Quantum Mechanics in 3-d

Relevant sections in text: Chapter 4

Quantum mechanics in three dimensions

Our study of a particle moving in one dimension has been instructive, but it is now time to consider more realistic models of nature by working in three dimensions. It is reasonably straightforward to generalize our model of a particle moving in one dimension to a particle moving in three dimensions. In terms of our postulates, the generalization is as follows.

The Hilbert space is the space of square-integrable, complex valued functions of three variables:

$$|\psi\rangle \leftrightarrow \psi(x, y, z) \equiv \psi(\mathbf{r}).$$

The scalar product is

$$\langle\phi|\psi\rangle = \int_{\mathbf{R}^3} \phi^*(\mathbf{r})\psi(\mathbf{r})\, d^3x \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi^*(x, y, z)\psi(x, y, z)\, dx\, dy\, dz.$$

The meaning of the wave function $\psi(\mathbf{r})$ is that $|\psi(\mathbf{r})|^2$ is the probability density for position measurements to find the particle at the position $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. The probability $P(\mathcal{R})$ for finding the particle in some region $\mathcal{R}$ is

$$P(\mathcal{R}) = \int_{\mathcal{R}} |\psi(\mathbf{r})|^2\, d^3x.$$

Normalization of the wave function now means

$$\int_{\mathbf{R}^3} |\psi|^2\, d^3x = 1.$$

Observables are constructed by tripling everything we did before. The position operator $\hat{\mathbf{r}}$ is[*]

$$(\hat{\mathbf{r}}\psi)(\mathbf{r}) = \mathbf{r}\psi(\mathbf{r}),$$

that is,

$$(\hat{x}\psi)(\mathbf{r}) = x\psi, \quad (\hat{y}\psi)(\mathbf{r}) = y\psi(\mathbf{r}), \quad (\hat{z}\psi)(\mathbf{r}) = z\psi(\mathbf{r}).$$

Momentum operators are

$$(\hat{\mathbf{p}}\psi)(\mathbf{r}) = \frac{\hbar}{i}\nabla\psi,$$

---

[*] Note that we are using the hat ˆ to denote linear operators – not unit vectors!

that is,

$$(\hat{p}_x\psi)(\mathbf{r}) = \frac{\hbar}{i}\frac{\partial\psi}{\partial x}, \quad (\hat{p}_y\psi)(\mathbf{r}) = \frac{\hbar}{i}\frac{\partial\psi}{\partial y}, \quad (\hat{p}_z\psi)(\mathbf{r}) = \frac{\hbar}{i}\frac{\partial\psi}{\partial z}.$$

Let us note that the different components of the position commute, as do the different components of momentum. Moreover, each *canonical pair* of coordinates and momenta has the usual *canonical commutation relation*:

$$[\hat{x}, \hat{p}_x] = i\hbar\hat{1}, \quad [\hat{y}, \hat{p}_y] = i\hbar\hat{1}, \quad [\hat{z}, \hat{p}_z] = i\hbar\hat{1}.$$

Consequently, we can determine all three components of the position of the particle to arbitrary statistical accuracy and likewise for the momentum – but not both, of course. Corresponding components of the position and momentum are incompatible. But non-corresponding components are compatible. So, for example, while there is a non-trivial uncertainty relation for x and p_x it is possible to have states in which both x and p_y are determined with arbitrary statistical accuracy.

The energy operator, *i.e.,* the Hamiltonian, is given by

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \hat{V}(\hat{\mathbf{r}}, t),$$

where

$$\hat{p}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2 = -\hbar^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) = -\hbar^2\nabla^2,$$

so that

$$(\hat{H}\psi)(\mathbf{r}) = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}) + V(\mathbf{r}, t)\psi(\mathbf{r}),$$

where $V(\mathbf{r}, t) \equiv V(x, y, z, t)$ is the potential energy function and ∇^2 is the Laplacian.

A significant new feature of particle motion in three dimensions is that there is a new set observables constituting *angular momentum.* Recall that classically the angular momentum is given by

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}.$$

This vector observable represents three observables:

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x.$$

The quantum operator representative of this triplet of observables is taken to be

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}},$$

that is,

$$(\hat{\mathbf{L}}\psi)(\mathbf{r}) = \mathbf{r} \times \left(\frac{\hbar}{i}\nabla\psi(\mathbf{r})\right) = \frac{\hbar}{i}\mathbf{r} \times \nabla\psi,$$

with, for example,

$$\hat{L}_x\psi = \frac{\hbar}{i}\left(y\frac{\partial\psi}{\partial z} - z\frac{\partial\psi}{\partial y}\right),$$

and so forth.

I will have a lot more to sat about angular momentum later. A good exercise for you at this point is to verify that the ordering of the position and momentum operators does not matter in the angular momentum. So, for example, we could just as well define the operator as $\hat{L} = -\hat{\mathbf{p}}\times\hat{\mathbf{r}}$.

The Schrödinger equation in three dimensions

The Schrödinger equation,

$$\hat{H}\Psi = i\hbar\frac{\partial\Psi}{\partial t}$$

in 3-d is a partial differential equation for $\Psi = \Psi(\mathbf{r}, t)$:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V(\mathbf{r}, t)\Psi = i\hbar\frac{\partial\Psi}{\partial t}.$$

For time independent potentials, $V = V(\mathbf{r})$, separation of the time variable works as before. With

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-\frac{i}{\hbar}Et},$$

we get the time independent Schrödinger equation:

$$\hat{H}\psi = E\psi.$$

As usual, the spectrum of $\hat{H}$ may be discrete, continuous, or both. Normalizable solutions of the TISE – corresponding to eigenvectors of $\hat{H}$ – will have discrete eigenvalues E_n, and corresponding stationary states $\Psi_n(\mathbf{r}, t)$. The corresponding stationary states Ψ_n evolve in time via the (physically irrelevant) phase factor $e^{-\frac{i}{\hbar}E_n t}$.

The stationary states (the energy eigenfunctions) are states where the energy is determined with certainty. The general solution to the SE is given by a superposition over *all* the stationary states (from the whole spectrum). Assuming for simplicity that the spectrum is discrete, we have

$$\Psi(\mathbf{r}, t) = \sum_n c_n\psi_n(\mathbf{r})e^{-\frac{i}{\hbar}E_n t}.$$

The superposition which occurs – *i.e.,* the coefficients c_n – are determined by the initial state in the usual way. With a discrete spectrum and the energy eigenfunctions chosen to form an orthonormal set we have:

$$c_n = \int_{\mathbf{R}^3}\psi_n^*(\mathbf{r})\Psi(\mathbf{r}, 0)\, d^3x.$$

Solving the TISE for central forces

Let us turn to a simple but important class of interactions in which the potential energy function for the particle depends only upon the distance from a given point. Placing our origin at this point the potential energy function depends upon (x, y, z) through a single function of

$$r = \sqrt{x^2 + y^2 + z^2}.$$

I will use the traditional abuse of notation and call this function V as well:

$$V = V(|\mathbf{r}|) \equiv V(r), \quad .$$

These potentials correspond to *central forces*, since the classical force field is given by (exercise)

$$\mathbf{F}(\mathbf{r}) = -\nabla V(r) = -V'(r)\frac{\mathbf{r}}{r}.$$

Three good examples of central force potentials are the *isotropic oscillator*

$$V(r) = \frac{1}{2}m\omega^2 r^2,$$

the *Coulomb* or *Kepler* potential

$$V(r) = \frac{k}{r},$$

and the *Yukawa potential*

$$V(r) = \alpha e^{-\mu r}.$$

The isotropic harmonic oscillator gets its name from the fact that, for any displacement in a radial direction, the restoring force is proportional to the displacement. In particular, isotropic indicates that the restoring force is independent of (radial) direction of displacement. The Coulomb potential is of course quite familiar to you and is the basis for the simplest model of an atom. The Yukawa potential models the strong, short-ranged attractive force between nucleons. Without this force nuclei would not exist.

The nice thing about central potentials is that they allow a straightforward solution of the TISE using spherical coordinates (r, θ, ϕ) and separation of variables. Let us now investigate this.

To begin, the Laplacian in spherical polar coordinates is

$$\nabla^2\psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2\psi}{\partial\phi^2}\right).$$

We use this in the TISE

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi(r, \theta, \phi) + V(r)\psi(r, \theta, \phi) = E\psi(r, \theta, \phi).$$

Note that we use μ to denote the mass of the particle; this will avoid a clash of conventions later. It is not obvious, but this partial differential equation can be solved by separation of variables. So, let us start by attempting to write a solution as a product of radial and angular functions:

$$\psi(r,\theta,\phi) = R(r)Y(\theta,\phi).$$

We substitute this into the TISE, and (i) divide both sides by ψ, (ii) multiply both sides by $\left(-\frac{2\mu r^2}{\hbar^2}\right)$ to find (exercise)

$$\frac{1}{R}\left(r^2 R'\right)' - \frac{2\mu r^2}{\hbar^2}\left[V - E\right] = -\frac{1}{Y\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) - \frac{1}{Y\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}.$$

We see the familiar separation of variables result: the left-hand side is a function of r, the right hand side is independent of r. Therefore, each side must equal a constant, which we denote for later convenience by $l(l+1)$:

$$\frac{1}{R}\left(r^2 R'\right)' - \frac{2\mu r^2}{\hbar^2}[V - E] = l(l+1) = -\frac{1}{Y\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{Y\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}.$$

The strategy now is to solve these two equations separately. Although I won't explain this until later, it will turn out that l must be a non-negative integer,*

$$l = 0, 1, 2, \dots\ .$$

For each choice of that integer there is a solution $(R(r), E)$ of the radial equation. Note that it is the radial equation that will be used to find the allowed energies. Of course, to solve the radial equation the potential energy function has to be specified.

Let us turn the angular equation. Note that it is "universal" in the sense that it is the same for all choices of the central potential energy function, and for all possible energies E. You probably have encountered this equation before. Note, in particular, if we consider the TISE with $E = V = 0$, then we are solving the Laplace equation in spherical polar coordinates. In this case the radial and angular equations are the same as you may have encountered in electrostatics. In the quantum mechanical case the radial equation is different – it depends upon the choice of potential energy and energy eigenvalue, but the angular equation is unchanged. Thus the angular part of the separation of variables type of solution to the TISE for central forces is the same as the angular part of separation of variables solutions to the Laplace equation for an electrostatic potential in a charge free region.

---

* Ultimately, this comes from requiring the angular momentum is represented as a self-adjoint operator.

When we solve the angular equation we are solving

$$-\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)-\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}=l(l+1)Y.$$

This is an eigenvalue equation for the linear differential operator on the right hand side. As mentioned earlier, this operator is the angular part of the Laplacian. We shall see a little later that the angular part of the Laplacian represents the observable $\frac{1}{\hbar^2}L^2$, so the possible values of the magnitude of angular momentum will be $\hbar\sqrt{l(l+1)},\ l=0,1,2,\ldots$.

As you may recall from electrostatics, the angular part of the Laplace eigenvalue problem can also be solved by separation of variables. The result is

$$Y(\theta,\phi)=AP_l^m(\cos\theta)e^{im\phi},$$

where (as already stated) l is a non-negative integer, and m is any integer such that

$$-l\leq m\leq l,\quad \text{i.e.,}\quad m=-l,-l+1,\ldots,l-1,l.$$

The functions P_l^m are the *associated Legendre functions*

$$P_l^m(x)=\frac{1}{2^l l!}(1-x^2)^{|m|/2}\left(\frac{d}{dx}\right)^{|m|+l}(x^2-1)^l.$$

The associated Legendre functions look a little messy when written in this general form, but they are not so bad when taken individually. For example,

$$P_0^0=1,$$

$$P_1^0(x)=x\Longrightarrow P_1(\cos\theta)=\cos\theta,$$

$$P_1^1(x)=P_1^{-1}(x)=(1-x^2)^{1/2}\Longrightarrow P_1^1(\cos\theta)=P_1^{-1}(\cos\theta)=\sin\theta,$$

and so on. See if you can obtain these from the general formula above. See your text for more examples.

