism, too, but it's important to note the difference. Scientific realism typically involves at least two ingredients — epistemic and ontological. The first says that we can have rational beliefs about a realm of unobservable entities. Opposition to this aspect of realism is a form of scepticism. Duhem and van Fraassen, for example, both share this form

of anti-realism with the minimal statistical interpretation of QM. On the other hand, neither Kant nor Bohr are sceptics about unobservable entities, but, with Kuhn, pragmatists, and verificationists, they reject the second aspect of realism which holds that the world we are trying to learn about exists independently of us. These are sometimes subtle distinctions; we'll have to develop them carefully.

EPR

The anti-realism of the Copenhagen interpretation was met head-on by the beautiful EPR thought experiment. The argument proceeds by first characterizing some key notions.

> *Completeness:* A theory is complete if and only if every element of reality has a counterpart in the theory. Thus, if an electron, for example, has both a position and a momentum, but the theory only assigns a value to one and not the other, then that theory is incomplete.
>
> *Criterion of Reality:* If, without disturbing the system, we can predict with probability one the value of a physical magnitude, then there is an element of reality corresponding to the magnitude. The qualification – without disturbing the system – is central. The Copenhagen interpretation holds that measurements do disturb the system (they collapse the wave function), so ascribing an independent reality to any magnitude cannot be naively based on the outcome of a (direct) measurement.
>
> *Locality:* Two events that are space-like separated (i.e., outside each other's light cones) have no causal influence on one another. They are independent events. This follows from special relativity which holds that nothing, including causal connections, travels faster than light.

The EPR argument (we will follow the more perspicuous version of David Bohm, which is now standard) starts with a system such as an energetic particle which decays into a pair of photons; these travel in opposite directions along the z-axis. Each photon, call them L and R (for left and right), is associated with its own Hilbert space. The polarization or spin eigenstates will be along any pair of orthogonal axes, say, x and y, or x' and y'. In any given direction a measurement (which, recall, only yields eigenvalues) will result in either a $+1$ for the spin-up state or a -1 for the spin-down state. We can represent these as $|+\rangle_L$ and $|-\rangle_L$, respectively, for the L photon, and $|+\rangle_R$ and $|-\rangle_R$ for R.

The spin of the system is zero to start with and this must be conserved in the process. Thus, if L has spin magnitude $+1$ in the x-direction then R must have -1 in the same direction to keep the total equal to zero. A composite system such as this (in the so-called singlet state), is represented by

$$|\psi_{LR}\rangle = \frac{1}{\sqrt{2}}(|+\rangle_L \otimes |-\rangle_R - |-\rangle_L \otimes |+\rangle_R)$$

If we measure the spin of the L photon we then know the state of R, since the measurement of $|\psi_{LR}\rangle$ immediately puts the whole system into one or other of the two eigenstates. Suppose our measurement resulted in L being polarized in the x-direction (i.e., has spin up in the x-direction). This means the state of the whole system is $|+\rangle_L \otimes |-\rangle_R$, from which it follows that the remote photon is in state $|-\rangle_R$, i.e., it has spin down in the x-direction. (Choosing the x-direction is wholly arbitrary; any other direction could have been tested for.) While it might be conceded that the measurement on L may have disturbed it or created rather than discovered the measurement result, the same cannot be said of R. We are able to predict with complete certainty the outcome of the distant measurement on the R photon, and since (by the locality principle) we could not have influenced it in any way with a measurement of L, it follows (by the criterion of reality) that the R magnitude exists independently of measurement. Since this is not

reflected in $|\psi\rangle$, it follows (by the criterion of completeness) that QM does not completely describe the whole of reality. EPR then concludes (though this is a bit vague) that the theory must be supplemented with hidden variables in order to give a full description.

The Bell Results

Compelling though EPR is, it can't be right. This is the upshot of several related findings known collectively as the Bell results. John Stuart Bell's original argument was rather complicated, but versions are now so simple that those with only elementary algebra can easily comprehend the argument. We'll begin with a simple derivation of a Bell-type inequality (due to Eberhard 1977), then briefly describe its experimental refutation.

