Lecture 2

Relevant sections in text: §1.2

Quantum theory of spin 1/2 – some preliminaries

We now try to give a quantum mechanical description of electron spin which matches the experimental facts described previously.

Let us begin by stating very briefly the rules of quantum mechanics. We shall show what they mean as we go along. But it is best to know the big picture at the outset.

Rule 1

Observables A are represented by self-adjoint operators A on a (complex) Hilbert space $\mathcal{H}$.

Rule 2

States are represented by unit vectors in $\mathcal{H}$. The expectation value $\langle A\rangle$ of the observable A in the state $|\psi\rangle$ is given by the diagonal matrix element

$$\langle A\rangle = \langle\psi|A|\psi\rangle.$$

Rule 3

Time evolution is a continuous unitary transformation on $\mathcal{H}$.

Note that there is an important logical distinction between an observable and its mathematical representation. For example, in the formula for the expectation value above, on the left hand side of the formula the symbol represents something you can measure in the laboratory, independently of any particular theory. You don't have to have even heard of quantum mechanics to measure it. The right hand side tells you how quantum mechanics computes this physical prediction. It would be nice, but is too cumbersome to make this distinction notationally, but you should keep it in mind.

We will now use Rules 1-2 to create a model of a spin 1/2 particle. We will not need Rule 3 for a while (until Chapter 2). We suppose that a spin 1/2 system is completely described by its spin observable $\vec{S}$, which defines a vector in 3-d Euclidean space. As such, $\vec{S}$ is really a collection of 3 observables, which we label as usual by S_x, S_y, S_z, each of which is to be represented by a (self-adjoint) linear operator on a (Hilbert) vector space. We have seen that the possible outcomes of a measurement of any component of $\vec{S}$ is $\pm\hbar/2$. As we will see, because the set of possible outcomes of a measurement of one these observables has two values, we should build our Hilbert space of state vectors to be two-dimensional. A two dimensional Hilbert space* is a complex vector space with a Hermitian scalar product. Let us explain what all this means.

I won't bother with the formal definition of a vector space since you have no doubt seen it before. You might want to review it. We denote the elements of the vector space by the symbols $|\alpha\rangle$, $|\beta\rangle$, *etc.* . Of course, they can be "added" to make new vectors, which we denote by

$$|\alpha\rangle + |\beta\rangle = |\gamma\rangle.$$

We denote the set of scalars (complex numbers) by the symbols a, b, c, *etc.* Scalar multiplication of a vector $|\psi\rangle$ by c is another vector and is denoted by

$$c|\psi\rangle \equiv |\psi\rangle c.$$

We denote the Hermitian scalar product of two vectors $|\alpha\rangle$ and $|\beta\rangle$ by the notation

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*.$$

The scalar product is linear as a function of its second argument:

$$\langle\alpha|(a|\beta\rangle + c|\gamma\rangle) = a\langle\alpha|\beta\rangle + c\langle\alpha|\gamma\rangle,$$

and satisfies

$$\langle\alpha|\alpha\rangle \geq 0,$$

with equality only if $|\alpha\rangle = 0$, *i.e.,* is the zero vector.

If all this leaves you feeling a bit dazed, then your linear algebra background probably needs strengthening. A quick fix is to study some simple examples to give meaning to the symbols, which we shall do now.

Vector spaces – elementary examples

Here are a couple of simple examples of the foregoing material.

First, consider the set of position vectors relative to an origin in 3-d Euclidean space. They form a *real** vector space with addition being defined component-wise or via the parallelogram rule. Scalar multiplication is defined component-wise or by scaling of the length of the vector (and reversing its direction if the scalar is negative). The scalar

* A Hilbert space is a vector space with scalar product and a certain completeness requirement. As it happens, this completeness requirement is redundant for finite dimensional vector spaces, so we don't need to worry with it just yet.

* A real/complex vector space uses real/complex numbers for the scalars.

product of two vectors is just the familiar dot product. Since this is a real vector space, the complex conjugation business is trivial.

