

NOTES ON THE MATHEMATICAL STRUCTURE OF QUANTUM MECHANICS

G. B. Folland. Math 480C, 4/16/08

To make a mathematical model of a physical system, you need to specify the possible *states* of the system, and the description of the “state space” (set of possible states) varies from system to system.

For a classical system of N particles moving according to Newton’s laws, the state can be specified by giving the positions and velocities (or momenta) of all the particles at a particular time. Thus the state space can be taken as $(\mathbb{R}^3)^{2N} \cong \mathbb{R}^{6N}$; a point in the state space looks like $(\mathbf{x}_1, \mathbf{v}_1, \dots, \mathbf{x}_N, \mathbf{v}_N)$ where \mathbf{x}_j and \mathbf{v}_j are the position and momentum of the j th particle.

Here’s a simple classical system that may be useful to think about: a collection of N coins. You put the coins in a bag, shake it up, and pour them out onto a table and see whether each of them comes up heads or tails. Here the state space is the set of N -tuples (s_1, \dots, s_N) where each s_j is either H (heads) or T (tails). Thus the state space here is finite; it consists of 2^N points.

For a quantum system, things are different. *The state space for any quantum system is the set of all unit vectors in some Hilbert space \mathcal{H} , with the understanding that two unit vectors that are scalar multiples of each other define the same state.* In more detail: a Hilbert space is a complex scalar product space that is complete (has no holes) with respect to the notion of convergence defined by the norm. Let \mathcal{H} be a Hilbert space, and let \mathcal{U} be the set of unit vectors in \mathcal{H} . We put an equivalence relation on \mathcal{U} by saying that $v \sim w$ if there is a complex number c (necessarily with $|c| = 1$) such that $w = cv$. Then \mathcal{U}/\sim is a candidate for being the state space of a quantum system. (Equivalently, instead of \mathcal{U} we could consider the set \mathcal{V} of all nonzero vectors in \mathcal{H} , with the equivalence relation that $w \sim v$ if $w = cv$ for some nonzero complex number c , i.e., if w and v lie on the same complex line through the origin. The equivalence classes here are called *rays*, and the set of equivalence classes is called the *projective space* associated to the vector space \mathcal{H} . The book makes a fairly big deal of using these terms, but don’t be frightened of them.)

In practice, we are usually a little sloppy with terminology. We will refer to the whole Hilbert space \mathcal{H} as the state space (although states should be given by unit vectors), and we will refer to unit vectors as states (although two unit vectors that differ by a phase factor define the same state).

Now, just specifying the states isn’t enough; we have to know how to use this specification to get some physical results. So let’s start by thinking about yes-no questions concerning the system at hand. These are questions you can ask about the system such that whenever you go out and observe the system, you always get the answer “yes” or “no.” For the classical system of N particles, such questions are things like “Is the kinetic energy of the 3rd particle equal to 5 joules, to within an accuracy of 0.1 joule?”, or “Is the x -coordinate of the position of the first particle positive?” For the system of N coins, such questions are things like “Are there 27 heads showing?” or “Is the number of tails odd?”

For a classical system, any yes-no question divides the state space into two parts: the states where the answer is yes, and the states where the answer is no. These are complementary subsets of the state space: each state is in one set or the other, and no state is in both.

Here is where quantum mechanics is really different: For a yes-no question about a quantum system, there are some states where the answer is always yes, some states where the answer is always no, and some states where the answer can turn out to be either yes or no. That is, if you make repeated observations of the system, sometimes you'll get the answer yes, sometimes no, and all you can predict in advance is the *probability* that the answer is yes. More precisely: *For any yes-no question there is a closed vector subspace V of the state space \mathcal{H} such that the answer is always yes for any state in V and always no for any state in V^\perp . In general, if v is any unit vector, let v_y and v_n be the orthogonal projections of v onto V and V^\perp , respectively; then the probability that the answer is yes when the system is in state v is $\|v_y\|^2$ and the probability that the answer is no is $\|v_n\|^2$. This makes sense because $\langle v_y, v_n \rangle = 0$, so*

$$\|v_y\|^2 + \|v_n\|^2 = \|v_y\|^2 + 2\langle v_y, v_n \rangle + \|v_n\|^2 = \|v_y + v_n\|^2 = \|v\|^2 = 1.$$

This is weird. But it's how the world works.

