Introduction to Quantum Mechanics and Quantum Computing

1. Abstract

Quantum Mechanics (QM) is the underlying physical theory of the very small; it has at least three highly bizarre features, including a notion of parallel worlds. Nevertheless, it is the direct source of the most exacting and precise measurements and predictions ever made, and is the basis for much of our understanding of the world around us.

Quantum Computing (QC) is the use of QM to perform (a bizarre kind of) parallel computation.

2. Light

We begin with light, which has for centuries been at the forefront of much thinking about the nature of the universe.

What is light? Answers, or parts of answers, have been given by the following thinkers, among others: Newton, Huygens, Maxwell, Einstein, Schrodinger. In the 17th century, Newton and Huygens proposed radically different ideas about the nature of light. Newton argued that it consists of tiny particles that speed about; but Huygens' view was that light is a kind of wave, not unlike a water wave. This latter view was the accepted one for over two centuries, culminating in Maxwell's astonishing mathematical demonstration in the 19th century (i) that electricity and magnetism are two aspects of the same thing: "electromagnetic waves" – this built on earlier results of others such as Faraday – and (ii) that the speed of such waves is precisely the speed of light, providing dramatic evidence that light too is an electromagnetic wave. It seemed conclusive in favor of Huygens, and indeed some physicists at the time stated that physics was pretty nearly over as an active science since all major questions were settled: Newton had explained motion and gravity ($F = ma$ etc) and Maxwell had explained electricity, magnetism, and light.

But it was not to be. The early 20th century found glaring gaps in these explanations. What had seemed like minor difficulties turned out to be major ones, and wholly new conceptual frameworks were in store.

Some of the problems can easily be illustrated:

2.1. Single spots of light.

If we prepare a room with only one light source (eg a lightbulb) and dim the light more and more, eventually only tiny "bits" of light will be produced even in a long time period, eg one hour. More specifically, the dimness can be arranged so that a highly sensitive photographic film placed in that room will, after one hour, reveal only one or two tiny spots where it has been exposed to light. This means that, whatever light is, it arrives at the film in very tiny pieces. If we dim the bulb even more, the spots do not get dimmer, they simply become less frequent; any one spot remains just as bright as always, but may occur once a day, or once a year, instead of once an hour. It is as if light comes in little bundles and cannot be further reduced. For now, these pieces themselves may or may not best be thought of as little bullets (particles) or little waves. Let's use a neutral term: light bundles.

2.2. Polarized light.

At any photography store one can buy polaroid filters. These serve to polarize light. From Maxwell, we know that light has both electric and magnetic parts, called fields, that vary in space and time, in a wavelike manner. These two fields are perpendicular to one another, eg one might "wave" vertically as it moves forward, and the other horizontally, the first like a jumprope being waved up and down, and the other like a snake slithering through the grass. A polaroid filter has very fine parallel lines that are sensitive to the electric field, so that any light "bundles" that pass through it must come out with their electic fields parallel to those lines; the light that passes through is said to be polarized, eg in the vertical direction. (The filter does not directly affect the magnetic field, which simply goes along with whatever the electric field does.) Also, we can rotate the filter, and thus make it a horizontal polarizer, a 45-degree slanted polarizer, etc. Here is a diagram of light passing through a vertical polarizer:

```
                             -----
ordinary randomly           |  ^  | vertically polarized
polarized light  ===>       |  |  |  ===>  light (50% of the light gets through)
                            |  v  |
                             -----
                        vertical filter
```

If we now place a second filter after the first, also vertically polarizing, all the vertically polarized light coming out of the first filter passes through the second as well, and remains unchanged (vertically polarized). On the other hand, if we rotate the second filter 90-degrees to make it a horizontal polarizer, no light from the first filter passes through the second at all. This does not surprising at this point.

What does seem surprising is the amount of light that gets through the first polarizer (whether or not there are other polarizers after it): fully 50% of the light that comes in on the left comes out on the right. This seems odd: for that incoming light is random, eg just from lightbulbs overhead, and presumably oriented in all possible directions. So it would seem that only a tiny fraction should pass through: those few bundles that happened to have their electric fields already vertical. Maybe something special about the lightbulbs has polarized 50% of them in advance? But if we rotate the filter, by any angle, we still get 50% coming through.