It can be shown that the associated Legendre functions $P_l^0(x)$ form a complete set of orthogonal polynomials (like the Hermite polynomials). This means in particular that these functions form a basis for the vector space of square-integrable functions on the interval $[-1,1]$. These polynomial functions $P_l^0(x)$ are usually called the *Legendre functions* or the *Legendre polynomials.*

The integer values of l and m are required so that the wave function is suitably non-singular, single-valued, and define a domain for the angular momentum operators such that they are self-adjoint. The "single-valued" requirement is

$$\psi(r,\theta,0)=\psi(r,\theta,2\pi).$$

This single-valuedness requirement might not appear to be necessary *a priori* (like it would be for electrostatics) since ψ is only defined up to an overall phase, and we only require that ψ be square-integrable. However, in order for ψ to be in the domain of self-adjoint operators H and $\mathbf{L}$ this restriction is in fact needed. It is also for this reason that l should be an integer.

We normalize the solutions of the TISE via

$$1 = \int_0^\infty dr\, |R(r)|^2 r^2 \int_0^{2\pi} d\phi \int_0^\pi d\theta\, |Y(\theta,\phi)|^2 \sin\theta.$$

It is conventional to normalize the radial and angular parts separately:

$$1 = \int_0^\infty dr\, |R(r)|^2 r^2\, dr, \quad 1 = \int_0^{2\pi} d\phi \int_0^\pi d\theta\, |Y(\theta,\phi)|^2 \sin\theta.$$

If we define

$$Y_l^m = \epsilon\sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}} e^{im\phi} P_l^m(\cos\theta), \quad \epsilon = (-1)^m,\ m \geq 0, \quad \epsilon = 1,\ m \leq 0.$$

then the angular functions are normalized. In fact, they form the orthonormal set of functions on the unit sphere called *spherical harmonics*. See table 4.2 in your text for a list of several spherical harmonics. The orthonormality relation for spherical harmonics is

$$\int_0^{2\pi}\int_0^\pi (Y_{l'}^{m'})^* Y_l^m \sin\theta\, d\theta\, d\phi = \delta_{ll'}\delta_{mm'}.$$

Exercise: Show that

$$(Y_l^m)^* = (-1)^m Y_l^{-m}.$$

Answer:

We have

$$(Y_l^m)^* = \left\{\epsilon\sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}} e^{im\phi} P_l^m(\cos\theta)\right\}^*.$$

The normalization coefficient has been chosen real, and the associated Legendre functions are real, hence

$$\begin{aligned}(Y_l^m)^* &= \epsilon\sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}} \left\{e^{im\phi}\right\}^* P_l^m(\cos\theta)\\ &= \epsilon\sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}} e^{-im\phi} P_l^m(\cos\theta)\\ &= Y_l^{-m}\end{aligned}$$

The radial equation

We now have a look at the radial equation,

$$\left(r^2R'\right)' - \frac{2\mu r^2}{\hbar^2}(V-E)R = l(l+1)R.$$

The detailed nature of the solutions will depend upon the form of the potential energy and on the boundary conditions, but we can make some progress without committing to either of these. Make a change of dependent variables

$$u = rR,$$

and you will find that u satisfies

$$-\frac{\hbar^2}{2\mu}u'' + \left(V + \frac{\hbar^2 l(l+1)}{2\mu r^2}\right)u = Eu.$$

This is *mathematically* the same as the 1-d TISE for a "particle" moving on the positive real line in a potential

$$V_{\text{eff}} = V + \frac{\hbar^2 l(l+1)}{2\mu r^2}.$$

Indeed, our normalization convention for R becomes

$$\int_0^\infty |u|^2\,dr = 1.$$

Without loss of generality you can assume that u (and hence R) is a real function. You can think of u as defining the radial position probability distribution for a particle moving in a central force field.

The *effective potential* V_{eff} governing the radial part of the wave function includes the central force potential and the "centrifugal potential" $\frac{\hbar^2 l(l+1)}{2\mu r^2}$. You can check that the centrifugal potential corresponds (classically) to a repulsive force. This kind of effective potential also arises in classical mechanics. It is worth spending a moment to recall this.

In classical mechanics, a particle moving in a central force field always moves in a plane orthogonal to the conserved angular momentum vector $\mathbf{L}$. If you introduce polar coordinates (r,ϕ) in the plane orthogonal to $\mathbf{L}$, the equations of motion for $(r = r(t), \phi = \phi(t))$ are

$$\mu\frac{d^2r}{dt^2} = -\frac{d}{dr}\left(V(r) + \frac{L^2}{2\mu r^2}\right),$$

and

$$\frac{d}{dt}(\mu r^2\frac{d\phi}{dt}) = 0.$$

The latter equation is conservation of the component of angular momentum perpendicular to the plane,

$$L = \mu r^2 \frac{d\phi}{dt},$$

which is the only component of **L**, so this is conservation of angular momentum. The first "radial equation" governs the radial motion. You see that the effective potential energy governing the radial motion involves a centrifugal term representing the effect of the angular motion on the radial motion. This classical picture suggests that in the quantum theory we might identify $\hbar^2 l(l+1)$ with the squared-magnitude of the angular momentum in the quantum description. This is correct, and will be shown later.

To proceed further, we need to specify the central force potential energy function. Let us briefly consider a couple of elementary examples, and then have a more detailed look at the Coulomb potential.

Example: The free particle in 3-d

A free particle certainly has a central force (rather trivially), since $V(\mathbf{r}) = 0$. The radial equation is

$$-\frac{\hbar^2}{2\mu}u'' + \frac{\hbar^2 l(l+1)}{2\mu r^2}u = Eu.$$

This differential equation is a form of the *spherical Bessel equation.* The solutions, labeled by l, are r times the *spherical Bessel functions* and *spherical Neumann functions* of order l. Only the spherical Bessel functions are non-singular as $r \to 0$, so we only use them. They are denoted by $j_l(x)$. In terms of these special functions, the solution to the radial equation is $u = rj_l(kr)$, where

$$k = \frac{1}{\hbar}\sqrt{2\mu E}$$

can take any non-negative value.

The spherical Bessel functions are oscillatory in functions with decreasing amplitude as the argument ($x = kr$) increases. They all vanish at $r = 0$ except the function defined by $l = 0$ (see below for more on this function). For details on the behavior of the spherical Bessel functions see the text. We will look at the simplest of them in a moment.

Keeping in mind that $u = rR$, we have then the solution to the TISE of the form

$$\psi(r, \theta, \phi) = Aj_l(kr)Y_l^m(\theta, \phi).$$

Note that these solutions are determined by (i) a choice of E (through k), (ii) a choice of l (non-negative integer) and (iii) a choice of m (an integer such that $-l \leq m \leq l$). Thus there is quite a bit of *degeneracy*, that is, there are many solutions (idealized stationary states) with the same energy. This corresponds to the fact in classical mechanics that a free particle can move in a variety of ways (same speed, different directions) and still have the same energy (exercise). We saw (in a homework problem) that for *normalizable* solutions of the TISE in 1-d there can be no degeneracy. Here this result doesn't apply for two reasons. First, the TISE was solved in 3-d. Still, one might argue, the radial equation was equivalent to a TISE in 1-d. We escape that objection by pointing out that (1) the degeneracy does not appear via the radial equation, for each fixed l and E there is only one solution to that equation, (2) the solutions in this example cannot, in any case, be normalizable (since, *e.g.,* the spectrum of H is continuous).

We will content ourselves with examining the simplest of the solutions to the TISE. Suppose we choose $l = 0$. Then the radial TISE is (exercise)

$$-\frac{\hbar^2}{2\mu}u'' = Eu.$$

This is the familiar harmonic oscillator equation, with solutions

$$u = A\cos(kr) + B\sin(kr).$$

Since $R = \frac{u}{r}$, we will only get a non-singular solution to the TISE if $A = 0$. The spherical Bessel function of order 0 is in fact

$$j_0(kr) = \frac{\sin(kr)}{kr},$$

in agreement with our direct solution when $l = 0$.

When $l = 0$ we must have $m = 0$ (exercise), in which case the angular functions are just a constant so the corresponding solution of the TISE with energy E is of the form

$$\psi_{k,l=0=m} = A\frac{\sin(kr)}{kr}.$$

Thus the probability distribution is spherically symmetric, and oscillatory with amplitude decreasing with increasing radius. The frequency of the oscillation increases with increasing energy. You can check that this solution is not normalizable, as expected (exercise).

It is interesting to think about what the particle is actually "doing" in this family of stationary states, especially in comparison to the classical mechanics behavior. What do we know? Well, we have a free particle with some statistically definite energy and vanishing angular momentum (relative to the origin). In classical mechanics with that information, since $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, the particle is either at rest (in which case the energy vanishes) or its momentum is parallel to $\mathbf{r}$. Just given the energy and vanishing of angular momentum, any radial motion at constant speed could occur. Of course, we know that in classical mechanics the particle traces out a definite path from one of the possibilities. We could in principle use additional measurements in position and momentum to prepare the particle in a state where it follows a single radial trajectory with the given energy. In the quantum state you can imagine similar possibilities for motion given the energy and angular momentum. But you should keep in mind that the state is already completely determined by the specification of energy and vanishing angular momentum, so the particle cannot really be said to take this or that radial path.* Indeed, you have to keep in mind the uncertainty relation between position and momentum which prohibits the classical path of motion to be determined with statistical certainty. Indeed, in this (idealized) stationary state the angular momentum is determined with certainty, which implies maximum uncertainty in angular position. The angular momentum and linear momentum are also incompatible, leading to maximum uncertainty in the direction of motion. So, while the particle is in some sense undergoing radial motion, the position and direction of motion are not determined! One can also prove that in this state (with proper care about its non-normalizability) the radial motion is equally likely to be toward the origin as away from it. One way to check this is to compute the integrand for, say, the expectation value for the x component of momentum. You can check that this integrand is an odd function of x. Consequently any symmetric integral over an arbitrarily large (but finite) region of this integrand will vanish. One can conlude that the expectation value of momentum is zero, even though the magnitude of the momentum need not be zero. This is consistent with equal probability for motion toward or away from the origin. Finally, I remind you that this idealized state is an eigenfunction of energy and hence determines an idealized stationary state. In contrast to the classical motion, no observable quantities for the particle will change in time in this state. The probability current vanishes for this wave function.

Particle in a spherical box

Let us now consider the stationary states of a particle confined to a region $r < a$. The potential energy function is that of an "infinite spherical well". We can view the free particle of the last section as a limiting case $a \to \infty$. We set $\psi(r, \theta, \phi) = 0$ when $r > a$. Inside the well, the radial equation is exactly as before. The only difference now is that we must take account of the boundary condition that

$$R(a) = 0.$$

In general, this leads to a transcendental equation involving the spherical Bessel functions:

$$j_l(ka) = 0.$$

---

* Notice that the wave function is spherically symmetric, so perhaps it is not surprising to learn that all directions of motion are equally likely.

The zeros of the spherical Bessel functions are known; they form a discrete set. Thus k is determined by these discrete values. In this way we get a discrete set of allowed energies, as expected.

Let us revisit the $l = 0$ case. We found that

$$R(r) = A\frac{\sin(kr)}{kr}$$

solves the radial equation (for $r < a$). To impose the boundary condition is now easy (exercise)

$$k = \frac{n\pi}{a}, \quad n = 1, 2, \dots .$$

(Exercise: why do we exclude $n = 0, -1, -2, \dots$?). The allowed energies are

$$E_{n,l=0} = \frac{n^2\pi^2\hbar^2}{2\mu a^2},$$

which are identical to the 1-d particle in a box! This is really not so surprising. By ignoring angular motion effects ($l = 0$), we have essentially reduced the problem to a 1-d problem (exercise). The normalized stationary states that have this energy are (exercise)

$$\psi_{n,l=0} = \frac{1}{\sqrt{2\pi a}}\frac{1}{r}\sin(\frac{n\pi r}{a}).$$

Keep in mind that these states and their energies correspond to choosing $l = 0$. For each $l \neq 0$ there will also be a sequence of allowed energies; these energies will depend upon the choice of l and another integer labeling the zeros of the spherical Bessel function. This integer is usually called the *principal quantum number*, n. When $l \neq 0$, it is also possible to have stationary states with $m \neq 0$. Because m does not enter the radial equation, it will play no role in determining the energy. This is quite general: associated with each l there will be $2l+1$ states, corresponding to $m = -l, -l+1, \dots, l-1, l$ (exercise), all having the same energy. One says that a stationary state with a given value of l is "$(2l + 1)$-fold degenerate". This degeneracy reflects the rotational invariance of the system.