Once we have analyzed yes-no questions, we can proceed to analyze *observables*, which for the purposes of this discussion are real-valued variables associated to the system whose values can be determined by observations or experiments. For the classical system of N particles, the common observables are things like the total kinetic energy of the system, the x-component of the angular momentum of the 4th particle, or the y-component of the center of mass of the system. For the system of N coins, observables are things like the number of heads minus the number of tails. In general, in any classical system an observable is just a real-valued function defined on the state space of the system. But in a quantum system this description doesn't work, because of the probabilistic aspects of the problem.

To see what's going on, we can analyze quantum observables in terms of the yes-no questions "Does the value of the observable lie in the set E ?", where E is any (nice) subset of \mathbb{R} . So as not to be too technical, let's think of an observable O that can only take on a finite number of values, say $\lambda_1, \dots, \lambda_N$. (The "number of heads minus number of tails" for the system of coins is a good example, if you imagine the coins as a quantum system. You can make real quantum systems of a similar nature by considering a finite collection of electrons and replacing "heads" and "tails" by "spin up" and "spin down.") For each $j = 1, \dots, N$, you can ask the question "Is the value of O equal to λ_j ?" There will be a certain subspace V_j of the state space where the answer is yes, and the answer will be no on the orthogonal complement V_j^\perp . Since the possibilities "The value of O is λ_j " and "The value of O is λ_k " are mutually exclusive if $j \neq k$, we have $V_j \perp V_k$ if $j \neq k$. On the other hand, since we assume that O must assume one of the values λ_j , the spaces V_j must span the whole state space \mathcal{H} . In short, \mathcal{H} must be the orthogonal direct sum

$$\mathcal{H} = V_1 \oplus V_2 \oplus \dots \oplus V_N.$$

Now, the observable O is some sort of "quantity" that takes on the value λ_1 on the subspace V_1 , the value λ_2 on the subspace V_2 , etc. This is strongly reminiscent of a mathematical

structure that should be familiar. To wit, the observable O should be represented not by a function on the state space \mathcal{H} but by the *linear operator* for which the λ_j and V_j are the eigenspaces and eigenvalues. That is, for any $v \in \mathcal{H}$ we can write $v = v_1 + \cdots + v_N$ with $v_j \in V_j$, and we define

$$Tv = T(v_1 + \cdots + v_N) = \lambda_1 v_1 + \cdots + \lambda_N v_N.$$

In this situation, if v is a unit vector we have $\|v_1\|^2 + \cdots + \|v_N\|^2 = \|v\|^2 = 1$ as before, and the numbers $\|v_j\|^2$ represent the probabilities that the observable O will have the value λ_j when the system is in the state v . The mean value, or expectation in the sense of probability theory, of the observable O in the state v is

$$\lambda_1 \|v_1\|^2 + \cdots + \lambda_N \|v_N\|^2 = \langle v, Tv \rangle.$$

The operator T that we get this way is not just any old linear operator: it is Hermitian, because

$$\langle Tv, w \rangle = \sum \langle \lambda_j v_j, w_j \rangle \sum \langle v_j, \lambda_j w_j \rangle = \langle v, Tw \rangle.$$

(We are using the fact that the λ_j are real here.) Conversely, *if* $\dim \mathcal{H} < \infty$, every Hermitian operator has this form — that is, its eigenvalues are real and the whole space \mathcal{H} is the orthogonal direct sum of its eigenspaces; this is the finite-dimensional *spectral theorem*. Things are more complicated in infinite-dimensional spaces (which are the ones you need to do many quantum problems), but the basic correspondence remains valid:

In quantum mechanics, real-valued observables are represented by Hermitian operators. The states where the observable T has a definite value λ are the eigenvectors of T with eigenvalue λ . In any state represented by a unit vector v , the average value or expected value of T in the state v is $\langle v, Tv \rangle$.

Now, to bring things to a more concrete level, let's describe the state space for a particle moving in \mathbb{R}^3 . The Hilbert space in question is $L^2(\mathbb{R}^3)$, and for analyzing the motion of the particle, the fundamental yes-no questions are “Is the particle located in E ?”, where E ranges over all (nice) subsets of \mathbb{R}^n . The subspace on which the answer to the question “Is the particle in E ?” is yes is

$$V_E = \{f \in L^2(\mathbb{R}^3) : f(x) = 0 \text{ for } x \notin E\}.$$

The orthogonal projection onto E is given by $[\Pi_E f](\mathbf{x}) = \chi_E(\mathbf{x})f(\mathbf{x})$, where $\chi_E(\mathbf{x})$ is equal to 1 for $\mathbf{x} \in E$ and equal to 0 for $\mathbf{x} \notin E$. Thus, if the particle is in state f (where f is a unit vector in $L^2(\mathbb{R}^3)$), the probability that the particle lies in E is

$$\|\Pi_E f\|^2 = \int_{\mathbb{R}^3} |\chi_E(\mathbf{x})f(\mathbf{x})|^2 d^3\mathbf{x} = \int_E |f(\mathbf{x})|^2 d^3\mathbf{x}.$$

In other words, $|f(\mathbf{x})|^2$ represents the *probability density* of the position of the particle in the state f .