It seems clear that the filter is somehow changing the incoming light, polarizing 50% of it vertically. What is more, the light that does not get through not only bounces off the filter but itself is polarized in the horizontal direction! So all of the original light is affected by the filter and ends up in one of two states: vertical or horizontal. But then why does no vertical light get through a horizontal filter? There seems to be a pattern of sensitivity to the direction of polatization of the incoming light. This then might explain the 50% result for random polarization: if lots of light bundles come in with polarizations in lots of more or less uniformly varied directions, then some get through because they already have the right (vertical) polarization, some do not get through at all (horizontal) and most get through with varying probabilities.

There is a second somewhat curious feature: if we again place a second filter after the first, but with a 45-degree rotation, 50% of the light that passes the first filter will succeed in passing the second, ie 25% of the original light gets through both. And if we then place a third horizontal filter after these, 50% of the light coming to it gets through, ie 12.5%. But if we already agree that filters can change the polarization of incoming light, then this is not so surprising, except that we do not have an explanation yet for why the particular amounts pass through – that will come later, and turns out to be a far more general issue than that of light and filters.

Finally, if we again do the dimming trick, we still get the same results: half the bundles get through the first filter, etc, except now we have to wait a long time. In any one hour there is only one bundle, and it either gets through or it doesn't, and on average 50% go each way; and each ends up polarized vertically if it passes, and horizontally if it doesn't.

2.3. **Diffracted Light.** The above experiments could not be done in the 17th century because of the lack of filters and light dimmers and photographic film. But the following one could have – and in essence it was.

If light coming from the left falls on a wall with one small hole and then whatever light gets through the hole falls on a second wall, one sees on the second wall (if there is no other source of light in the room) not just a single spot of light but rather a pattern of concentric light and dark regions, with the brightest being directly opposite the hole:

```
                           |                               *|  dark
                           |                             ***| bright
=====>                     |                               *|  dark
=====>                   hole                          *****|brightest
=====>                     |                               *|  dark
=====>                     |                             ***| bright
                           |                               *|  dark
```

It is a bit surprising that one gets dark areas and then lighter ones further from the center. But if we suppose light is a wave something like a water wave, and the first wall is a seawall that protects the beach (second wall) from the waves, then a hole in the seawall will in fact produce a series of smaller waves at the hole, that spreads out and hits the beach with points of more intensity and others of less intensity, just as in the above diagram. In fact, this can be explained mathematically in terms of peaks and troughs in the spreading waves; the math is simple trigonometry.

So it seems that the wave-theory wins out here. In fact, it is even more dramatic. If we now put a second hole in the wall, close to the first, we get an even more surprising pattern, where some regions that had been bright are now dark! It is hard to see how, if light were little particles, opening another hole could prevent particles from hitting a place on the second wall where they had been going before. But the wave theory explains this in terms of more troughs and peaks, that can actually interfere with one another, just as with water waves.

But, if we now start to dim the light (this won't work with water! – and it was not do-able in the 17th century) we get a different surprise: we can dim the light as before so that only one light "bundle" arrives every hour – and we can check this by replacing the second wall with a photographic film. Now we only get one individual dot of brightness per hour, always the same intensity, but at lots of places as the hours go by. And if we graph all the results, we get back the exact same picture as above! So now what are we to think? The bundles seem very different from water waves, they do not appear to spread out, they do not have different intensities at different places. And yet, taken all together, they have a group behavior very much like water – even when they are an hour apart from each other! Some scientists have speculated that each bundle does manage to spread out somehow, in fact splits and goes through each of the two holes and then interferes with itself (or its twin) as it (or they) get back together again upon hitting the second wall. This is pretty wild stuff, and no one really knows. But the mathematics that has been developed in quantum mechanics is a kind of description that sounds very much like this speculation, and that very accurately allows prediction of what is observed, not only in this case but also for polarized light, for electrons, protons, entire atoms, and far more.

Let us state some lessons.

1. Observation of a system can sometimes change it so that a property (eg polarization) has a value (eg vertical) it did not have before. This seems clearly to have happened in the case of polarized light, where fully 50% of it passes through even though far less than 50% already had the correct polarization for the filter. In fact, 50% can get through even when *none* of the light has the correct polarization, as in the case of the 45-degree filter placed behind the vertical one. And yet the system does not always change: if we place a second vertical filter after the first vertical filter, the second filter introduces no changes.

