Physics 200-04

Heisenberg Quantum Theory

Instead of trying to follow history at this point, let us vear off and discuss modern quantum theory. I will for the next large number of lectures be following Heisenberg rather than Schrödinger. This is not least because I feel that the Heisenberg development is much more true to our modern concepts of what quantum mechanics is. The Schroedinger approach (which is mathematically completely equivalent, as Schroedinger showed in 1927) is I believe misleading as to the content and essential structure of quantum mechanics. This position is somewhat of a minority. Most textbooks follow Schroedinger, as being more intuitive. Only Feynmann in his Feynman lectures follows the Heisenberg course that I know of, for very similar reasons to mine. Toward the end of the course I will introduce you to the Schrödinger approach, the wave mechanics, rather than the matrix mechanics approach.

Let us start with the system that Stern and Gerlach found. They found that silver atoms seemed to have a very strange property. They had a magnetic moment– they acted like a little magnet, but that magnet it seemed could have only two orientations. An inhomogeneous magnetic field would either exert a force up on the atom, or and equal force down. Those two discrete forces were all that seemed to exist.

This was allied with the fact that the atom seemed to have two (low) energy levels, which were degenerate if there was no external magnetic field, but became non-degenerate (ie had different energies) if a magnetic field was placed onto the atom. Now one usually would associate a little magnet with having a variety of orientations– there should one would think be a whole variety of orientations of the magnet with respect to the magnetic field, which would produce a whole variety both of energies in a magnetic field, and forces in an inhomogeneous field.

Thinking a bit, it becomes clear that the two-valuedness of the forces fits in with the two-valuedness of the energy. Imagining the particle to be in the lower energy state. Now if the field were inhomogeneous, the particle would like to lower its energy still more. Ie, it would like to move into an even stronger field, where the internal energy was even lower. Ie, there would be a force on the particle pushing it toward a stronger field.

Similarly, if it were in the upper energy level, it would want to decrease that energy. Ie, it would want to move into a region where the field were weaker. Since there are only two energy levels, there would be only two

possible forces on the particle.

In Matrix mechanics, we can represent the state of the particle by a vector in a two dimensional space. The vector $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ would be used to represent the upper energy level, and the vector $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ will represent the lower energy level. To represent the energy itself, which is called the Hamiltonian for reasons we will not go into now, we write a matrix

$$H = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \tag{1}$$

We note that

$$\begin{aligned} Hv_1 &= E_1 v_1 \\ Hv_2 &= E_2 v_2 \end{aligned} \tag{2}$$

Ie, the vectors v_1 and v_2 are the eigenvectors of the matrix H.

Following Dirac, we will also use a special symbol to represent these vectors. The notation we will use is

$$\begin{aligned} |E_1\rangle &= v_1 \\ |E_2\rangle &= v_2 \end{aligned} \tag{3}$$

This symbol is to be taken as a whole. The $\rangle$ is NOT a greater than symbol. The whole thing, $|...\rangle$ is the symbol for the vector. The contents between the $|$ and the $\rangle$ is used to identify which particular vector one is talking about. This whole symbol $|...\rangle$ is called a ket-vector.

One of the things which we will be using a lot of is another type of vector, the Hermitean adjoint, or the Dirac adjoint of these vectors. Thus if $v = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is some vector with complex coefficients, then the vector $v^\dagger = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix}$ is the Hermitean or Dirac Adjoint. Dirac gave these vectors a special symbol and name, namely $\langle ...|$, which he called bra-vectors. This had nothing to do with women's undergarments. Rather they come from bracket– bra-ket, with the idea that $\langle |...| \rangle$ forms a sort of bracket around the stuff inside.

Thus

$$\langle v| = (|v\rangle)^\dagger = (|v\rangle)^{*T} \tag{4}$$

The symbol stuck inside the ket or bra vector is simply there to tell the reader which vector is being refered to. It is the name of the vector. Thus the above equations can be written

$$H|E_1\rangle = E_1|E_1\rangle \tag{5}$$

which is to say that the ket vector $|E_1\rangle$ is the eigenvector of the matrix H with eigenvalue E_1. The vector $|E_1\rangle$ is just another name in this case for $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Thus in the case of the energy H and the two eigenvectors, what we have done is to represent the possibilities for the two physical states by two dimensional complex vectors, and the values of the energy by the eigenvalues of the matrix H.