Second, and much more importantly for our present purposes, consider the *complex* vector space $\mathbf{C}^2$, the set of 2-tuples of complex numbers. Elements of this vector space can be defined by column vectors with complex entries, *e.g.,*

$$|\psi\rangle \Longleftrightarrow \begin{pmatrix} a \\ b \end{pmatrix}.$$

Addition is defined component-wise in the familiar way. Scalar multiplication is also defined component-wise, by multiplication of each element of the column, *e.g.,*

$$c|\psi\rangle \Longleftrightarrow \begin{pmatrix} ca \\ cb \end{pmatrix}.$$

The scalar product of two vectors,

$$|\psi_1\rangle \Longleftrightarrow \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \quad |\psi_2\rangle \Longleftrightarrow \begin{pmatrix} a_2 \\ b_2 \end{pmatrix},$$

is given by

$$\langle\psi_1|\psi_2\rangle = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}^\dagger \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1^* & a_2^* \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = a_1^* a_2 + b_1^* b_2.$$

In each of these two examples you should verify the various abstract properties listed above. Pay particular attention to the scalar product.

Note that we have refrained from literally identifying the vectors with the column vectors. The reason for this is that we will be adopting the point of view that a column vector is really defined as the components of a vector in a given basis. One can change the basis and thereby change the column vector which corresponds to one and the same vector. For now, you can safely ignore this subtlety and just think of the notation $|\psi\rangle$, *etc.* as a fancy way to manipulate column and row vectors. Later we will tighten up the interpretation of the notation.

As it happens, every complex 2-dimensional vector space can be expressed in the above form (column vectors, *etc.*). So the vector space used to model the spin 1/2 system has this mathematical representation. We shall make use of this extensively in what follows.

Bras and Kets

We have already used the notation $|\alpha\rangle$ for the elements of a vector space. Of course, we can call these things "vectors", but it is also customary – following Dirac – to call them "kets", which we shall explain in a moment. Associated to every ket in a Hilbert space there is another kind of vector which lives in a different "dual" vector space to $\mathcal{H}$. We call this dual vector space $\mathcal{H}^*$ and the elements of this space are called "dual vectors" or "bras" (!). If you think of kets as column vectors, then the bras are row vectors and the correspondence between a ket and bra is that induced by the Hermitian adjoint † (complex-conjugate-transpose). More generally, every vector $|\psi\rangle$ determines a *linear* function F_ψ on $\mathcal{H}$, which is defined using the scalar product by

$$F_\psi(|\alpha\rangle) = \langle\psi|\alpha\rangle. \tag{1}$$

Conversely, every linear function on a Hilbert space can be constructed via the scalar product as above.

The set of all linear functions on a vector space is itself a vector space – the dual vector space. (Linear functions can be added and scalar multiplied in the obvious way to make new linear functions, *etc.* Try and see if you can make this into a proof!) For a Hilbert space the dual vector space can be identified with the original vector space in the manner indicated above and so it becomes pretty obvious that the set of linear functions is a vector space. Instead of F_ψ, we use the notation $\langle\psi|$ for the linear function defined by (1). Thus a ket $|\psi\rangle$ defines a bra $\langle\psi|$ which is a linear function on $\mathcal{H}$ via

$$|\alpha\rangle \to \langle\psi|\alpha\rangle.$$

Sometimes we write

$$\langle\psi| = (|\psi\rangle)^\dagger,$$

which makes good sense if you are thinking in terms of column and row vectors. Note that because the bras form a complex vector space they can be added and scalar multiplied as usual. We use the obvious notation:

$$\langle\alpha| + \langle\beta| = \langle\gamma|, \quad c(\langle\alpha|) = c\langle\alpha| = \langle\alpha|c,$$

and so forth. Note in particular, though, that the bra corresponding to the ket $c|\psi\rangle$ involves a complex conjugation:

$$(c|\psi\rangle)^\dagger = \langle\psi|c^*.$$

To see this we consider the linear function defined by $|\gamma\rangle = c|\psi\rangle$. Evaluated on some vector $|\alpha\rangle$ we get

$$\langle\gamma|\alpha\rangle = \langle\alpha|\gamma\rangle^* = (\langle\alpha|c|\psi\rangle)^* = c^*\langle\alpha|\psi\rangle^* = c^*\langle\psi|\alpha\rangle.$$

The origin of the terminology "bra" and "ket" comes from the pairing between vectors and linear functions via the scalar product $\langle\psi|\alpha\rangle$, which uses a *bracket* notation. Get it? This terminology was introduced by Dirac, and the notation we are using is called "Dirac's bra-ket notation".