2. We cannot predict just where a given light bundle will end up, but we can predict the pattern that large numbers of them will make. That is, we can state probabilities, and hence expected values, for where they end up.

Now let us move ahead a bit, to a theory that was developed in the 20th century, largely due to work of Einstein and of Schrodinger.

Einstein, and before him Planck, worried about an experimental result that was hard to explain: when zinc is exposed to ultrviolet light, electrons shoot out of the metal. This alone is not so surprising; but if the intensity of the light is changed (ie we dim it!), the rate of electrons (how many per second) changes but not how fast they move. It is as if the light always contributes a fixed unit of energy (or momentum) to an electron, no matter how dim the light is. This seems to say that the light passes on energy to the metal in fixed bundles. Planck argued that light energy is absorbed in special energy bundles he called quanta; to mathematically explain the precise experimental data he needed to introduce a special constant, h, called Planck's constant, that gave precise meaning to these bundles of energy. Specifically he posited $E = h\nu$ where E is the energy of each energy bundle, and ν is the frequency of the light wave (related to the wavelength λ by $\lambda\nu = c$).

Einstein went further, taking the results to be evidence that light *always* consists of little fixed bundles of energy, even when it is not hitting a zinc plate. We also saw this idea of light bundles in all the other cases of dimmed light above, but this one is more dramatic since it also involves a generally accepted kind of particle, the electron, with a known mass. And this new work of Planck and Einstein also had mathematical details that related light bundles to energy in a precise way. (Note that this work of Einstein – which incidentally won him the Nobel Prize – is not his relativity theory at all, but rather quite different work.)

So it seems that light does consist of particles, but ones that obey strange rules, rules that give them a wavelike behavior when very very many of the light particles are looked at, so that statistical summaries of all of the behaviors can be used. Others then began to postulate such wavelike behavior in all particles, not just light particles (now called photons); de Broglie in particular argued that there is a precise relation between the momentum p of an electron and its associated wavelength (the wavelength of its wavelike behavior, fuzzy as that notion is): $\lambda = h/p$. And such wavelike behaviors for particles were found, in glorious accuracy. Specifically, much the same experiments as above for light were done for electrons, with much the same results. So here we see a prime example of fuzzy conceptual underpinnings and yet a precisely measured phenomenon.

Schrodinger next wondered what the strange rules might be that give rise to such wavelike behavior. What sort of mathematical representation might work for particles (photons, electrons, whatever), allowing precise expression of these rules so they could be compared to experimental data? He proposed a "wave function" that is, essentially, just a mathematical expression that has variables for position (x y z) and time (t), and that varies in a wavelike way, eg with sines and cosines. An example is easy to construct, and in fact in an elegant way using the identity

$$e^{i\theta} = cos\theta + isin\theta$$

where $i = \sqrt{-1}$. This simple case of a wavefunction has the form

$$\Psi(x,y,z,t) = ae^{i(k_x x + k_y y + k_z z - \omega t)}$$

Here a, ω, and the k's are special constants. The k's in particular have a wave-like interpretation in accord with the de Broglie equation above: The three k's together form a multiple of the momentum vector $\mathbf{p} = \hbar(k_x, k_y, k_z)$. (Here $\hbar$ is $h/2\pi$ where h is Planck's constant.)

Now this is well and good, but without rules it does little good. Schrodinger noted that if we take the derivative of Ψ with respect to, say, x, we get

$$d/dx\Psi = aik_x e^{i(k_x x + k_y y + k_z z - \omega t)}$$

ie

$$d/dx\Psi = ik_x\Psi$$

and since k_x is just $\hbar$ times the x-component p_x of momentum, and since $1/i = -i$, then

$$(-i\hbar d/dx)\Psi = p_x\Psi$$

This is a simple case of the famous time-independent Schrodinger equation, or TISE. It tells us that an actual physically observable quantity, momentum (in the x-direction), can be computed from the wavefunction. And it is computed by means of an operator – $-i\hbar d/dx$ – that is applied to the wavefunction, to yield back both the observable quantity (the momentum in this case) and Ψ again.

An equation of the form $Opfunction = const \times function$ is called an eigenvalue equation. Usually the operator is fixed and we look for functions (so-called eigenfunctions) and constants (eigenvalues) that make the equation true. But typically an eigenvalue equation has more than one possible solution (ie more than one eigenfunction and eigenvalue pair). This will play a key role below.

