12. The Principles of Quantum Mechanics

So far this course has focused exclusively on the quantum behavior of a single, structureless particle in one dimension. We've described this system in terms of wavefunctions, found the wavefunctions that correspond to definite values of position, momentum, and energy, predicted how the wavefunctions evolve in time, and explored the energy level structure and scattering behavior of particles subject to various specific potential energy functions.

Along the way, however, I've sneakily introduced some more abstract language for describing these systems: eigenfunctions and eigenvalues, operators, complete and orthonormal sets of basis functions, and even vectors and matrices. This is the powerful language of *linear algebra*, and it is just what we need to generalize the principles of quantum mechanics to a broader variety of systems: multiple dimensions of space, multiple particles, and particles that have internal degrees of freedom.

In this lesson I will lay out the *principles* of quantum mechanics once and for all, so you can see them all in one place. To somewhat alleviate the abstraction, I'll also describe how these principles pertain to the specific one-dimensional systems we've been studying so far.

Principle 1: The states of a quantum system correspond to vectors in a vector space.

Recall from linear algebra that a *vector space* is any collection of objects that you can both add to each other and multiply by scalars (numbers). These operations need to obey the usual commutative, associative, and distributive rules, and the space needs to be *closed* under these operations, so that the result is always another vector in the space. In general the scalars can be either real or complex numbers, but in quantum mechanics they are always complex. The vectors can be expressed as n-tuples of numbers, functions of one or more variables, or more complicated objects—so long as you can define legitimate ways of adding and scalar-multiplying them. The particular vector space that corresponds to a given quantum system will depend on the system. More often than not, the vector space will be infinite-dimensional.

In quantum mechanics we further require that the vector space have an *inner product* operation that obeys the same algebraic rules as the familiar dot product. The only twist is that the inner product of vectors ψ_1 and ψ_2 can yield a complex number, and reversing their order must then yield the complex conjugate:

$$\langle \psi_2, \psi_1 \rangle = \langle \psi_1, \psi_2 \rangle^*. \tag{1}$$

The inner product of a vector with itself must therefore be real, and the *norm* or "length" of a vector is the square root of its inner product with itself.

For a structureless particle in one dimension, the appropriate set of vectors is the complex-valued square-integrable functions of a single variable: $\psi(x)$. By "square-integrable" I simply mean that when you integrate the square modulus of the function over all x, the result has to be finite (that is, the function is normalizable). The inner product of two such vectors is defined as

$$\langle \psi_1, \psi_2 \rangle = \int_{-\infty}^{\infty} \psi_1^*(x)\psi_2(x)\,dx, \tag{2}$$

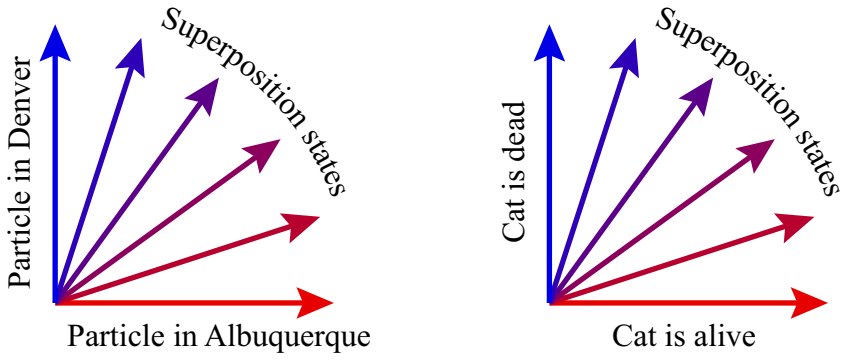
so the square-integrable requirement is needed for inner products to always be finite. The set of *all* functions of x would be a vector space but would not have a sensible inner product. On the other hand, the set of all normal*ized* functions of x is not a vector space, because it isn't closed under addition or scalar multiplication. The precise rule, then, is that *all* multiples of a given function $\psi(x)$ actually correspond to the *same* physical state of this one-dimensional system.

Similarly, for any quantum system, a given physical state actually corresponds to a specific *direction* in the vector space, that is, to some vector ψ and to all scalar multiples of that vector. For some purposes, however, we will require that our vectors be normalized, that is, $\langle \psi, \psi \rangle = 1$. Even then, we can multiply a vector by an arbitrary phase factor $e^{i\phi}$ without changing the corresponding physical state and without affecting the vector's normalization.