Hydrogen atom

The simplest successful model of a hydrogen atom is the quantum mechanical system associated with the central force potential

$$V(r) = -\frac{e^2}{4\pi\epsilon_0}\frac{1}{r}.$$

Here e is the charge on the electron, r is the value of the observable which can be interpreted as "distance from proton to electron", and ϵ_0 is the "permittivity" of the vacuum, familiar from electrostatics.

The Coulomb potential being used here is everywhere negative. It diverges as $r \to 0$, and it vanishes as $r \to \infty$. The behavior as $r \to 0$ suggests, at least classically, that the system is not stable since the energy is unbounded from below. The behavior as $r \to \infty$ means that positive energy, unbound (scattering) states can exist. This is good, since our model should be able to describe ionized states and/or scattering states. We shall restrict our attention to the bound states, so as to understand the famous success of quantum mechanics in predicting the spectrum of hydrogen.

The TISE for the hydrogen atom separates into the angular equations, which we solve using spherical harmonics, and the radial equation:

$$-\frac{\hbar^2}{2\mu}u'' + \left[-\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{\hbar^2}{2\mu}\frac{l(l+1)}{r^2}\right]u = Eu.$$

The effective potential energy for the radial equation includes a "centrifugal barrier" when $l \neq 0$. Recall that a similar term arises in classical mechanics and guarantees that bound states are kept away from $r = 0$. When $l = 0$ the classical particle may reach $r = 0$. We shall see how these features fare in the quantum domain.

In the radial equation above, the mass μ can be interpreted as the *reduced mass*

$$\mu = \frac{m_e m_p}{m_e + m_p},$$

built from the electron mass m_e and the proton mass m_p. Because $m_p >> m_e$ we have $\mu \approx m_e$, but one must eventually go beyond this approximation to get quantitative agreement with atomic spectra.

We seek solutions to the radial equation that are well behaved everywhere and that vanish as $r \to \infty$, so that the solutions define normalizable functions, *i.e.,* bound states. As usual, we must also determine the allowed values of the energy E; these energy eigenvalues will form a discrete spectrum since we are looking for normalizable solutions. The text provides all the gory details; we shall be content with understanding the results.

Energy spectrum

To begin, the allowed energies are labeled by a positive definite integer $n = 1, 2, \ldots$. This integer is called the "principal quantum number". In terms of it the energy eigenvalues are

$$E_n = -\left\{\frac{\mu}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right\}\frac{1}{n^2}.$$

We can write this as (exercise)

$$E_n = -\frac{e^2}{8\pi\epsilon_0 n^2 a},$$

where

$$a = \frac{4\hbar^2 \pi \epsilon_0}{\mu e^2} \approx 0.53 \times 10^{-10} m.$$

The quantity a, which has dimensions of length, is the *Bohr radius.* It gets its name because, in the Bohr model of the atom, this is the radius a classical electron must maintain to be in the ground state. Indeed, you can check that when $n = 1$ the atom is in its lowest allowed energy state and the resulting energy is precisely that of a classical particle in a circular orbit at the Bohr radius. More generally, the energies for any n are the same you would get for a classical particle at the radius na. This is why the Bohr model was successful for the hydrogen atom. The Bohr model is now known to be physically incorrect as a model of nature since (a) it fails to describe any other atom, and (b) it does not provide a satisfactory means for describing any phenomenon besides the energy levels of hydrogen. The energy, expressed in terms of the principal quantum number and the Bohr radius is, nevertheless, often called the *Bohr formula* for the hydrogen energies. I emphasize that this formula was obtained by Bohr by using a voodoo blend of classical mechanics and preliminary ideas from the – not yet formed – quantum theory. It is unfortunate that modern textbooks propagate this intermediate step (Bohr theory) on the road to quantum theory as if it were a viable model of nature*.

The ground state, when $n = 1$, has energy $E_1 = -13.6\, eV$. Keep in mind that the potential energy is being measured relative to $r \to \infty$ and is always negative. The total energy we have found is negative, but this is just for the bound states; positive energies correspond to unbound (scattering states). This ground state energy, therefore, is the energy needed to ionize the atom, in good 1 agreement with experiment. For this reason, $|E_1|$ is often called the *binding energy* of the atom in its ground state. Excited states have, of course, greater energy and hence *less* binding energy.

Stationary states

The stationary states are, of course, labeled by n. But they also depend upon l and m. One can anticipate a relation between l and n since l appears in the radial equation. What one finds when solving the radial equation and imposing boundary conditions is that, for each value of n, one can have any value of l between 0 and $n-1$, inclusive:

$$l = 0, 1, 2, \ldots, n-1.$$

---

* In my mind, using the Bohr model to help explain the quantum theory of an atom is analogous to using the Ptolemaic model of the universe (geocentric, epicycles and all that) as a viable model of the universe just because it appeared on the road from the theory in which heavenly bodies were gods to the Copernican theory of the solar system. I concede that I might be a little too zealous in this point of view.

Keep in mind also that for each l there are the $2l+1$ allowed values of m. Thus, the stationary states are labeled as follows. Pick any $n>0$ this determines the "energy level"; pick any integer l such that $0 \leq l \leq n-1$; pick any integer m such that $-l \leq m \leq l$. As we shall see, the choices of l and m correspond to specifying the angular momentum of the stationary state.

With this algorithm in mind, the orthonormal basis of stationary states is, in all of its glory:

$$\psi_{nlm}(r,\theta,\phi) = N_{nlm} e^{-\frac{r}{na}} \left(\frac{2r}{na}\right)^{l} L^{2l+1}_{n-l-1}(2r/na) Y^m_l(\theta,\phi).$$

Here N_{nlm} is a normalization constant (see the text for a formula) and L^s_r is the *associated Laguerre polynomial*, defined by

$$L^s_r(x) = (-1)^s \frac{d^s}{dx^s}\left(e^x \frac{d^{r+s}}{dx^{r+s}}(e^{-x}x^{r+s})\right).$$

Note that this special function appears in ψ_{nlm} with $x = \frac{2r}{na}$.

The radial dependence of the stationary states is that of a polynomial in r times a decaying exponential $e^{-\frac{r}{na}}$ in r. Here are a few of the associated Laguerre polynomials:

$$L^0_0(x) = 1, \quad L^0_1(x) = 1-x, \quad L^2_2(x) = 12x^2 - 96x + 144;$$

See your text for some more. Such functions – polynomials times decaying exponentials – are normalizable because of the decaying exponential. Note that the exponential decay of the wave function is characterized by n times the Bohr radius, so we can say that measurements of the "size" of the atom (via observables that characterize the position probability distribution) will yield values on the order of magnitude of na.

One amusing feature of the stationary states stems from the fact that the states with $l=0$ have wave functions that do not vanish at the origin (exercise), since $R(0) \neq 0$. For example, in its ground state we have

$$\psi_{n=1,l=0,m=0} = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}.$$

When $l \neq 0$ it is not hard to see that $R(0)=0$ (exercise). Does this means that the atom has (in states with $l=0$) a non-zero probability for finding the electron occupying the same space as the proton? Let's see. The probability $P(\epsilon)$ for finding a ground state electron in, say, a spherical region of radius ϵ about the origin is

$$P(\epsilon) = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \int_0^{\epsilon} dr \, |\frac{1}{\sqrt{\pi a^3}} e^{-r/a}|^2 r^2 \sin(\theta) = 1 - \left[2\left(\frac{\epsilon}{a}\right)^2 + 2\left(\frac{\epsilon}{a}\right) + 1\right] e^{-2\epsilon/a}$$

As $\epsilon \to 0$ this probability vanishes. On the other hand, there is a real sense in which the nucleus (here, a proton) has a finite size (on the order of Fermis). From the preceding computation we can expect that the probability for finding the electron at the "surface" of the proton is non-zero, and by continuity we expect that probability to remain non-zero inside the proton. Explicit computations with a finite-sized charge density confirm this. Thus, figuratively speaking, it is possible that the electron is inside the proton! Again we see the tension between the quantum meaning of the word "particle" and the classical, macroscopic meaning. Because the nucleus has a finite size, the overlap between the electron wave function and the nuclear charge density indicates that we could improve our description of the atom by modifying the potential energy function ($1/r$ outside the proton, r^2 inside). This results in a small change in the $l = 0$ energy spectrum (thus partially lifting some degeneracy (see below)). For a hydrogen atom this effect is very small, on the order of $10^{-9}eV$. For "hydrogenic atoms" with larger atomic number, the effect is more pronounced. This effect has been measured experimentally.

More on the spectrum of hydrogen

There is a good bit of degeneracy in the energies of hydrogen. Recall that an eigenvalue is called *p-fold degenerate* if there are p linearly independent eigenvectors with that eigenvalue. Physically, this means that the outcome of a measurement corresponding to that eigenvalue can occur with probability one via a p-dimensional set of states. For the hydrogen atom, the energy is only determined by n. For a given n, there are n possible l values (exercise). For each l there are $2l+1$ possible m values. All together, for each energy one has (exercise) n^2 different (*i.e.,* linearly independent) energy eigenfunctions obtained by varying l and m in their allowed ranges. The freedom to vary m reflects the rotational symmetry of the problem: by rotating the atom we can obtain new states with the same energy. The freedom to vary l is a special feature of the Coulomb potential and reflects a "hidden symmetry" in the Coulomb (and Kepler) dynamical motion. Classically, this hidden symmetry is responsible for the fact that the motion of a particle in a $\frac{1}{r^2}$ central force field has bounded orbits that are closed (in fact, ellipses).*

One can probe the energy levels of hydrogen experimentally by adding energy to a hydrogen atom and looking at the electromagnetic radiation that comes out when the atom decays. For example, one can pass current through a container containing hydrogen gas. This has the effect of transferring energy to and from to the atom. The atom absorbs energy from the perturbation and moves to an excited state. The perturbation will then

---

* You may have encountered the "Laplace-Runge-Lenz" vector in the study of the 2-body inverse-square central force problem in classical mechanics. This conserved vector, like all conservation laws, comes from a symmetry and it is this symmetry which is responsible for the degeneracy of energies in hydrogen associated to varying l for a given n.

cause the atom to emit energy (via one or more photons) as it makes a transition to a state of lower energy. This process is called "stimulated emission". This picture of stimulated emission is understood by looking at the time dependent Schrödinger equation. We will develop techniques for understanding stimulated emission via the SE next semester. (Why the atom decays at all in the absence of a perturbation (spontaneous emission) is another story. . . wait until next semester.)

The following ideas will be given a full quantum mechanical treatment next semester. The bottom line is as follows. The atom will absorb any energy equal to the difference between any two allowed energies. The effect is to raise that atom into a higher energy state. After a certain amount of time, the atom will decay to a lower energy state by emission of a photon. The energy of this photon is the energy difference between the stationary states before and after the emission of the photon. Typically we measure the energy of the photons by using a spectrometer, which discerns the wavelength λ of the photon. The relation between the energy level difference

$$\Delta E = (\frac{1}{n_{final}^2} - \frac{1}{n_{initial}^2})(-13.6\ eV),$$

and the photon wavelength is (c is the speed of light)

$$\Delta E = \frac{2\pi c\hbar}{\lambda}.$$

Transitions from excited states to the ground state ($n_{final} = 1$) give off the most energy; the emitted light is ultraviolet. Transitions to the first excited state yield visible light. Transitions down to the second excited state give off infrared radiation, *etc.* The wavelengths predicted by quantum mechanics are in very good agreement with experiment. There *are* differences, however, since there are other physical effects not accounted for in this simplest of models of the hydrogen atom.

Angular momentum

So far we have focused on the stationary states of a particle moving in a central force field. In particular, we have focused on the energies of the hydrogen atom. Another very important observable that one can study for a particle moving in three dimensions is its angular momentum. We shall spend a fair amount of time talking about this observable.