This brings us to our next lesson:

3. Values q of a physically measurable (or observable) quantity Q of a physical system can be computed by applying an appropriate mathematical operator $\hat{Q}$ to the wavefunction for that system. The result will be Q times that same wavefunction. This can be written as a more general version of TISE:

$$\hat{Q}\Psi = q\Psi$$

But we still a bit more to describe, before we will be able to give an explanation for phenomena such as the polarized light experiment.

3. States, Observables, and Rules

We have by now assembled enough ideas to be able to give a more thorough statement of the rules of elementary quantum mechanics.

A. At any moment, a given physical system S is in some *state*, which is (described by) a wavefunction Ψ.

B. This state can change over time, since it is a function of time, as well as of position.

C. Normally this change over time is a well-defined and predictable evolution given by the time-dependent Schrodinger equation (TDSE). We will not describe this equation in detail, except to say that it specifies exactly how Ψ changes under "normal" conditions, ie, when the system is not being subjected to a measurement.

D. However measurement in general tends to cause the system to change suddenly and unpredictably from whatever state Ψ it was in to a new state $\Psi\prime$, such that $\Psi\prime$ satisfies TISE (ie, is an eigenfunction or eigenstate) for the operator $\hat{Q}$ corresponding to the observed quantity Q; and observed value q of Q is the corresponding eigenvalue.

E. If Ψ already is an eigenfunction of $\hat{Q}$ then there is no change at all resulting from the measurement, ie the system stays in state Ψ.

F. But if Ψ is not an eigenfunction for the quantity Q being measured, then the state will definitely change into one of the eigenfunctions Q_j of $\hat{Q}$. But which one it changes into (and therefore which eigenvalue q is observed) is not determined but is rather a matter of chance (probabilities).

G. The probability that Ψ changes to a particular eigenstate, say Q_n, can be computed very precisely as follows. We write Ψ as a "superposition" (linear combination) of all the eigenstates Q_j of $\hat{Q}$:

$$\Psi = a_1 Q_1 + a_2 Q_2 + ... + a_n Q_n + ...$$

where the sum can be infinite (and may even in some cases need to be written as an integral instead). This is the "superposition principle", namely that for any observable Q the eigenstates form a "complete" set so that any wavefunction at all can be written uniquely as such a linear combination (or weighted sum) of them. And now the probability that a given Q_n is the one Ψ changes into is simply $|a^2|$. Absolute values are needed since the coefficients a_j are allowed to be (and in general need to be) complex numbers.

The coefficients a_j can be thought of as weights indicating how strongly each Q_j is part of Ψ, or to what extent Q has value q_j, althought the precise meaning is in terms of the absolute value squared.

Thus a particular state that a system S might be in at some time t, might have more than one "value" of some variable (such as momentum) at that time. This is so even if the system is a single particle. People sometimes speak as if the particle had many "incarnations" all at once, or parallel existences. But this also means that we cannot "see" (measure) these values, until we set up a measuring apparatus which collapses that superposition state into just one and the rest disappear. It is as if we were to say "there is an orange and purple hippopotamus in this room whenever no one is there to see it." And yet, although we never observe the superposed states, although all but one disappears when we try to observe the given property, the math works out in perfect agreement with experiment! Nature seems to require that the superposed states be there, in order to interfere with each other in a wavelike manner, so that when they collapse into just one observed state and value, the results when averaged over many cases give the wavelike pattern that we find experimentally.

At last we are done with our exposition of elementary quantum mechanics. We will apply it, below, to the polarization experiment in order to get a feel for how it works.

But it have already seen the essential ingredients for quantum computation, namely that (i) a system may be in a state that does not have a well-defined value for some observable quantity; but (ii) that state will change into a new state that does have such a value if we try to observe that quantity; and (iii) that in a sense the old state has all possible values of that observable implicitly represented as a superposition of all the eigenstates for that quantity; and finally (iv) until we perform that measurement, the system will evolve predictably (and therefore with all the implicit eigenstates evolving along with it).

The trick of quantum computation is to get the system to evolve unmeasured, in a way we want, so that all the implicit eigenstates of the quantity that we will measure *eventually* also evolve as we want, together – ie in parallel – until we actually do the measurement at the end (at which time all but one of those eigenstates "collapse" away leaving only the one we actual observe).