Now, if we stick the particle into the state $|E_1\rangle$, sometimes a physical process will change it into the state $|E_2\rangle$ and sometimes something can change $|E_2\rangle$ into $|E_1\rangle$.

The simplest kind of transformation is to use the matrix

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{6}$$

Now we have $S|E_1\rangle = |E_2\rangle$ and $S|E_2\rangle = |E_1\rangle$. (recall what $|E_1\rangle$ means in terms of matrices.)

Interlude

Consider A a Hermitean matrix. Then all of the eigenvalues of A are real.

Proof: Consider the eigenvalue equation

$$A|\lambda\rangle = \lambda|\lambda\rangle \tag{7}$$

Then if we take the Dirac adjoint of both sides, we get

$$\begin{aligned} \langle\lambda|A = (|\lambda\rangle)^\dagger A^\dagger = (A|\lambda\rangle)^\dagger \\ = (\lambda|\lambda\rangle)^\dagger = \lambda^*(|\lambda\rangle)^\dagger = \lambda^*\langle\lambda| \end{aligned} \tag{8}$$

Ie, $\langle\lambda|$ is also an eigenvector, but from the left side, of A with eigenvalue λ^*. Now consider the expression $\langle\lambda|A|\lambda\rangle$. We can operate either to the left or to the right with A. Is we can first multiply the bra times A and then the ket or first A times the ket and then the bra. In the first case we get

$$(\langle\lambda|A)|\lambda\rangle = \lambda^*\langle\lambda||\lambda\rangle \tag{9}$$

In the second we get

$$\langle\lambda|(A|\lambda\rangle) = \lambda\langle\lambda||\lambda\rangle \tag{10}$$

$$\langle\lambda||\lambda\rangle = \begin{pmatrix}\alpha^* & \beta^*\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \alpha^*\alpha + \beta^*\beta > 0 \tag{11}$$

we must have that $\lambda^* = \lambda$, as required.

Continued with Matrix mechanics

Just as we represented the energies of the system by a matrix, with the possible energies being the eigenvalues of that matrix, we can imagine representing other physical qualities by matricees, and the values which we associate with that quality by eigenvalues of that matrix. In particular, the transformation matrix S is a Hermitean matrix. Its eigenvalues turn out to be plus or minus 1. and the two eigenvectors

$$\begin{aligned} S|S,1\rangle &= +1|S,1\rangle \\ S|S,-1\rangle &= -1|S,-1\rangle \end{aligned} \tag{12}$$

Ie, the physical quality associated with the matrix S, whatever that is, can have two possible values, ± 1. If they have a value of 1, then the vector which represents the state of the system with that eigenvalue is the vector $|S,1\rangle$. But if we write our what the equation $S|S,1\rangle = +1|S,1\rangle$ means, we get

$$\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} \tag{13}$$

This gives us the equations

$$\begin{aligned} \alpha &= \beta \\ \beta &= \alpha \end{aligned} \tag{14}$$

Ie, the eigenvector $|S,1\rangle$ is given by a vector of the form

$$|S,1\rangle = \begin{pmatrix}\alpha\\ \alpha\end{pmatrix} = \alpha\begin{pmatrix}1\\ 1\end{pmatrix} \tag{15}$$

This of course has an arbitrary constant. The usual way to specify that constant is to also demand that the

$$\langle S,1||S,1\rangle = 1 \tag{16}$$

which would then require that $|\alpha|^2 = \frac{1}{2}$.

Thus, we can represent both the particle having an energy E and the particle having a quality S with the same mathematics and the same space of vectors.

However there is something strange here. We can write the vector $|S,1\rangle$ as

$$|S,1\rangle = \frac{1}{\sqrt{2}}|E_1\rangle + \frac{1}{\sqrt{2}}|E_2\rangle \tag{17}$$

Ie, the eigenvector which represents the quantity S having a value of 1 seems to be some sort of combination of the two vectors representing the energy as having the value E_1 and E_2. What could this mean? How can the system have a combination of energies?

The first guess was that these coefficients represented the probability that the the particle had energy E_1 and E_2. But Max Born, in a footnote of a paper he wrote, suggested rather that it was the square of these coefficients which was the probablility. Ie, if the particle has a value of 1 for the property S then it will have energy E_1 with a probability of $\frac{1}{2}$ and will have energy E_2 with probability $\frac{1}{2}$.