Representing system states by vectors in a vector space is radical, because it means that for any two states, corresponding to vectors ψ_1 and ψ_2, you can also have arbitrary *mixtures* that correspond to the linear combinations

$$c_1\psi_1 + c_2\psi_2, \tag{3}$$

where c_1 and c_2 are complex numbers. If ψ_1 is a state for which the particle is in Albuquerque and ψ_2 is a state for which the particle is in Denver, then you can also have states for which the particle is partly in Albuquerque and partly in Denver. If ψ_1 is a state for which the cat is alive and ψ_2 is a state for which the cat is dead, then you can also have states for which the cat is part alive and part dead.



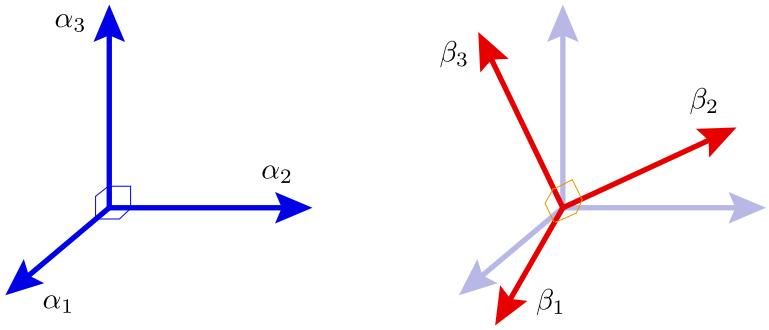
Principle 2: The observable quantities for a quantum system correspond to operators that act in the vector space.

An operator, in general, is a transformation that converts any vector into another vector. In quantum mechanics we consider only *linear* operators, for which (denoting the operator as A)

$$A(c_1\psi_1 + c_2\psi_2) = c_1 A\psi_1 + c_2 A\psi_2, \tag{4}$$

for any two vectors ψ_1 and ψ_2 and any two scalars c_1 and c_2.

Any operator that corresponds to an observable quantity must also have only *real* eigenvalues that equal the possible values of that quantity, and must have an associated set of eigenvectors that form a complete orthogonal basis for the vector space; these are the special state vectors for which the corresponding quantity is well defined. In general, however, two different observables, with corresponding operators A and B, will have two different sets of orthogonal basis vectors, $\{\alpha\}$ and $\{\beta\}$. It's hard to visualize all these basis vectors in an infinite-dimensional vector space, but visualizing just three dimensions at a time is better than nothing:



An operator whose eigenvalues are real and whose eigenvectors are orthogonal is said to be *Hermitian*. I'll explain this term more fully in an upcoming lesson.

For a structureless particle in one dimension, represented by vectors that are functions $\psi(x)$, we've already seen that the position operator is x and the momentum operator is $-i\hbar\, d/dx$. Every real number is an eigenvalue of either of these operators, but they're both a bit awkward because their eigenvectors are non-normalizable functions that therefore lie outside our vector space. What's important, however, is that each of these sets of eigenvectors still forms a basis, so we can express any state vector in terms of them.

We've also seen that from the position and momentum operators we can build more complicated operators such as kinetic energy, potential energy, and total energy. The last of these can have a discrete spectrum of eigenvalues, and then the corresponding eigenvectors do lie within the vector space.

Rather than focusing on these operators themselves, it's more vivid (and often more useful) to focus on their complete sets of eigenvectors and corresponding

eigenvalues. For each observable quantity (position, momentum, energy, and so on) there is a complete, orthonormal set of eigenvectors that correspond to the states for which that quantity has a definite value. The specific value that corresponds to a given eigenvector is its associated eigenvalue. The collection of eigenvectors is complete, so any other vector can be expressed as a linear combination of them, that is, as a specific mixture of the definite-value vectors.

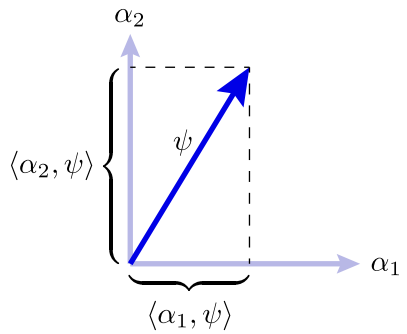
Principle 3: Measurement probabilities are given by the squares of the appropriate vector components.

Suppose the current normalized state vector of a system is ψ and you decide to measure observable A, whose normalized eigenvectors are α_n with corresponding eigenvalues a_n. Then the component of ψ along the α_n direction is the inner product $\langle \alpha_n, \psi \rangle$, and the probability of obtaining a_n when you make your measurement is the square modulus of this component:

$$(\text{Probability of } a_n) = |\langle \alpha_n, \psi \rangle|^2. \tag{5}$$

Thus, the greater the degree to which ψ lies along the α_n direction, the greater the chance of obtaining a_n. In the special case where $\psi = \alpha_n$, the probability is 100%.