Let us begin by considering an EPR-type set-up. Unlike EPR, however, we will consider measurements of spin in different directions, say along a and a' for the L photon and b and b' for R. (See Fig. 4) There are four possible measurements that could be made:

$$(a, b), (a', b), (a, b'), (a', b')$$

(where (a, b) means the L photon is measured for spin along the a-direction and R along the b-direction). A spin up result of a measurement has value $+1$, and spin down -1. Now define a correlation function, $c(x, y)$ as follows:

If $a = 1$ and $b = 1$, then $c(a, b) = 1 \times 1 = 1$

If $a = 1$ and $b = -1$, then $c(a, b) = 1 \times -1 = -1$

and so on for a', b',etc.

(where $a = 1$ means that the result of measuring the L photon in the a direction is $+1$, etc.)

We imagine running the experiment many times. After N tests, with a_i being the i^{th} result, we have

$$c(a, b) = \frac{1}{N} \sum\nolimits_i a_i b_i.$$

We will make two key assumptions.

> *Realism:* Each photon has all of its properties all of the time; in particular, each has a spin up or spin down magnitude in every direction whether there is a spin measurement made in that direction or not.

This assumption is embedded in the mathematics as follows. Let a_i (or a'_i, b_i, b'_i, respectively) be the result of the i^{th} measurement, if made in the a (or a', b, b', respectively) direction. The value is either $+1$ or -1 and this value exists whether a measurement is made or not. In particular, if photon L is measured in the a direction then it cannot be measured in the a' direction. Nevertheless, even thought we can't know what the value is, we still assume that it has one value or the other. This is the core of realism — measurements do not create, they discover what is independently there.

> *Locality:* The results of measurement on one side of the apparatus do not depend on what is happening at the other side. The outcome of a spin measurement on photon L is independent of the direction in which R is measured (i.e., the orientation of the apparatus); it is independent of the outcome of that measurement; and it is independent of whether R is measured at all.

Formally, the locality assumption is captured by having the value of a_i be independent of the values of b_i and b'_i. So if a measurement of L in the a direction would result in $+1$ if R were measured in the b-direction, it would still be $+1$ if R were measured in the b'-direction instead. Recall that Bohr holds that a micro-entity has its properties only in relation to a macro-measuring device — different settings may create different micro-properties. Locality does not completely deny this, but it does deny that the settings of a remote macro-device have any influence that could be transmitted faster than the speed of light.

Now define the following formula which I'll call F for convenience:

$$F = a_i b_i + a_i b'_i + a'_i b_i - a'_i b'_i$$

Rearranging terms we have

$$F = a_i(b_i + b'_i) + a'_i(b_i - b'_i)$$

Since the a terms equal $+1$ or -1, and since one of the terms in parentheses equals 0 while the other equals either $+2$ or -2, we have

$$F = +2 \text{ or } -2$$

Thus, taking the absolute value, we have

$$|a_i b_i + a_i b'_i + a'_i b_i - a'_i b'_i| = 2$$

This holds for the i^{th} measurement result. The generalization for N measurements is therefore

$$|\frac{1}{N}\sum_i a_i b_i + a_i b'_i + a'_i b_i - a'_i b'_i| \leq 2$$

In terms of the correlation function we have

$$|c(a,b) + c(a,b') + c(a',b) - c(a',b')| \leq 2$$

This is one form of Bell's inequality. It means that when spin measurements are done for arbitrary directions a and a' on the L photons and b and b' on the R photons, we can expect this degree of correlation. After many tests the correlations between the L and R photons, taken a pair at a time, must satisfy this inequality — if the assumptions of realism and locality both hold.

It is important to stress that the inequality is derived by a simple combinatorial argument based on two common sense assumptions, realism and locality. QM, however, makes a different prediction. An experimental test of QM and Local Realism (as it is often called) is thus possible.

To get specific QM predictions we need to specify directions for the spin measurements to be made. Let $a = b$, otherwise the orientations of a and b can be arbitrary; furthermore, let a' be -45 degrees and b' be $+45$ degrees from the common a/b direction.

According to QM the correlation functions for the case of particles in the singlet state have the following values:

$$\begin{aligned} c(a,b) &= -\cos 0 = -1 \\ c(a,b') &= -\cos 45 = -\frac{1}{\sqrt{2}} \\ c(a',b) &= -\cos 45 = -\frac{1}{\sqrt{2}} \\ c(a',b') &= -\cos 90 = 0 \end{aligned}$$

What this means is that if L is measured in the a direction and has, say spin up, then R measured in the b $(= a)$ direction will not have spin up. They are perfectly negatively correlated. In the fourth case immediately above when L and R are measured at right angles to each other, the results of measurement are completely uncorrelated. The other two cases yield results in between.