Recall that we defined the operator representative of angular momentum via

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}.$$

What this means is that, on a suitably well-behaved wave function ψ, the linear operation is

$$\hat{\mathbf{L}}\psi = \frac{\mathbf{\hbar}}{\mathbf{i}}\mathbf{r} \times \nabla\psi.$$

More explicitly,

$$\begin{aligned}
L_x &= \frac{\hbar}{i}(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}),\\
L_y &= \frac{\hbar}{i}(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}),\\
L_z &= \frac{\hbar}{i}(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}).
\end{aligned}$$

Note that, as is always the case with the cross product (and curl), once you know the first of these formulas the other two follow by *cyclic permutations.*

Probably the most important aspects of angular momentum are (1) it is conserved by central forces (more generally, we believe that the total angular momentum of a closed system is always conserved); (2) the 3 components of **L** are *incompatible*, that is, they do not commute. To establish (1) is to check

$$[\hat{\mathbf{L}}, \frac{\hat{p}^2}{2m} + \hat{V}(r)] = 0,$$

which you will do in your homework. As for (2), a direct computation shows that

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= i\hbar\hat{L}_z,\\
[\hat{L}_y, \hat{L}_z] &= i\hbar\hat{L}_x,\\
[\hat{L}_z, \hat{L}_x] &= i\hbar\hat{L}_y.
\end{aligned}$$

These are the very important *angular momentum commutation relations*, or *angular momentum algebra.* Note again that any two formulas follow from a third by cyclic permutations. If you have ever studied the Lie algebra of infinitesimal rotations, you will notice a similarity with the angular momentum algebra. This is no accident, the linear transformation defined by the angular momentum operators can be viewed as the transformation of a state under an infinitesimal rotation.

Let me show you how to prove one of these. Keeping in mind that the only non-vanishing commutators among position and momentum are $[\hat{x}^i, \hat{p}_j] = i\hbar\delta^i_j\hat{1}$, that the commutator is anti-symmetric and linear in its arguments, and the identity

$$[A, BC] = [A, B]C + B[A, C],$$

we have

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z]\\
&= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z]\\
&= \hat{y}[\hat{p}_z, \hat{z}]\hat{p}_x + \hat{x}[\hat{z}, \hat{p}_z]\hat{p}_y\\
&= -i\hbar\hat{y}\hat{p}_x + i\hbar\hat{x}\hat{p}_y\\
&= i\hbar\hat{L}_z.
\end{aligned}$$

Because angular momentum is conserved for central potentials, the angular momentum is a useful observable for describing states and their dynamical evolution in this setting. The incompatibility of the components of **L** means that angular momentum is going to have some surprising properties relative to its classical behavior (recall that position and momentum are also incompatible, hence the uncertainty principle, *etc.*) In particular, we will find that *the possible values of any component of angular momentum do not form a continuum but rather a discrete set.* The spectra of the angular momentum operators (L_x, L_y, L_z) are discrete. One says that angular momentum is "quantized".

Because of the incompatibility of (L_x, L_y, L_z), it is generally impossible to have a particle in a state in which two or more components of **L** are known with certainty, unless the angular momentum is in fact zero. To see this, simply apply our general form of the uncertainty principle to, say, L_x and L_y to find (exercise)

$$\sigma_{L_x}\sigma_{L_y} \geq \frac{\hbar}{2}|\langle L_z\rangle|.$$

As long as the state is such that $\langle L_z\rangle \neq 0$, the variances σ_{L_x} and σ_{L_y} cannot vanish, as they should if one knows either of these observables with probability unity. Indeed, suppose that the state $|\psi\rangle$ is an eigenvector of L_x and L_y, then it is easy to see that (exercise)

$$[L_x, L_y]|\psi\rangle = 0.$$

But, by the angular momentum commutation relations, this implies

$$L_z|\psi\rangle = 0.$$

So, in fact, the state must be a state of definite value for all three components. By using the other two commutation relations you can then show that

$$L_x|\psi\rangle = 0 = L_y|\psi\rangle,$$

so that the angular momentum is zero with probability unity in such a state.

Each of the components (L_x, L_y, L_z) has its basis of eigenvectors, where the value of that observable is known with probability one. But because the three angular momentum operators do not commute, these bases are not the same. If the particle is in a state in which the angular momentum component, say, L_x is known to be (non-zero) with certainty, then in general there will be statistical uncertainty in the values of L_y and L_z. When working in a basis of eigenvectors of angular momentum the component which is determined with certainty is conventionally chosen to be L_z. There is nothing special about the z-axis, since we can call any direction in space the z direction.

Classically we are used to thinking of the angular momentum as a vector. Quantum mechanically, there are some vectorial aspects to angular momentum (*e.g.,* it has three components), but it is impossible to say with certainty which way the vector points. As it turns out, the length of the angular momentum vector and any one component (*e.g.,* L_z) can be determined with statistical certainty. This is because the operator representing the length-squared,

$$L^2 = L_x^2 + L_y^2 + L_z^2,$$

commutes with any component of $\mathbf{L}$. For example (exercise),

$$\begin{aligned}[L_z, L^2] &= [L_z, L_x^2 + L_y^2 + L_z^2] \\ &= [L_z, L_x]L_x + L_x[L_z, L_x] + [L_z, L_y]L_y + L_y[L_z, L_y] \\ &= i\hbar\left(L_yL_x + L_xL_y - L_xL_y - L_yL_x\right) \\ &= 0.\end{aligned}$$

Recall that commuting Hermitian operators always admit a simultaneous basis of eignvectors. Thus it is possible to find a basis of vectors that are eigenvectors of both one component (conventionally taken to be L_z) and L^2. The states defined by these *angular momentum eigenvectors* are states in which the value of L_z and L^2 are known with probability unity. Our task now is to determine these eigenvectors. We will see that they correspond to the spherical harmonics.

Angular momentum eigenvalues and eigenfunctions

It is possible to systematically deduce the eigenvalues of L^2 and L_z just using the angular momentum commutator algebra. This is logically the same game we played when finding the spectrum of the harmonic oscillator Hamiltonian using raising and lowering ("ladder") operators. See the text for details. The result is that the operator L^2 has eigenvalues $\hbar^2 l(l+1)$, where $l = 0, 1, 2, \ldots$. Given the value l, any given component, say L_z, has eigenvalues $m\hbar$, where $m = -l, -l+1, \ldots, l-1, l$. This is true for L_x and L_y, too. The catch is that the eigenfunctions for each component are not, in general, the same. We can pick any one component, find its eigenfunctions, and arrange it so that these eigenfunctions are also the eigenfunction of L^2. It turns out that the simultaneous eigenfunctions of L_z and L^2 are given by any function of r times the spherical harmonics Y_l^m.

Let us see a little bit more how this works. If we write the angular momentum operators $\mathbf{L}$ and L^2 in terms of spherical polar coordinates we find (via a nice, but long exercise)

$$L_x = \frac{\hbar}{i}\left(-\sin\phi\frac{\partial}{\partial\theta} - \cos\phi\cot\theta\frac{\partial}{\partial\phi}\right).$$

$$L_y = \frac{\hbar}{i}\left(\cos\phi\frac{\partial}{\partial\theta} - \sin\phi\cot\theta\frac{\partial}{\partial\phi}\right).$$

$$L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi}.$$

You can see why one likes to use the component L_z; the spherical polar coordinates are adapted to the z and L_z takes a very simple form. You really should try to at least verify the formula for L_z.

It is easy to see that eigenfunctions of L_z are any function of r and θ times $e^{im\phi}$:

$$\psi_m(r,\theta,\phi) = F(r,\theta)e^{im\phi}, \quad m = 0, \pm1, \pm2, \ldots,$$

$$\hat{L}_z\psi_m = m\hbar\psi_m.$$

Here m is any integer. The restriction to integer values is so that the eigenfunction is a well-defined function in space. You can see that there is a lot of degeneracy here. Since F is arbitrary, there are many, many functions that have the same eigenvalue $m\hbar$ for L_z. We can narrow down the possibilities by demanding that ψ_m also be an eigenfunction of L^2.

The operator L^2 is a little more complicated to write out in spherical polar coordinates. Up to a factor of r^2, it turns out to be the angular part of the Laplacian:

$$L^2\psi = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\left(\frac{\partial^2\psi}{\partial\phi^2}\right)\right].$$

This differential operator is the restriction of the Laplace operator to a (unit) sphere (exercise). To find the eigenfunctions is a story we have already touched upon when we solved the TISE for central forces via separation of variables (exercise). Indeed, we recover our results from that analysis. In particular, we have that (i) the eigenvalues of L^2 are $l(l+1)\hbar^2$, where $l = 0, 1, \ldots$; (ii) for a given choice of l, the eigenvalues of L_z are restricted by $m = -l, -l+1, \ldots, l-1, l$; (iii) the eigenfunctions $\psi_{l,m}(r,\theta,\phi)$ for a given choice of l and m are built from the spherical harmonics Y_l^m:

$$\psi_{l,m}(r,\theta,\phi) = f(r)Y_l^m(\theta,\phi).$$

$$\hat{L}^2\psi_{l,m} = l(l+1)\hbar^2\psi_{l,m}.$$

We see that there is still some degeneracy owing to the appearance of the arbitrary function $f(r)$. Thus, many different states correspond to the same values of L^2 and L_z. These states all differ by the choice of $f(r)$. To uniquely specify the states one must find another observable which commutes with L_z and L^2 and demand that $\psi_{l,m}$ also be an eigenfunction of that operator with a specific eigenvalue. For example, in the central force problems it turns out (see homework) that

$$[H, \mathbf{L}] = [\frac{p^2}{2m} + V(r), \mathbf{L}] = 0.$$

This fact is equivalent to the statement that the Hamiltonian operator is rotationally invariant. As you will see in a homework problem, the fact that H commutes with $\mathbf{L}$ means that angular momentum is conserved (probabilities for outcomes of momentum measurements will not change in time.) This is important in the present discussion because it means that it is possible to further specialize the eigenfunctions of angular momentum to also be energy eigenfunctions. How is this done? Well, you have already done it! The stationary states in a central force problem are radial functions times the spherical harmonics, as above. The radial functions are determined by the radial equation, which also determines the allowed energies. Thus $f(r)$ can be determined by a choice of energy. Indeed, the stationary states of the hydrogen atom (more generally, any central force system) are simultaneously (i) energy eigenfunctions, (ii) eigenfunctions of a component (L_z) of angular momentum and (iii) eigenfunctions of the total angular momentum L^2. We now see the physical significance of l and m.

Infinitesimal translations and rotations

There is a nice (if a little sophisticated) way to think of momentum-like quantities in classical and quantum mechanics, namely, as "generators" of transformations. In fact, linear momentum generates translations while angular momentum generates rotations. Let us briefly explore this.

We begin with translations. Let $\psi(\mathbf{r})$ describe the state of a system. How does the value of ψ change as we "translate" from $\mathbf{r}$ to the point $\mathbf{r}+\mathbf{a}$? Simple:

$$\psi(\mathbf{r}) \to \psi(\mathbf{r}+\mathbf{a}).$$

Now, we can built any translation by a sequence of many small translations. Suppose that the magnitude a is very small (infinitesimal), then a standard result from calculus is that

$$\psi(\mathbf{r}+\mathbf{a}) \approx \psi(\mathbf{r}) + \mathbf{a}\cdot\nabla\psi(\mathbf{r}) = \psi(\mathbf{r}) + \frac{i}{\hbar}\mathbf{a}\cdot(\hat{\mathbf{p}}\psi)(\mathbf{r}).$$

We say that the momentum is the *infinitesimal generator of translations*, since a finite translation can be built by many infinitesimal translations associated to $\mathbf{p}$ as above. In general, the infinitesimal change in a wave function associated with a translation along the infinitesimal vector $\mathbf{a}$ is given by $\frac{i}{\hbar}\mathbf{a}\cdot\mathbf{p}\psi(\mathbf{r})$.

A similar result is available for angular momentum and infinitesimal rotations. Every rotation is a rotation by some angle about some axis. For a given rotation, we can choose our z axis to be along the axis of rotation. In these coordinates a rotation by an angle ϕ about z is (exercise)

$$x \to x' = x\cos\phi - y\sin\phi,$$

$$y \to y' = y\cos\phi + x\sin\phi,$$

$$z \to z'.$$

Again, we can think of any rotation as the net result of a sequence of many small rotations. Suppose $\phi << 1$, then to first order in ϕ (exercise)

$$x' \approx x - \phi y, \quad y' \approx y + \phi x.$$

Consider a function $\psi(\mathbf{r})$, how does it change under an infinitesimal rotation? We have (exercise)

$$\psi(x', y', z') \approx \psi(x, y, z) - \phi y \frac{\partial\psi(x, y, z)}{\partial x} + \phi x \frac{\partial\psi(x, y, z)}{\partial y}.$$

We can write this as

$$\psi(x', y', z') \approx \psi(x, y, z) + \frac{i}{\hbar}\phi(\hat{L}_z\psi)(x, y, z).$$

Better yet, we can write

$$\psi(x', y', z') \approx \psi(x, y, z) + \frac{i}{\hbar}\phi\mathbf{n}\cdot(\hat{\mathbf{L}}\psi)(x, y, z),$$

where $\mathbf{n}$ is a unit vector along the rotation axis. We say that angular momentum is the infinitesimal generator of rotations since $\mathbf{n}\cdot\mathbf{L}$ gives the infinitesimal change in the wave function due to a rotation about the axis along $\mathbf{n}$. A finite rotation can be built by many infinitesimal rotations, generated by $\mathbf{L}$.