Here's the standard way of visualizing these vector components, showing only two of the basis vectors for simplicity:



Because ψ is normalized and the eigenvectors are orthonormal, the Pythagorean theorem (or its higher-dimensional equivalent) guarantees that the total probability for all possible measurement outcomes is always exactly 1.

I've written equation 5 for the case where the eigenvectors and eigenvalues are discrete. If instead the parameter n varies continuously, then (with appropriate normalization) the probability of obtaining a result in any range of a values is

$$\begin{pmatrix} \text{Probability of result} \\ \text{between } a(n_1) \text{ and } a(n_2) \end{pmatrix} = \int_{n_1}^{n_2} |\langle \alpha_n, \psi \rangle|^2 \, dn. \tag{6}$$

Equation 6 is just an abstract version of the more familiar formulas for calculating position and momentum probabilities for a single-particle, one-dimensional

system. For position, the continuous parameter n is actually x, the eigenvectors α_n are delta functions, and the inner product is simply the wavefunction evaluated at x. For momentum, the continuous parameter is p, the eigenvectors α_p are the sinusoidal functions e^{ipx} (times a "normalization" constant), and the inner product is the momentum-space wavefunction, $\Phi(p)$. Equation 5, on the other hand, applies to energy measurements of trapped particles with discrete energy eigenvalues.

Principle 4: After a measurement, the state vector "collapses" to the eigenvector that corresponds to the result obtained.

So, for example, when you measure the position of a particle in one dimension, its wavefunction collapses to a delta function located at the place where you found the particle. Similarly, when you measure the momentum or energy of a particle, its state vector collapses into the corresponding eigenvector.

There's a slight complication when two or more orthogonal eigenvectors have the same associated eigenvalue. We then say that the eigenvalue is *degenerate*. A familiar example in one dimension is kinetic energy, for which each eigenvalue ($p^2/2m$ if the particle is nonrelativistic) corresponds to two independent wavefunctions, one with positive momentum and the other with negative momentum. When you measure a quantity and obtain a degenerate eigenvalue, the state vector collapses to its projection onto the subspace of eigenvectors that correspond to that eigenvalue.

Please note that the "collapse" of the state vector is absolutely necessary for consistency: if you measure the same quantity twice in rapid succession, the second measurement must give the same result as the first, with probability 100%.

On the other hand, this principle is the source of much of the confusion and mystery of quantum mechanics, because nowhere does the theory define exactly what constitutes a "measurement." This ambiguity isn't much of a problem in practice, because in the laboratory there's usually a clear distinction between the quantum system we're studying and the apparatus we're using to study it. The apparatus is generally big and heavy enough that we can treat it classically, and the "measurement" process is then just the interaction between the quantum system and the classical apparatus.

The problem is instead conceptual: Fundamentally, the laboratory apparatus must also obey quantum mechanics, so in principle we could construct a state vector for the larger system that includes the apparatus. Then the interaction between the apparatus and the smaller quantum system would *not* constitute a "measurement," and would not trigger any "collapse," so the collapse would presumably occur when we read the instruments to determine the measurement outcome. But why stop there? Quantum mechanics must also apply to *us*, so there must be a state vector that describes everything in the laboratory including the experimenters, and now it can't collapse until someone or something outside this larger system performs a "measurement." In the most extreme case we would apply quantum mechanics

to the entire universe, and then there's nothing outside the system to perform a measurement and cause the state vector to collapse.

Physicists have argued for almost 90 years over the correct interpretation of the collapse principle. The arguments continue to this day, and I certainly won't try to resolve the issue here. My personal opinion, however, is that the correct resolution must have something to do with the complexity of the measuring apparatus and the practical impossibility of determining its exact quantum state.

Principle 5: The state vector of a quantum system evolves in time according to the time-dependent Schrödinger equation.

As long as we express the TDSE in terms of the Hamiltonian operator, it looks exactly the same for all quantum systems:

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi. \tag{7}$$

For a structureless particle in one dimension, ψ is a square-integrable function of x and, if the particle is nonrelativistic, $H = p^2/2m + V(x)$, where p is the momentum operator, $-i\hbar\,\partial/\partial x$. For a general quantum system, ψ is a vector in some vector space that might be entirely different, while the Hamiltonian operator H is whatever operator corresponds to the total energy of the system.

In any case, the TDSE is a differential equation that describes the continuous changes in ψ as time passes. Moreover, if we can first solve the time-*in*dependent Schrödinger equation to find the system's energy eigenvalues and eigenvectors, we can then expand $\psi(t = 0)$ in terms of these eigenvectors and immediately slip wiggle factors, $e^{-iE_nt/\hbar}$, into each term to obtain the solution to the TDSE. This procedure works in exactly the same way for an arbitrary quantum system as it does for a particle in one dimension.