We now put these values derived from QM into the same form as the Bell inequality:

$$|-1-\frac{1}{\sqrt{2}}-\frac{1}{\sqrt{2}}-0| = 1+\frac{2}{\sqrt{2}} > 2$$

Thus, at these angles, QM and Local Realism diverge in their predictions, making an empirical test possible[3]. Just how remarkable this situation is cannot be stressed too much. For years EPR was attacked by the empiricist-minded for being "idle metaphysics," since it was thought that it made no detectable difference. Even defenders of EPR were willing to concede that the realism/anti-realism debate has no empirical import. Now it turns out that all were wrong. Abner Shimony calls it "experimental metaphysics", and the phrase is exactly right. From the original EPR argument to Bell's derivation of his inequality, to the experimental tests, to the reaction to those tests, there is an inextricable

[3]It can be shown that the quantum correlations are bounded at the value $2 \times \sqrt{2}$ which is not the theoretical maximum imaginable which could be 4 if the terms of $F = a_i b_i + a_i b'_i + a'_i b_i - a'_i b'_i$ have maximum values (look up Tsirelson bound ...). It has been conjectured that this limitation has something to do with information being the basis of reality!?

mix of physics and metaphysics. The whole situation is highly reminiscent of the 17^{th} century when philosophy and physics were intertwined and at their best.

If the possibility of performing an empirical test on realism was surprising, the outcome was even more surprising — common sense has taken a beating. There have been several tests of the inequality. In almost every one, QM has made the right predictions and Local Realism the wrong ones. Of all these tests, the ones carried out by Aspect *et. al.* (1981, 1982a, 1982b) have been the most sophisticated.

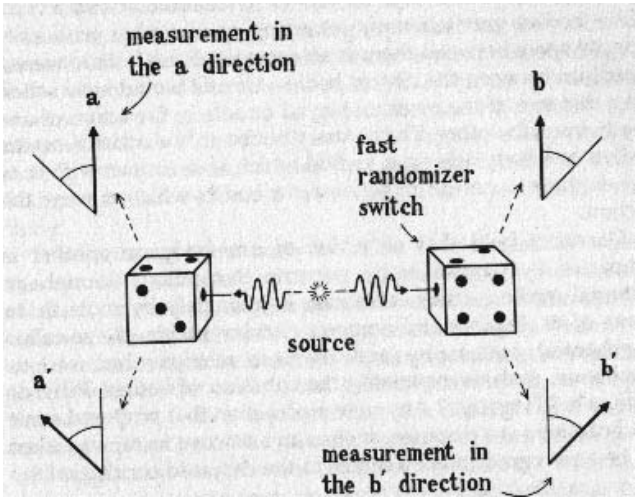

Fig. Schematic version of the Aspect experiment.

The crucial feature of the Aspect experiment is the presence of a very fast optical switch which directs L photons to either a or a' and R photons to either b or b' measurements. It picks a direction randomly, *while the photon is in flight.* The reason this is considered important is that in earlier experiments the setting of the distant measuring device was fixed long before the measurement, thus allowing the possibility of a subluminal signal between the distant wings of the apparatus and hence the possibility that they could "communicate" with one another. Of course, that may seem bizarre, but the QM world is so weird that it is always nice to have one more possibility ruled out, however farfetched it may seem to common sense.

References

Alain Aspect, Philippe Grangier, Gérard Roger (1981), 'Experimental Tests of Realistic Local Theories via Bell's Theorem', Phys. Rev. Lett. 47 (7): 460–3.

A. Aspect, Dalibard, Roger: (1982a) 'Experimental test of Bell's inequalities using time-varying analyzers' Phys. Rev. Letters 49 #25, 1804.

A. Aspect, P. Grangier, G. Roger (1982b) 'Experimental realization of Einstein-Podolsky-Rosen-Bohm gedanken experiment; a new violation of Bell's inequalities' Phys. Rev. Letters 49 #2, 91.

P. H. Eberhard (1977), 'Bell's theorem without hidden variables', *Il Nuovo Cimento*, 38B 1, 75—80.