Physically, you can think of the translations and rotations as transformations of the physical system (here a particle). Thus the operators that represent momentum and angular momentum give the infinitesimal change in the state of the system as one performs the indicated transformation (translation and/or rotation) of the system.

Exercise: What operator is the infinitesimal generator of time translation of a given system?

The Stern-Gerlach experiment

I am hoping you have had an introduction to the phenomenology of "intrinsic spin" in an earlier class in modern physics. Here I just want to introduce you to the quantum mechanical basic formalism and a couple of nice results.

To begin, let me describe an experiment conducted by Stern and Gerlach in the early 1920's. It gives us a valuable demonstration of the kind of phenomenon that needs quantum mechanics to explain it.* It also provides an example of what is probably the simplest possible quantum mechanical model. As you probably know, this experiment involves the property of particles known (perhaps misleadingly) as their *spin*, which is an intrinsic angular momentum possessed by the particles. Note, though, at the time of the experiment, neither intrinsic spin nor quantum mechanics was very well understood!

The Stern-Gerlach experiment amounts to passing a beam of particles through a region with a magnetic field which has the same direction everywhere, but a varying magnitude. Recall that the classical potential energy of interaction of a magnetic moment $\vec{\mu}$ with a magnetic field $\vec{B}$ is $-\vec{\mu}\cdot\vec{B}$. Thus the force (as opposed to the torque) exerted on the particle with a magnetic moment is non-zero if the magnetic field varies in space. We have

$$\vec{F} = \nabla(\vec{\mu}\cdot\vec{B}).$$

If the magnetic field varies only in one direction, then the force is parallel or anti-parallel to that direction, depending upon the orientation of the magnetic moment relative to the fixed direction of $\vec{B}$. Thus the force experienced by the particle with a magnetic moment $\vec{\mu}$ allows us to measure the component of $\vec{\mu}$ along the direction of $\vec{B}$; we simply have to observe which way the particles are deflected. Of course, this reasoning was purely classical, so we are initially confident it will work with macroscopic objects, but it can be justified/explained quantum mechanically. For now we just note that the correct quantum model of the interaction of a magnetic moment with a magnetic field does indeed use the potential energy function shown above, and since the atom is sufficiently massive its motion can be well approximated using a classical mechanics description. Here we have a nice example of how classical, macroscopic reasoning gives us important clues to modeling a quantum mechanical, microscopic system.

Stern and Gerlach passed a beam of silver atoms through such an apparatus. These atoms have a magnetic moment (thanks to the electron spin) and so are deflected by the apparatus. Based upon a classical model of the atom's magnetic moment as coming from motion of the (electronic) charge distribution, and given that the atoms in the beam have random orientations of their magnetic moments, one expects a continuous deflection of the beam, reflecting a continuous spread in the projections of the magnetic moment along the inhomogeneity axis. Instead, what was observed was that the beam splits into two parts!

The explanation for this phenomenon is that the magnetic moment vector $\vec{\mu}$ of the atom is due to the presence of an electron which carries intrinsic angular momentum $\vec{S}$ – its "spin" — and hence a magnetic moment. (Here $\vec{\mu}$ is proportional to $\vec{S}$: $\mu = \frac{q}{mc}S$) The electron is in an atomic state which does not have any orbital angular momentum, so its

---

* This also gives a good illustration of how one can be completely wrong in one's theoretical expectations only to be summarily corrected by experimental data. Read the history of this experiment, *e.g.,* on Wikipedia.

motion about the atom does not provide a contribution to the magnetic moment.* In this (initially bizarre) explanation, the electron can be in one of two "spin states" relative to any given direction so as to explain why the beam split into two parts. More precisely, the projection of the spin of an electron along any axis can only take two values: $\pm\hbar/2$. Such particles are said to have "spin 1/2", an intrinsic (state-independent) property of the electron. If each atom is randomly selected, one expects that these two alternatives occur with 50-50 probability, and this is what is observed. Half the beam gets deflected in each direction. To explain this discreteness – indeed, two-valued nature – of an angular momentum is one challenge faced by any putative theory. Soon I will show you how to create a mathematical model of this two-state behavior using the rules of quantum mechanics. But there is much more to the experiment which should be contemplated first.

Using Stern-Gerlach apparatuses, we can measure any component of the magnetic moment – equivalently, component of the spin vector – of a spin 1/2 particle by aligning the SG magnetic field direction along the axis of interest, then passing the beam of particles through and seeing which way the particles are deflected, corresponding to spin "up" or "down" along that direction. Let us therefore try to model the behavior of the spin vector $\vec{S}$ in such an experiment, ignoring all the other "degrees of freedom" that the atoms might have. Thus the atom is modeled as a "spin 1/2 particle". Let us call an SG apparatus that measures the spin along an axis characterized by a unit vector $\hat{n}$ "SG_n". Thus the apparatus SG_n measures $\vec{S}\cdot\hat{n}$. The empirical fact is that if you measure $\vec{S}\cdot\hat{n}$ you always get $\pm\hbar/2$. Let us pass a beam of spin 1/2 particles through SG_n and keep, say, only the particles that deflect according to $\vec{S}\cdot\hat{n}$ having the value $+\frac{\hbar}{2}$. If we pass this *filtered* beam through another such SG_n/filter device we see that 100% of the beam passes through. We say that we have "determined the spin along $\hat{n}$ with certainty" for all the particles in the filtered beam. We model this situation by saying that all the particles in the (filtered) beam are in the state $|\vec{S}\cdot\hat{n},+\rangle$. We say that we have "prepared" many particles all in the same state by passing a beam of particles through an SG apparatus and only keeping those deflected up or down.

With the notation and terminology understood, let us proceed to a more concrete situation. Suppose we pass a beam through the apparatus SG_z and only keep one spin projection: "spin up". We now have a beam consisting of electrons all prepared in the state $|S_z,+\rangle$. Let us try to pin down the value of S_x that these electrons possess. Pass the beam (all particles in the state $|S_z,+\rangle$) through another Stern-Gerlach apparatus SG_x. Particles are now deflected according to the projection of their magnetic moments (or spin vectors) along the x direction. What you find in this experiment is that the beam splits in half. This is perfectly reasonable; we have already decided that any component of the

---

* Of course, there are 47 electrons in a silver atom. However 46 of them are in a state with no net angular momentum and hence no net contribution to the magnetic moment.

spin has just two projections along any given axis. Since there is nothing special about the x or z directions; we should get similar behavior for both. In the SG_z filtered beam we did not "prepare" S_x in any special way, so it is not surprising that we get the beam to split in half.

Let us continue our investigation as follows. We have passed our beam through SG_z and kept the "spin up" particles. We then pass these spin up particles through SG_x; let us focus on the beam that gave $\hbar/2$ for the S_x measurement. Therefore, roughly half of the beam that entered the SG_x apparatus is kept, and we now have 1/4 of the original particles left in our prepared beam. After this filtering process we can, if we like, verify the value of S_x using SG_x again; and find all the particles do indeed have $S_x = \hbar/2$. Thus we can represent the state of the particles by by $|S_x, +\rangle$.[*]

Now we have a beam of electrons that have been measured to have the following properties: first , S_z is $+\hbar/2$; second S_x is $+\hbar/2$.[†] Given (1) and (2) above, it is reasonable to believe that the electrons we have kept now have definite values for S_z and S_x since we have filtered out the only other possibilities. This belief does not agree with experiment! Suppose you go back to check on the value of S_z. Take the beam that came out of SG_z with value $\hbar/2$ and then SG_x with the value $+\hbar/2$ and pass it through SG_z again. You may expect that all of the beam is found to have a value $+\hbar/2$ for S_z, but instead you will find that beam splits in two! This is despite the fact that we supposedly filtered out the spin down components along z.

So, if you measure S_z and get, say, $\hbar/2$, and then you measure it again, you will get $\hbar/2$ with probability one (assuming no other interactions have taken place). If you measure S_z and get, say, $\hbar/2$, then measure S_x and then measure S_z, the final measurement will be $\pm\hbar/2$ with a 50-50 probability. This should get your attention! The values that you can get for the observable S_z in two measurements depends upon whether or not you have determined the value of S_x in between the S_z measurements.

Given this state of affairs, it is hard to make sense of the classical picture in which one imagines the electron to have given, definite values of all its observables, *e.g.,* S_x and S_z. One sometimes says that the measurement of S_x has somehow "disturbed" the value of S_z. This point of view is not incorrect, but is not a perfect description of what is going on. For example, as we shall see, the quantum mechanical prediction is unambiguously independent of the way in which we make the measurements. Nowhere do we really need to know how the SG devices worked. Moreover, the "disturbance" in S_z due to the S_x measurement is not a function of how carefully we make the S_x measurement, that is,

---

[*] Of course, one naturally prefers to write the state as something like $|S_z, +; S_x, +\rangle$, but we shall see that this is not appropriate.

[†] We could now go and measure S_y in this doubly filtered beam; you will find that half the beam has spin up along y, half has spin down (exercise). But let us not bother with this.

one cannot blame the strange behavior as coming from some “experimental error”, the measurements can, ideally, be perfect and we still get the same result. The fact of the matter is that one shouldn’t think of observables (such as S_z and S_x) has having given, fixed, values that “exist” in the object of interest. This may be philosophically a bit sticky (and psychologically a bit disturbing), but it seems to be quite alright as a description of how nature actually works.

If all this seems perfectly reasonable to you, then you probably don’t understand it too well. Our macroscopic experience with matter just doesn’t give any hint that this is the way nature works microscopically.

Electrons (and other elementary particles) are not like tiny baseballs following classical trajectories with tiny spin angular momentum arrows attached to them, and there is no reason (experimentally) to believe that they are. It is a purely classical prejudice that a particle has definite values for all observables that we can measure. Try to think this way: what is a particle? It has mass, (total) spin, charge, *etc.* and other intrinsic, “real” properties that do not change with the state of the particle. Based upon experiment, one may want to assign other observable properties such as position, energy, orbital angular momentum, spin component along an axis to the particle. But according to experiment, these properties change with the state of the particle and cannot be viewed as “existing” in the particle independently of the measuring process (which changes the state). As it turns out, according to the quantum mechanical explanation of this sort of phenomenon, all you are guaranteed to be able to “assign” to a particle is probability distributions for its various observables. Our next task is to build up the quantum mechanical model of the spin 1/2 system using the rules of quantum mechanics.