But to do this, we need to find ways to get a system to evolve, unmeasured, as we want. This can be extremely tricky.

For instance, suppose we have an array of n numbers, and we want to multiply each of them by 2. Ordinarily this will take n steps. But if we can encode each of the numbers as an eigenstate of some operator, and if we can get a physical system to be in a state Ψ that not only is a superposition of all those eigenstates but also is a state that we can get to evolve so that those eigenstates change into ones encoding 2 times each of our numbers, then we will "have" the n results all at once (ie as soon as the evolving has been done. But we are still not done since the answers are hidden inside the superposition, and if we try to observe any of the doubled data values then system changes into *one* of the eigenstates and the others disappear leaving us only with that one value rather than n values. So a problem that requires as much output data as input tends not to be a good one for quantum computation.

But this very fact makes "quantum information processing" good for sending encrypted messages. Here the encryption also involves a superposition; and then if anyone intercepts the message and decodes it, the message (the superposition) collapses entirely and so even the sender and intended receiver can tell it has been read.

Examples of problems that do seem to lend themselves to quantum computation (unlike that of doubling data values) are few and far between. So far the only one of any practical interest seems to be that of factoring large numbers. There is an algorithm due to Shor, that in principle allows what would normally be a vast number of computations on many individual pieces of data to be done as a single computation (evolution of the wavefunction that is a superposition of all the individual data) in such a way that the answer can be read out from any of the eigenstates at all, once the final measurement is made.

And quantum computers have in fact been built, and applied to Shor's algorithm. But the technology so far is very primitive, and the "large" numbers are not yet large. A few years ago one such machine was able to factor the number 6 using Shor's algorithm, and this year the number 15. But the difficulties are practical ones of machine design; it appears clear that the underlying principles are correct, so it may only be a matter of time until we have vast computational power at our hands, but only a handful (or less) of algorithms that we know how to encode.

To end our discussion, we briefly present two standard applications of quantum mechanics. Note well: these are *not* examples of quantum computation!

The first one is a quantum mechanical description of the polarization phenomenon. A given photon at any time is in a state Φ which may or may not be a vertical-filter eigenstate, ie it may or may not already be vertically polarized. But if not, then it is a superposition of eigenstates, of which there turn out to be exactly two key ones: vertical and horizontal. That is, observing the photon with a vertical filter means finding out whether it passes through (this vertical state we write as $|v>$) or not (the horizontal state, written $|h>$). Any weighted sum $a|v> +b|h>$ of these represents a superposed state in which the photon has a mix of the vertical and horizontal states (as long as it is normalized so that the total probability is right: $|a^2|+|b^2| = 1$). The verticality operator $\hat{V}$ in this case simply operates on this superposed state and returns the $|v>$ part only. Thus $|v>$ itself is an eigenstate (with eigenvalue 1, meaning "yes") and so is $|h>$ with eigenvalue 0 ("no"). Now what about a 45-degree polarization? A photon in that state has a perfectly even mix (superposition) of the vertical and horizontal: $1/\sqrt{2}|v> +1/\sqrt{2}|h>$. And the probability a photon in this initial state collapses into state $|v>$ (when we measure it with a vertical filter) we can compute: it is just the absolute value squared of the coefficient of $|v>$ in that superposition, namely $|\sqrt{2}^2| = 1/2$.

Our second application concerns chemistry, or more specifically the energy levels of electrons in an atom (which determine the chemical properties of that element). An electron in an atom is in a so-called bound state: it cannot move feely but rather is constrained to stay fairly near the nucleus by the attractive electrical force between it and the protons in the nucleus. One can write a version of TISE for the energy observable, and solve it for the various energy eigenfunctions and associated energy eigenvalues. It turns out that the mathematics gets very tricky, and so people make simplifying assumptions in order to get approximate solutions instead. But even so, the eigenvalues one gets are in amazingly good agreement with what is actaully observed in the laboratory. First, one finds that the energies are discrete: not all positive real numbers for electron energy can occur as energy eigenvalues; and this is also what is found experimentally, as is well-known – the electrons in an atom have only very specific energy levels and no others. Second, the actual energy values that are computed from TISE are virtually identical to the laboratory data, in all cases where people have been able to work out the math, eg for simpler atoms such as hydrogen and helium.