One of the strange features of the physics of atoms is that probabilities seemed to come naturally to the systems. Radioactive decay, for example, seemed to follow a probability, rather than occuring deterministically. In any interval on averge the same number of atoms would decay, but the actual number fluctuated without apparent cause.

Another Theorem

If one defines a dot product between two vectors by $|a\rangle \cdot |b\rangle = \langle a||b\rangle$ then the eigenvectors of a matrix with different eigenvalues are different.

Proof: Consider the two vectors $|1\rangle$ and $|2\rangle$ to be eigenvectors of a matrix A with different eigenvalues

$$\begin{aligned} A|1\rangle &= \lambda_1|1\rangle \\ A|2\rangle &= \lambda_2|2\rangle \end{aligned} \tag{18}$$

with $\lambda_1 \neq \lambda_2$. Then by the same procedure we followed above, we have

$$\langle 1|A|2\rangle = \lambda_1\langle 1||2\rangle = \lambda_2\langle 1||2\rangle \tag{19}$$

Since the eigenvalues are not equal, the only way this equation can be true is if $\langle 1||2\rangle = 0$.

Continued If we choose all of the eigenvectors always to have unit norm ($\langle\ ||\ \rangle = 1$) then we can always decompose the eigenvector of S in the following way

$$|S,1\rangle = \langle E_1||S,1\rangle|E_1\rangle + \langle E_2||S,1\rangle|E_2\rangle \tag{20}$$

Remember how to read this. The vector $|S,1\rangle$ can be written as a number $\langle E_1||S,1\rangle$ times the vector $|E_1\rangle$ plus another number $\langle E_2||S,1\rangle$ times a second vector $|E_2\rangle$. (since the vectors we have are all two dimensional, these two vectors will always be enough to write any vector in terms of.)

To show this, we can multiply both sides by $\langle E_1|$ which gives us

$$\langle E_1||S,1\rangle = \langle E_1||S,1\rangle\langle E_1||E_1\rangle + \langle E_2||S,1\rangle\langle E_1||E_2\rangle \tag{21}$$

But $\langle E_1||E_2\rangle = 0$ because the two eigenvalues of H, namely E_1 and E_2 are different, and thus by the theorem the two eigenvectors multiplied together are zero.

Also, by assumption, all eigenvectors have unit norm, so $\langle E_1||E_1\rangle = 1$. Thus the above equation is consistant, since we get $\langle E_1||S,1\rangle = \langle E_1||S,1\rangle$.

We can also multiply the vector $|S,1\rangle$ by itself. Since its norm is by assumption unity, we get

$$\begin{aligned} 1 &= \langle S,1||S,1\rangle = ((\langle E_1||S,1\rangle)^*\langle E_1| + (\langle E_2||S,1\rangle)^*\langle E_2|)\,(\langle E_1||S,1\rangle|E_1\rangle + \langle E_2||S,1\rangle|E_2\rangle) \\ &= \langle E_1||S,1\rangle)^*\langle E_1||S,1\rangle\langle E_1||E_1\rangle + \langle E_1||S,1\rangle)^*\langle E_2||S,1\rangle\langle E_1||E_2\rangle \\ &\quad +\langle E_2||S,1\rangle)^*\langle E_1||S,1\rangle\langle E_2||E_1\rangle + (\langle E_2||S,1\rangle)^*\langle E_2||S,1\rangle\langle E_2||E_2\rangle \\ &= |\langle E_1||S,1\rangle|^2 + |\langle E_2||S,1\rangle|^2 \end{aligned} \tag{22}$$

Thus, if we interpret the square of the coefficient of the $|E_1\rangle$ term as the probability that the particle has energy E_1 if it is in the 1 eigenstate of S, and the square of the coefficient of the $|E_2\rangle$ term as the probability that the particle has energy E_2 if it is in the 1 eigenstate of S, then the total probability adds up to 1, as it should. Ie, it should be in one or the other of the energy states.

It is thus at least consistant to argue that $|\langle E_1||S,1\rangle|^2$ is the probability that if the particle is in the 1 eigenstate of S then has energy E_1. Note that we can interpret it the other way as well. Since

$$\begin{aligned} |\langle E_1||S,1\rangle|^2 &= \langle E_1||S,1\rangle(\langle E_1||S,1\rangle)^* \\ &= (\langle S,1||E_1\rangle)^*\langle S,1||E_1\rangle \end{aligned} \tag{23}$$

we can also interpret this same expression as the probability of the particle having S with value 1 in the state $|E_1\rangle$.