Spin

Spin is an intrinsic property of an elementary particle, like its mass and charge. Spin is a form of angular momentum – conservation of angular momentum requires the spin to be taken into account. Charged particles with spin, *e.g.,* the electron, have an intrinsic magnetic moment which is proportional to the spin vector, but neutral particles can have spin, *e.g.,* the neutrino. Unlike mass and charge, which only take a single value for a given particle, the spin can be in various states. More precisely, the magnitude of the spin vector is fixed for a given type of particle – unlike the orbital angular momentum whose magnitude can vary – while any given component of that spin vector can take various discrete values. Here I will focus on “spin-1/2”, which pertains to the most significant matter in the universe, *e.g.,* particles such as electrons, protons, neutrons, positrons, neutrinos, quarks, *etc.*

The observable of interest is the spin vector, which represents angular momentum carried intrinsically by the particle. Like orbital angular momentum it can be represented by 3 Hermitian operators, $\mathbf{S} = (S_x, S_y, S_z)$ satisfying the angular momentum commutation relations:

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y.$$

As mentioned when we studied the orbital angular momentum, these relationships can be used to deduce the possible eigenvalues and eigenvectors for any given component. It is not too hard to do, but we don't have time for it. Let me just state the main results. First of all, while each component will have a basis of eigenvectors, the bases will be different for different components. The magnitude squared of the spin,

$$S^2 = S_x^2 + S_y^2 + S_z^2$$

is compatible with any component:

$$[S^2, S_x] = [S^2, S_y] = [S^2, S_z] = 0.$$

Consequently, we can find a basis of eigenvectors of S^2 and any component – conventionally we work with eigenvectors of S_z. It can be shown from the commutator algebra that the eigenvalues of S^2 are given by

$$s(s+1)\hbar^2, \quad s = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ 2,\ \ldots$$

where it is understood that for a given type of particle the s value is fixed. Conventionally, it is the s-value which is used to characterize the spin. The "spin 1/2" particles are those with $s = 1/2$. It is pedagogically unfortunate that the particles with "spin s" actually have their total angular momentum taking the value $\sqrt{s(s+1)}\hbar$. The eigenvalues of any one component, say S_z, are given by

$$m_s\hbar, \quad m_s = -s, -s+1, \ldots s-1, s.$$

This is all very much like orbital angular momentum, of course. But there are two key differences to keep in mind. First, the quantum number s controlling the total spin value is not subject to change with the change of state of the particle; it is fixed by the type of particle (as is the mass and electric charge). The electron may have different l values in an atom, but its s value is *always* 1/2. This means that every state has the value of S^2 determined with certainty, and the operator representing S^2 must take the form

$$\hat{S}^2 = s(s+1)\hbar^2\hat{1},$$

where s is fixed and $\hat{1}$ is the identity operator. Second, the orbital angular momentum operators require that only integer values of l (and m_l) can occur. This stems from the fact that $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ with the momentum acting via the gradient on wave functions. The integer values of l are needed so the wave functions are sufficiently well-defined for $\mathbf{L}$ to act on them. The spin operators are not constructed as $\mathbf{r} \times \mathbf{p}$ and do not have that restriction; half integral values of s (and m_s) may occur for elementary particles. Half-integral spin particles of a given type (*e.g.,* electrons) obey Fermi-Dirac statistics and are called *fermions.* Integral spin particles of a given type (*e.g.,* photons) obey Bose-Einstein statistics and are called *bosons.*

Even though the spin (s) is fixed for a given type of particle, if $s \neq 0$ the value of m_s can vary so that the particle can have more than one *spin state.* Of course the particle can also have a variety of states of motion, but we will ignore that for now and just consider the spin states as if that is all the particle can ever "do". This is for pedagogical reasons, but there *are* physical situations where the spin state is pretty much all one is interested in. For example it is the spin state of the electron and proton in interstellar hydrogen which explains the famous 21 centimeter hydrogen emission line, and it is the spin state of electrons in a lattice which is used to explain ferromagnetic phase transitions in crystalline solids.

Spin 1/2

Let's now specialize to the consideration of "spin 1/2", *i.e.,* particles with $s = 1/2$. This means we are studying the spin properties of particles like the electron, which is a fermion. The magnitude of the spin for an electron is therefore $\sqrt{3}\hbar/2$. The m_s values are $\pm\frac{1}{2}$, so that the component of the spin vector along any direction takes only two values, $\pm\hbar/2$. Therefore the basis of eigenvectors of any of the spin components will have only 2 elements and the space of spin states is two-dimensional.*

Because the space of states of a spin 1/2 particle is 2-dimensional, we can fruitfully use a matrix representation of vectors and operators. The vectors representing states are then column vectors with two components:

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}.$$

The scalar product between two vectors is the usual one from matrix algebra, but keep in mind there is a complex conjugation:

$$\langle\phi|\psi\rangle = (\phi_1^* \ \phi_2^*) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \phi_1^*\psi_1 + \phi_2^*\psi_2.$$

---

* Thus we have an example of a "2 state system". This does not mean there are only two states – there are infinitely many vectors in a 2-d vector space. What it does mean is that any observable can have (at most) 2 eigenvalues and 2 eigenvectors, so there are 2 possible outcomes of a measurement and two corresponding states.

The normalization condition for $|\psi\rangle$ is then

$$1 = \langle\psi|\psi\rangle = |\psi_1|^2 + |\psi_2|^2.$$

The spin operators can be constructed as $\hbar/2$ times the *Pauli spin matrices*:

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is straightforward, if a bit tedious (unless you use a computer) to check that these matrices do satisfy the angular momentum commutation relations. It is also straightfoward to check that

$$S^2 = S_x^2 + S_y^2 + S_z^2 = \frac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

as it should. Notice that the basis

$$|S_z, +\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |S_z, -\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

is a basis of eigenvectors of S_z:

$$S_z|S_z, \pm\rangle = \pm\frac{\hbar}{2}|S_z, \pm\rangle.$$

It is straightforward to show that the other spin components have the following column vectors representing their eigenvectors:

$$|S_x, \pm\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}, \quad |S_y, \pm\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm i \end{pmatrix}.$$

Incompatibility of the spin operators

We have already noted that the components of the spin vector, like any angular momentum, are incompatible observables. A simple physical manifestation of this is the phenomenology of the Stern-Gerlach experiment, which we can now analyze precisely using the mathematics presented above.

Suppose we have a device (a "Stern-Gerlach apparatus") which can measure the component of spin along any chosen axis. Let us take a beam of electrons and only keep those whose x component of spin has the value $+\frac{\hbar}{2}$. Let us now measure the z component of spin with a subsequent device. After the initial measurement, the electrons we retained are all in the state $|S_x, +\rangle$. The probability for getting this value of S_x is now 100%. The probability distribution $P(S_z, \pm)$ for the outcomes $\pm\frac{\hbar}{2}$ of a measurement of S_z is given by

$$P(S_z, \pm) = |\langle S_z, \pm|S_x, +\rangle^2 = \frac{1}{2}.$$

Evidently, there is a 50-50 chance of getting the two possible outcomes. Let us just keep the electrons which have $+\frac{\hbar}{2}$ for S_z. So, at this point we filtered the beam twice, once to select $S_x = \frac{\hbar}{2}$ and then to further select $S_z = \frac{\hbar}{2}$. Now let us measure S_x again. After the second measurement we have electrons in the state $|S_z, +\rangle$. The probability distribution $P(S_x, \pm)$ is therefore given by

$$P(S_x, \pm) = |\langle S_x, \pm | S_z, +\rangle^2 = \frac{1}{2}.$$

Evidently it is not possible for us to ascribe statistically certain values of S_x and S_z (or any two components) to the electron!

Larmor precession

You may recall from electrodynamics that a rotating charged body with angular momentum $\vec{L}$, magnetic moment $\vec{\mu} = \kappa\vec{L}$ immersed in a (for simplicity, uniform) magnetic field $\vec{B}$ experiences a torque

$$\vec{\tau} = \vec{\mu} \times \vec{B}.$$

Assuming $\tau \neq 0$, this torque causes the magnetic moment to precess about the magnetic field direction at a frequency κB, much as the angular momentum vector of a gyroscope will precess when the gyroscope is in a gravitational field.

The magnetic moment of the electron has a magnitude $\frac{\sqrt{3}e\hbar}{2mc}$, where e is the magnitude of the charge of the electron and m is its mass; I am using Gaussian electromagnetic units. Since we identify the magnetic moment as arising from the spin, the corresponding magnetic moment vector is

$$\vec{\mu} = -\frac{e}{mc}\vec{S}.$$

Because of this magnetic moment, the electron will interact with an applied magnetic field even if it is at rest. Let us consider the time evolution of the spin observable due to a uniform magnetic field $\vec{B}$. Of course, the electron may be moving and that will introduce additional observable behavior, but we do not model that here, just the dynamics of the spin and magnetic moment. Suppose the electron is at rest.

According to the rules of quantum mechanics, time evolution is always via the Schrödinger equation:

$$\hat{H}|\psi(t)\rangle = i\hbar\frac{d}{dt}|\psi(t)\rangle.$$

So to build that equation we need a Hamiltonian $\hat{H}$. Recall that the potential energy of a point magnetic dipole with an applied magnetic field is of the form $-\vec{\mu}\cdot\vec{B}$. We will use this interaction energy to define the Hamiltonian operator:

$$\hat{H} = \frac{e}{mc}\vec{B}\cdot\hat{\vec{S}}.$$

Since $\vec{B}$ is fixed in space, and since we have built our spinor technology to favor the z axis, let us choose our z axis along $\vec{B}$. In this reference frame the Hamiltonian is

$$\hat{H} = \frac{eB}{mc}\hat{S}_z.$$

As usual, since the Hamiltonian has no explicit time dependence, if we can solve the eigenvalue problem for the Hamiltonian we can solve the Schrödinger equation. Fortunately, the Hamiltonian is proportional to S_z so the S_z eigenvectors will be the energy eigenvectors:

$$\hat{H}|S_z, \pm\rangle = \pm\left(\frac{eB\hbar}{2mc}\right)|S_z, \pm\rangle.$$

Let us denote by ω the classical precession frequency (this angular frequency is called the *Larmor frequency* in this context):

$$\omega = \frac{eB}{mc}.$$

We then have energy eigenvalues

$$E_\pm = \pm\frac{1}{2}\hbar\omega.$$

Suppose the initial state of the system is given by

$$|\psi(0)\rangle = a|S_z, +\rangle + b|S_z, -\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \quad |a|^2 + |b|^2 = 1.$$

Notice that every state vector is of this form, so we are considering all possible initial conditions. The meaning of this state is that the probability for finding $\hbar/2$ for measuring S_z is $|a|^2$ and the probability for finding $-\hbar/2$ is $|b|^2$. The probability distribution $P(S_x, \pm\hbar/2)$ for S_x is

$$P(S_x, \pm\hbar/2) = |\langle S_x, \pm|\psi(0)\rangle|^2 = \frac{1}{2}[1 \pm (ab^* + a^*b)] = \frac{1}{2} \pm \Re(ab^*).$$

As a good exercise you should determine the probability distribution for S_y. Using our usual strategy for solving the Schrödinger equation via energy eigenvectors we have

$$|\psi(t)\rangle = ae^{-i\frac{\omega}{2}t}|S_z, +\rangle + be^{i\frac{\omega}{2}t}|S_z, -\rangle,$$

or, if you prefer,

$$\begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} = \begin{pmatrix} ae^{-i\frac{\omega}{2}t} \\ be^{i\frac{\omega}{2}t} \end{pmatrix}.$$

Probably you should directly verify that this formula does in fact solve the Schrödinger equation. Notice that the state vector repeats itself in time at a frequency which is 1/2 the Larmor frequency, that is, it takes twice as long as the classical magnetic moment takes to repeat itself. In one Larmor period the state vector changes sign. This is just a phase so the observable properties defined by the state vector actually are periodic at the Larmor frequency. As always, the energy probability distribution (which is also the S_z probability distribution in this example) is time independent – this is conservation of energy. Unless $ab = 0$, the system is *not* in a stationary state and so *something* observable must be changing in time. Let us consider the probability distribution for S_x. Letting $a \to ae^{-i\frac{\omega}{2}t}$ and $b \to be^{i\frac{\omega}{2}t}$ we can use our previous result on the probability distribution for S_x to get the probability distribution at time t, denoted by $P_t(S_x, \pm\hbar/2)$. We then get

$$P(S_x, \pm\hbar/2) = \frac{1}{2} \pm \Re(ab^* e^{-i\omega t}).$$

Thus the probabilities for S_x are periodic at the Larmor frequency. It is a straightforward calculation to compute the expectation values of the spin observables as a function of time. For simplicity, let us suppose that $a = b = 1$ so that the initial state is one where S_x is $\hbar/2$ with certainty. We then get (good exercise!)

$$\langle S_x\rangle(t) = \frac{\hbar}{2}\cos(\omega t), \quad \langle S_y\rangle(t) = \frac{\hbar}{2}\sin(\omega t), \quad \langle S_z\rangle(t) = 0.$$

A classical analog of this computation would be that the spin vector starts off pointing along the x axis and then precesses around the z axis at the Larmor frequency.

This very simple physical model is the basis for understanding the phenomenon of *spin resonance.** If a weak (compared to B) oscillating magnetic field is introduced a resonance phenomena will occur when the frequency of oscillation approaches ω. Another way to think about this is that the system preferentially emits and absorbs photons with angular frequency ω by the usual quantum field theory arguments.