One should note that there are no states in which the particle is definitely located at a particular point \mathbf{a} . The vector representing such a state would be the “delta-function” $\delta(\mathbf{x} - \mathbf{a})$, which is zero everywhere except at $\mathbf{x} = \mathbf{a}$, but delta-functions aren't genuine functions. (Physicists use such idealized states anyway, sometimes, and they refer to them as “non-normalizable states.”) However, there are states in which the particle is definitely located in a very small region such as $\{\mathbf{x} : |\mathbf{x} - \mathbf{a}| < 10^{-10}\}$.

The observables corresponding to the classical position variables x_1, x_2, x_3 (the three coordinates of the position of the particle) are the operators X_1, X_2, X_3 defined by

$$[X_j f](\mathbf{x}) = x_j f(\mathbf{x}) \quad (j = 1, 2, 3),$$

that is, X_j = multiplication by the j th coordinate function. The mean value of the observable X_j in the state f is $\int x_j |f(\mathbf{x})|^2 d^3\mathbf{x}$. If f vanishes outside a very small set E containing the point \mathbf{a} (i.e., if the particle is definitely located near \mathbf{a}), then $x_j \approx a_j$ on E , and hence this mean value is approximately $a_j \int |f(\mathbf{x})|^2 d^3\mathbf{x} = a_j$ as it should be.

The other ingredients we need to build the quantum analogue of a classical Newtonian system are the momentum variables. There are several ways to arrive at the right expression for them, but they all involve excursions into physics that would take too long to explain here. Suffice it to say that the states with a definite momentum \mathbf{p} are the “plane waves” $\phi_{\mathbf{p}}(\mathbf{x}) = e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}$ where \hbar is Planck’s constant. (Like the delta-functions, these are only idealized states because they aren’t in L^2 .) The operator that multiplies the function $e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}$ by p_j (the j th component of \mathbf{p}) is $(\hbar/i)\partial/\partial x_j$, so this is the operator corresponding to the classical observable p_j :

$$P_j f = \frac{\hbar}{i} \frac{\partial f}{\partial x_j}.$$

Now we can build other functions of position and momentum. The quantum version of the angular momentum variables

$$\mathbf{l} = \mathbf{x} \times \mathbf{p} = (x_2 p_3 - x_3 p_2, x_3 p_1 - x_1 p_3, x_1 p_2 - x_2 p_1)$$

are the operators (L_1, L_2, L_3) defined by

$$L_1 = X_2 P_3 - X_3 P_2, \quad \text{i.e.,} \quad [L_1 f](\mathbf{x}) = \frac{\hbar}{i} \left(x_2 \frac{\partial f}{\partial x_3} - x_3 \frac{\partial f}{\partial x_2} \right), \quad \text{etc.}$$

The classical total energy of a particle moving in a potential $V(\mathbf{x})$ is the sum of the kinetic and potential energies

$$E = \frac{1}{2} m |\mathbf{v}|^2 + V(\mathbf{x}) = \frac{|\mathbf{p}|^2}{2m} + V(\mathbf{x}).$$

The operator corresponding to the quantum energy is called the *Hamiltonian* and is denoted by H :

$$H = \frac{1}{2m} (P_1^2 + P_2^2 + P_3^2) + V(\mathbf{X}) = -\frac{\hbar^2}{2m} \nabla^2 + (\text{mult. by } V(\mathbf{x})),$$

that is,

$$[Hf](\mathbf{x}) = -\frac{\hbar^2}{2m} \nabla^2 f(\mathbf{x}) + V(\mathbf{x}) f(\mathbf{x}).$$

This is the operator whose eigenvectors are the states with definite energy. With the particular choice $V(\mathbf{x}) = -c|\mathbf{x}|^{-1}$, corresponding to the inverse-square-law force, it is the operator that we have to analyze to find the “energy shells” for the electron in a hydrogen atom.