It was on this question, the question of probabilities, that first Einstein, and eventually Schrödinger as well finally gave up on quantum mechanics. However, all experiments indicate that this is valid. One example is the Stern-Gelach experiment carried out consecutively.

Quantum Rules– Kinematics

1) Any physical observable is represented by a Hermitean matrix. The eigen-values of that matrix are the possible values that observable can have.

2) If a system has a certain value for a particular observable, then that situation is reprented by the eigenvector for that eigenvalue for that matrix. This eigenvector is called the **state** of the system. The state is always normalised (ie has unit norm $v^\dagger v = 1$ or $\langle v||v\rangle = 1$.

3) If the system has some generic state $|\psi\rangle$, and if $|a\rangle$ is the eigenvector for operator A with eigenvalue a, then the probability that the system has value a for A is

$$Prob(a) = |\langle a||\psi\rangle|^2 = \langle a||\psi\rangle^*\langle a||\psi\rangle \tag{24}$$

Example

Let us say that the matrix B represents some phyisical property of a "two-level" system (ie a system with at most two possible values for any physical property). Let

$$B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{25}$$

We can clculate its eigenvectors by solving the equation

$$Bv = \lambda v \tag{26}$$

or if we take $v = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$

$$\begin{aligned} \alpha + \beta &= \lambda\alpha \\ \alpha + \beta &= \lambda\beta \end{aligned} \tag{27}$$

Solving the second for β and substituting into the first

$$\alpha + \frac{1}{\lambda - 1}\alpha = \lambda\alpha \tag{28}$$

or

$$((\lambda - 1)^2 - 1)\alpha = 0 \tag{29}$$

This will only have a non-zero solution if $\lambda^2 - 2\lambda = 0$ or λ is either 0 or 2.

If λ is 0, then the eigenvector has $\alpha = -\beta$ or

$$|B, 0\rangle = \alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{30}$$

If we want it normalized, we need $\alpha = \frac{1}{\sqrt{2}}$. Thus the normalised eigenvector is

$$|B, 0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{31}$$

For the eigenvalue 2, we find that $\alpha = \beta$ and the normalized eigen-vector is

$$|B, 2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{32}$$

Let us say that the state of the system is given by the normalised vector

$$|\psi\rangle = \frac{1}{5} \begin{pmatrix} 4 \\ 3i \end{pmatrix} \tag{33}$$

Then the probability of the system having value 0 for B is

$$|\langle B, 0||\psi\rangle|^2 = \left| \frac{1}{\sqrt{2}} \begin{pmatrix} 1^* & (-1)^* \end{pmatrix} \frac{1}{5} \begin{pmatrix} 4 \\ 3i \end{pmatrix} \right|^2$$

$$= \left| \frac{1}{\sqrt{50}}(4-3i) \right|^2 = \frac{1}{2} \tag{34}$$

Clearly the probability of having value 2 for B is also $\frac{1}{2}$.

Comment

This is clearly a very different system than classical mechanics. Firstly in classical mechanics a system either has or does not have a certain value for any particular attribute. A particle is described by a generic position, x say, and a generic momentum p. These variables can take a variety of values, but the values are not specifically encoded in the form of the variable. For example you would use exactly the same variable to encode a particle which could only live in a universe with values from 0 to 1 as one with any possible value. In quantum mechanics the possible values that the variable can take are already encoded in the representation, the matrix that represents that attribute.

In classical physics the value of a variable is encoded in the theory by simply giving the variable that value. If you say that a particle with position represented by x is located at position 5 then you simply substitute 5 everywhere in the expressions where you see x. In quantum mechanics the mechanism is different. You do not change the the representation, the matrix, you change the state. If you want to represent the fact that the attribute associated with the matrix B has value 0 say, then you do not change B anywhere, you change the vector representing the state of the system to the eigenvector of B with eigenvalue 0. Ie, the representation of the state of the world is separated from the representation of the attributes. You do NOT substitute 0 everywhere where you see B.

Furthermore, a system need not have a value for B. In classical physics you can imagine that all attributes actually have a value, that anywhere in your expressions where you see x, you can imagine that it is some definite number. In quantum mechanics on the other hand, you cannot do that. If the state is not an eigenvector of B, then you cannot imagine that it really is.