Finally it is amusing to note that while all measurements of the spin will evolve in time at the Larmor frequency, the state vector itself evolves at one half the Larmor frequency. You can trace this 1/2 back to the fact that the spin is 1/2. Because of this 1/2, as mentioned earlier, after one Larmor period the state vector changes sign. An overall sign change is just a constant phase factor and so does not affect the outcome of measurements of the spin. It is tempting to think that this sign change is just a mathematical feature with no physical content, but this is not quite true. One can "see" this sign change if we consider two electrons, one in the magnetic field and one not in the magnetic field. Suppose they both start off in the same spin state. Now one electron's vector changes sign after the Larmor period relative to the other. This relative phase *can* be measured, *e.g.,* via an interference experiment. This experiment has been done (with neutrons) and the change in sign has been confirmed! It is this phenomenon which sometimes leads people

---

* Also known as *magnetic resonance.*

to say that if you rotate an electron by 360° that it is does not return to its original state, that it is instead a rotation by 720° which returns an electron to its original state.

Addition of angular momenta: two spin 1/2 particles

It is not too hard to imagine, at least conceptually, how to extend our treatment of spin 1/2 to treat two spin 1/2 particles. We will still be ignoring translational degrees of freedom and only study the spin states of the particles. The idea is just to consider two copies of what we did for a single spin 1/2 particle. To be concrete, let us suppose that the particles are an electron, labeled by "1" and a proton, labeled by "2". The state of the composite system is now built as a superposition of the possible states of each subsystem.* Let us use the basis of S_z eigenvectors for each particle. If we measure the S_z component for each particle you can easily see that there are four possible outcomes, corresponding to particle one being spin up or down and particle 2 being spin up or down. If we denote the two sets of S_z eigenvectors by $|\pm\rangle_1$ and $|\pm\rangle_2$, we have the 4 basis states:

$$|++\rangle = |+\rangle_1 \otimes |+\rangle_2, \quad |+-\rangle = |+\rangle_1 \otimes |-\rangle_2, \quad |-+\rangle = |-\rangle_1 \otimes |+\rangle_2 \quad |--\rangle = |-\rangle_1 \otimes |-\rangle_2$$

Here I have introduced the fancy *tensor product* symbol $\otimes$; for now just think of it as an elaborate way to juxtapose the two individual spin states. Since we always put particle one on the left side of the $\otimes$ and particle two on the right side, we can just drop the 1 and 2 if there is no ambiguity. The key step now is to use these 4 states as a basis for a *four-dimensional* Hilbert space – the space of states of the two spin system. This means the vector space is the set of linear combinations:

$$|\psi\rangle = a|++\rangle + b|-+\rangle + c|+-\rangle + d|--\rangle.$$

If you like, you can view the vector space as the usual complex vector space of column vectors with 4 rows:

$$|\psi\rangle = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

The scalar product between

$$|\psi\rangle = a|++\rangle + b|-+\rangle + c|+-\rangle + d|--\rangle$$

and

$$|\phi\rangle = e|++\rangle + f|-+\rangle + g|+-\rangle + h|--\rangle$$

---

* This would be a way to analyze the contribution of electronic and nuclear spin to the angular momentum of a hydrogen atom to understand, *e.g.*, the 21 cm emission line seen astrophysically.

is given by

$$\langle\phi|\psi\rangle = (e^* f^* g^* h^*)\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = e^*a + f^*b + g^*c + h^*d.$$

The normalization condition on $|\psi\rangle$ is therefore

$$1 = \langle\psi|\psi\rangle = |a|^2 + |b|^2 + |c|^2 + |d|^2.$$

You can easily check that the 4 basis vectors $|\pm,\pm\rangle$ form an orthonormal basis.

The meaning of the 4 complex numbers a, b, c, d in

$$|\psi\rangle = a|++\rangle + b|-+\rangle + c|+-\rangle + d|--\rangle.$$

is that $|a|^2$ is the probability that a measurement of S_z for each particle gives $+\hbar/2$ for both; $|b|^2$ is the probability that a measurement of S_z for each particle gives $-\hbar/2$ for the first particle and $\hbar/2$ for the second particle, and so forth. As usual the corresponding vectors represent states in which the outcome of those measurements is statistically certain. For example, $|-+\rangle$ is a state in which the electron has spin down along z and the proton has spin up along z, both with probability 1.

We can extend all the spin eigenvectors for a single spin to the case of two spins. For example, we know that

$$|S_x,\pm\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle).$$

The state vector $|\chi\rangle$ where particle one has $S_z = \frac{\hbar}{2}$ and particle two has $S_x = -\frac{\hbar}{2}$ is then given by

$$|\chi\rangle = |+\rangle\otimes|S_x,-\rangle = |+\rangle\otimes\left[\frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)\right] \equiv \frac{1}{\sqrt{2}}\Big(|+\rangle\otimes|+\rangle - |+\rangle\otimes|-\rangle\Big) = \frac{1}{\sqrt{2}}\Big(|++\rangle - |+-\rangle\Big).$$

In the last equality I have expressed $|\chi\rangle$ in our original orthonormal basis. If the two-spin system is in the state represented by $|\chi\rangle$ then the possible outcomes of a measurement of S_z for both particles is that particle one has $S_z = \hbar/2$ and particle two has $S_z = \pm\hbar/2$ with equal probability. One way to compute this is scalar products. For example, the probability P that particle one has spin up along z and particle two has spin up along z is given by

$$P = |\langle +,+|\chi\rangle|^2 = |\frac{1}{\sqrt{2}}\langle +,+|+,+\rangle|^2 = \frac{1}{2}.$$

According to the postulates of quantum mechanics each observable is represented by a Hermitian operator. We have been talking about measuring the spin of particle one and particle two; what is the operator representing those observables? The idea is simply to let the operator representing the spin of particle 1, $\vec{S}_1$, be the usual spin operator $\vec{S}$ acting on the particle one part of the total state, and to let the operator representing the spin of particle 2, $\vec{S}_2$, be the usual spin operator $\vec{S}$ acting on the particle two part of the total state. We have, for any basis vector

$$|\pm,\pm\rangle = |\pm\rangle \otimes |\pm\rangle$$

$$\hat{\vec{S}}_1|\pm,\pm\rangle \equiv (\hat{\vec{S}}|\pm\rangle) \otimes |\pm\rangle, \quad \hat{\vec{S}}_2|\pm,\pm\rangle \equiv |\pm\rangle \otimes (\hat{\vec{S}}|\pm\rangle).$$

We extend the definitions of these operators to *any* vector by decreeing they are linear operators, so they distribute across sums and commute with scalar multiplication. Since every vector is a linear combination of basis vectors, this defines these operators on any vector. A useful shorthand notation for writing these two operators is

$$\hat{\vec{S}}_1 = \hat{\vec{S}} \otimes 1, \quad \hat{\vec{S}}_2 = 1 \otimes \hat{\vec{S}}.$$

If you like to view the state vectors as 4-row column vectors then you should be able to check that

$$\hat{S}_{1z} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

and

$$\hat{S}_{2y} = i\frac{\hbar}{2}\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

So far we have only considered observables which pertain to each particle individually. Given a system consisting of two spin 1/2 particles, there is another class of observables we can consider: the *total* spin angular momentum $\vec{\mathcal{J}} = \vec{S}_1 + \vec{S}_2$ exhibited by the system. This observable is represented by the following Hermitian operator:

$$\hat{\vec{\mathcal{J}}} = \hat{\vec{S}}_1 + \hat{\vec{S}}_2 = \hat{\vec{S}} \otimes 1 + 1 \otimes \hat{\vec{S}}.$$

You can easily check that the eigenvectors of the i^{th} component, $\mathcal{J}_i$ are given by vectors of the form

$$|S_i,\pm\rangle \otimes |S_i,\pm\rangle.$$

For example,

$$\begin{aligned} \hat{\mathcal{J}}_x|S_x,+\rangle \otimes |S_x,+\rangle &= (\hat{S}_x|S_x,+\rangle) \otimes |\hat{S}_x,+\rangle + |S_x,+\rangle \otimes (\hat{S}_x|S_x,+\rangle) \\ &= (\frac{\hbar}{2}|S_x,+\rangle) \otimes |S_x,+\rangle + |S_x,+\rangle \otimes (\frac{\hbar}{2}|S_x,+\rangle) \\ &= \hbar|S_x,+\rangle \otimes |S_x,+\rangle. \end{aligned}$$

Similarly,

$$\begin{aligned}
\hat{\mathcal{J}}_x|S_x,-\rangle\otimes|S_x,+\rangle &= 0,\\
\hat{\mathcal{J}}_x|S_x,+\rangle\otimes|S_x,-\rangle &= 0,\\
\hat{\mathcal{J}}_x|S_x,-\rangle\otimes|S_x,-\rangle &= -\hbar|S_x,-\rangle\otimes|S_x,-\rangle.
\end{aligned}$$

As a nice exercise you might try to verify these results via the matrix representation mentioned earlier.

You may have noticed that the eigenvalues of $\mathcal{J}_x$ are consistent with a quantum number for $\mathcal{J}^2$ given by $s = 0, 1$. This is not an accident. To see this, first one calculates

$$[\mathcal{J}_x, \mathcal{J}_y] = i\hbar\mathcal{J}_z, \quad \text{and cyclic permutations.}$$

This calculation can be understood as follows:

$$\begin{aligned}
[\mathcal{J}_x, \mathcal{J}_y] &= \left(\hat{S}_x\otimes 1 + 1\otimes\hat{S}_x\right)\left(\hat{S}_y\otimes 1 + 1\otimes\hat{S}_y\right) - \left(\hat{S}_y\otimes 1 + 1\otimes\hat{S}_y\right)\left(\hat{S}_x\otimes 1 + 1\otimes\hat{S}_x\right)\\
&= [\hat{S}_x, \hat{S}_y]\otimes\hat{1} + \hat{1}\otimes[\hat{S}_x, \hat{S}_y]\\
&= i\hbar\hat{S}_z\otimes\hat{1} + i\hbar\hat{1}\otimes S_z\\
&= i\hbar\hat{\mathcal{J}}_z.
\end{aligned}$$

Thus the total angular momentum satisfies the commutation relations appropriate to any angular momentum operator. As we have discussed earlier, this means the eigenvalues of

$$\mathcal{J}^2 = S_1^2 + S_2^2 + 2(S_{1x}\otimes S_{2x} + S_{1y}\otimes S_{2y} + S_{1z}\otimes S_{2z})$$

are

$$s(s+1)\hbar^2, \quad s = 0, 1,$$

and the eigenvalues of $\mathcal{J}_z$ are

$$m\hbar = \begin{cases} \pm\hbar, 0, & s = 1\\ 0, & s = 0 \end{cases}.$$

The restriction on the s values follows from the initial observation that the eigenvalues of a component of $\vec{\mathcal{J}}$ are $0, \pm\hbar$ along with the observation that there are 4 such eigenvectors. Thus there are 3 states of the total system with "spin-1" and a single state of the total system with "spin-0". One sometimes refers to the spin-1 states as the "triplet states" and the spin-0 state as the "singlet state". The singlet and triplet states are

$$\begin{aligned}
s = 0, m = 0 \quad &\leftrightarrow \quad \frac{1}{\sqrt{2}}(|+,-\rangle - |-,+\rangle) \equiv |0,0\rangle\\
s = 1, m = 1 \quad &\leftrightarrow \quad |+,+\rangle \equiv |1,1\rangle\\
s = 1, m = 0 \quad &\leftrightarrow \quad \frac{1}{\sqrt{2}}(|+,-\rangle + |-,+\rangle) \equiv |1,0\rangle\\
s = 1, m = -1 \quad &\leftrightarrow \quad |-,-\rangle \equiv |1,-1\rangle.
\end{aligned}$$

Notice that these eigenvectors form a basis for the 4-dimensional vector space. The coefficients appearing in the expansion of one basis in terms of the other (as above) are known as *Clebsch-Gordan coefficients.*

To find these eigenvectors one just creates the matrix representatives of $\hat{\mathcal{J}}^2$ and $\hat{\mathcal{J}}_z$ and solves their eigenvalue problem in the usual way. I will spare you all that. But I will, at least, show you how these eigenvectors work. The key formulas we need (besides the definitions of $\hat{\mathcal{J}}^2$ and $\hat{\mathcal{J}}_z$) are

$$(\hat{S}_1)^2 = \frac{3}{4}\hbar^2\hat{1} = (\hat{S}_2)^2,$$

$$\hat{S}_x|\pm\rangle = \frac{\hbar}{2}|\mp\rangle, \quad \hat{S}_y|\pm\rangle = \pm i\frac{\hbar}{2}|\mp\rangle, \quad \hat{S}_z|\pm\rangle = \pm\frac{\hbar}{2}|\pm\rangle.$$

Then we have

$$\begin{aligned}\hat{\mathcal{J}}^2|+,-\rangle &= \frac{3}{4}\hbar^2|+,-\rangle + \frac{3}{4}\hbar^2|+,-\rangle + 2\left(\frac{\hbar^2}{4}|-.+\rangle + \frac{\hbar^2}{4}|-.+\rangle - \frac{\hbar^2}{4}|+.-\rangle\right)\\ &= \hbar^2|+,-\rangle + \hbar^2|-+\rangle,\end{aligned}$$

and

$$\begin{aligned}\hat{\mathcal{J}}^2|-,+\rangle &= \frac{3}{4}\hbar^2|-,+\rangle + \frac{3}{4}\hbar^2|-,+\rangle + 2\left(\frac{\hbar^2}{4}|+.-\rangle + \frac{\hbar^2}{4}|+.-\rangle - \frac{\hbar^2}{4}|-.+\rangle\right)\\ &= \hbar^2|-,+\rangle + \hbar^2|+,-\rangle,\end{aligned}$$

$$\begin{aligned}\hat{\mathcal{J}}^2|\pm,\pm\rangle &= \frac{3}{4}\hbar^2|\pm,\pm\rangle + \frac{3}{4}\hbar^2|\pm,\pm\rangle + 2\left(\frac{\hbar^2}{4}|\mp,\mp\rangle - \frac{\hbar^2}{4}|\mp,\mp\rangle + \frac{\hbar^2}{4}|\pm,\pm\rangle\right)\\ &= 2\hbar^2|\pm,\pm\rangle,\end{aligned}$$

Now you can easily check that

$$\hat{\mathcal{J}}^2|0,0\rangle = 0, \quad \hat{\mathcal{J}}^2|1,\pm1\rangle = 2\hbar^2|1,\pm1\rangle, \quad \hat{\mathcal{J}}^2|1,0\rangle = 2\hbar^2|1,0\rangle,$$

and

$$\hat{\mathcal{J}}_z|0,0\rangle = 0, \quad \hat{\mathcal{J}}_z|1,\pm1\rangle = \pm\hbar|1,\pm1\rangle, \quad \hat{\mathcal{J}}_z|1,0\rangle = 0.$$

We now have two physically important classes of basis vectors for the space of spin states of a system of two spin 1/2 particles. Firstly, there are the bases defined by eigenvectors of the individual particle spins. Physically these are states in which the component of the individual spins (along some chosen axes) are determined with certainty. Secondly, there is the basis defined by the total spin of the composite system. I must emphasize that the basis of total spin eigenvectors (the singlet and triplet states) is not the same as the basis formed by the individual spin eigenvectors (the original basis we stared with). This corresponds to the fact that the observable $\mathcal{J}^2$ is not compatible with S_{1z} and S_{2z}. States where the total spin is determined with certainty may have the individual spins having statistical uncertainty, and *vice versa.* Notice that this only happens for states in which $\mathcal{J}_z = 0$ with certainty.

Spin correlations, the EPR paradox

I now present to you another striking phenomenon in quantum mechanics having to do with remarkable effects of quantum mechanical correlations between particles due to, ultimately, the existence of incompatible observables. Many results stem from the discussion we will have here; for example, entanglement, quantum computation, the Einstein-Podolsky-Rosen paradox, "spooky action at a distance". Our focus will be on the sense in which the experiments performed in one location may "affect" the experimental results found in another location, even though the two experiments cannot be causally related (*e.g.,* they are performed simultaneously). This is a very strange feature of quantum mechanics, to be sure, but we shall see that it does not violate the principle of causality and, in particular, the light speed limit. There are various ways to see all this, but perhaps the simplest way is to use a system of two spin 1/2 particles and just focus on spin measurements.

Spin Correlations

Consider a pair of spin 1/2 particles (*e.g.,* an electron and a proton) in a spin singlet eigenstate of $\mathcal{J}^2$ and $\mathcal{J}_z$:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|z+, z-\rangle - |z-, z+\rangle).$$

As usual, our notation means, *e.g.,* in the state $|z+, z-\rangle$ particle 1 has spin up along z and particle 2 has spin down along z with probability unity.

For example, you can imagine a pair of particles being created via a decay process which starts in a zero angular momentum state. If the decay products – the two particles – move in opposite directions, their orbital angular momentum is zero and so, by conservation of angular momentum, their total spin angular momentum must vanish so that the total angular momentum is still zero after the reaction.

Now suppose the particles are well-separated and each enters a detector capable of measuring the spin state along any desired axis. We shall consider the statistical predictions of quantum mechanics for various measurement scenarios.

First, let us consider the probability that particle 1 has spin up along the z axis. It is easy to see the answer from inspection of the singlet state vector, but let us proceed systematically so we do not have to guess. A basis for the degenerate subspace of states in which particle 1 has spin up along z is given by

$$|z+,z+\rangle, |z+,z-\rangle.$$

If the state of the 2-particle system is the vector

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|z+,z-\rangle - |z-,z+\rangle),$$

then the probability for getting spin up is

$$|\langle z+,z+|\psi\rangle|^2 + |\langle z+,z-|\psi\rangle|^2 = \frac{1}{2}.$$

Next, let us suppose that particle 1 is detected to have spin up along z. What is the probability that particle 2 has spin down along z? There are a few ways to approach answering this question. First, one can say that the measurement of particle 1 yielding spin up along z means that in the superposition $|\psi\rangle$ the state has been filtered to be $|z+,z-\rangle$. Now it is clear that the probability for getting spin down along z is unity. Alternatively, since the state of the system is one of zero total S_z, if one of the particles has spin up along z, the other must have spin down along z with certainty. Yet another computation one could do is to compute the probability of finding particle one with spin up and particle 2 with spin down. This gives

$$|\langle z+,z-|\psi\rangle|^2 = \frac{1}{2}.$$

Since we already know that the probability for finding particle one with spin up is 1/2, this means the probability is unity for finding particle 2 with spin down when particle 1 has spin up. Of course, you can reverse the roles of particle 1 and 2 and get the same results (exercise).

Now imagine experimentally verifying these predictions. Particle 1's detector sits in one location and particle 2's detector sits in another. Particle 1's spin measurement gives a random distribution of spin up and spin down results with 50-50 probability. Likewise, particle 2's detector sees a random distribution with 50-50 probability. Nonetheless, every time one particle has spin up, the other has spin down. Thus the two – apparently random – sets of data are in perfect anti-correlation. A fancy way to characterize this anti-correlation is to compute the $S_{1z}S_{2z}$ *correlation function*. If you have had a decent course in statistics you probably have heard that the statistical correlation of two random variables, A and B, can be computed via their correlation function* $C(A,B)$:

$$C(A,B) = \frac{\langle (A - \langle A\rangle)(B - \langle B\rangle)\rangle}{\sigma_A \sigma_B}.$$

---

* This correlation function is only defined if the standard deviations σ_A, σ_B do not vanish.

If the variables are uncorrelated then $C(A, B) = 0$. It they are correlated in some form then $C(A, B) \neq 0$. If they are maximally correlated or anti-correlated, then $C(A, B) = \pm 1$. In the singlet state $|\psi\rangle$ we have (nice exericse!)

$$\langle S_{1z}\rangle = \langle S_{2z}\rangle = 0, \quad \sigma_{S_{1z}} = \sigma_{S_{2z}} = \frac{\hbar}{2}, \quad \langle S_{1z}S_{2z}\rangle = -\frac{\hbar^2}{4}.$$

The correlation function for the spins along z are then given by

$$C(S_{1z}, S_{2z}) = \frac{\langle\psi|(S_{1z} - \langle S_{1z}\rangle)(S_{2z} - \langle S_{2z}\rangle)|\psi\rangle}{\sigma_{S_{1z}}\sigma_{S_{2z}}} = -1.$$

Ultimately, you can blame this complete anti-correlation on the conservation of angular momentum, so this is in itself not really surprising. One might try to argue that, while quantum mechanics is not adequate to tell you *which* particle has spin up and which has spin down (it can only give a 50-50 probability), it is good enough to make sure the spins are anti-correlated. Of course, there is nothing special about the z axis. One can use any axis to define the singlet state and the perfect anti-correlation will hold. Since the two particles can be very far apart one can imagine setting things up so that there should be no causal relationship between the respective measurements, *i.e.,* one can try to argue that the measurement of one spin cannot affect the measurement of the other. This is a classical notion of *locality* for the spin observables. It is now tempting to believe that the particles are created with all 3 components of spin taking their particular values with certainty. Indeed, one can determine whether particle 2 has spin up or down along any desired axis by measuring particle 1's spin along that axis. The locality idea should guarantee that this measurement does not affect particle 2. This is the idea of a classical *reality* of the values of the spin observables. But we know that different components of spin for a particle are incompatible; it is not possible to have states in which more than one component is defined with certainty. Therefore either or both of the classical notions of locality and reality must be false.

Let us examine this conclusion from a different perspective. Suppose that particle 1's detector is set so that it measures spin along the x-axis. What is the probability for particle 1 to have spin up (now along x)? There are two (non-unique) linearly independent states where particle 1 has spin up along x, for example $|x+, z+\rangle$, $|x+, z-\rangle$. The probability that particle one has S_x measured to be $\frac{\hbar}{2}$ is

$$|\langle x+, z+|\psi\rangle|^2 + |\langle x+, z-|\psi\rangle|^2 = \frac{1}{2},$$

which you may have already figured out by inspection of the singlet state vector. (You can check that any other appropriate linearly independent states (*e.g.,* $|x+, x+\rangle$, $|x+, x-\rangle$) will give the same result.) Suppose particle 1 is found to have spin up – now along x. What is the probability for finding spin up along z for particle 2? We can compute this via

$$|\langle x+, z+|\psi\rangle|^2 = \frac{1}{4},$$

which implies the probability for getting particle 2 with spin up along z is 1/2 given that particle 1 had spin up along x (because there was a probability of 1/2 of particle 1 having spin up along x). Moreover, the results of the two spin measurements are now completely *uncorrelated*. Indeed, the $S_{1x}S_{2z}$ correlation function is (exercise)

$$C(S_{1x}, S_{2z}) = 0.$$

This is the central point. One often tacitly assumes that, given the state of a system, all its observable characteristics are there, waiting to be measured. From this point of view, even though the experimenter for particle one did not choose to measure S_z, that value is still whatever it is and the results of particle 2 should therefore still be correlated with the measurements of particle 1. But this isn't what happens. If the experimenter running the detector for particle 1 measures spin along z, then detector 2 (measuring spin along z) will (anti) correlate perfectly. But if the experimenter running the detector for particle 1 starts measuring spin along x then detector 2 sees random, uncorrelated results. This is a manifestation of the incompatibility of S_x and S_z measurements – if the value of S_x for particle 1 is determined with certainty then the value for S_z is not determined with statistical certainty. But this raises a striking issue. It *appears* as if the choice of measurement at detector 1 somehow "controls" the results found at detector 2 in the sense that particle 2 must somehow know whether to correlate or not to correlate with particle 1. Again, these results occur no matter how far apart are the two detectors, whether one measurement is made before, or after, or simultaneously with the other. It *appears* like one can send signals instantaneously over arbitrarily large distances!

Of course, appearances can be deceiving here. As described above there is a sense in which the results in detector 2 "know" about the results from detector 1. So one might be justified in saying that there is a certain non-locality being exhibited by nature. But this non-locality does not violate causility in the sense of superluminal communication. To understand this, consider that in both scenarios, detector 2 sees a random sequence of spin up and spin down results with 50-50 probability. From these data alone there is no way to know which measurements were performed on particle 1. The only way detector 2 can be "told" which measurement detector 1 has performed is if one compares the two results — such a comparison can only be made within the usual confines of causality.

Still, this "non-local" feature of quantum mechanics has caused much discussion. It was first brought to light by Einstein-Podolsky-Rosen (EPR paradox) who used this feature to argue that quantum mechanics could not be a complete description of what is going on. They hoped a more complete theory could tell you precisely what spin properties each particle carried — with no funny business about the properties – the reality – of one particle being "controlled" at a distance by the determination of properties of the other particle. However, theories involving classical locality and reality have been shown to be incompatible with experiment – this is done through the *Bell inequalities.* So, thus far, quantum mechanics is still the theory of choice and we simply have to get used to this weird quantum reality brought about by incompatible observables.