PHYS 116: Quantum Mechanics
## Connor Neely, Spring 2025

1 Quantum States

1.1 Stern-Gerlach Experiments

We'll begin our discussion of quantum mechanics with a classic experiment that has been used to measure the intrinsic spin angular momentum of an atom. To get here, though, we first need to know how spin interacts with magnetic fields. In classical mechanics, all angular momentum is orbital and has $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. So for a charge-$q$, mass-$m$ point particle moving with speed v in a circular orbit of radius r, the magnetic moment is

$$\mu = \frac{I \cdot \pi r^2}{c} = \frac{1}{c}\frac{qv}{2\pi r} \cdot \pi r^2 = \frac{q}{2mc}L \implies \boldsymbol{\mu} = \frac{q}{2mc}\mathbf{L},$$

expressed in Gaussian units. It turns out that the magnetic moment due to spin angular momentum $\mathbf{S}$ is identical up to a unitless, experimentally-determined constant of proportionality g: $\boldsymbol{\mu} = (gq/2mc)\mathbf{S}$.

There's another, much larger difference between orbital and spin angular momentum, though. Suppose we shoot a beam of silver atoms through the magnetic field generated by a pointed north pole and a flat south pole; a cross section of this arrangement (called a Stern-Gerlach device) is provided at right. We can see that this magnetic field is inhomogeneous.
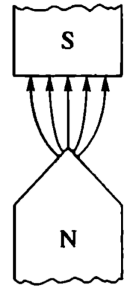
Let the $+z$-direction point upward toward the south pole. When a neutral atom with magnetic moment $\boldsymbol{\mu}$ enters this device it experiences a force $\mathbf{F} = \nabla(\boldsymbol{\mu} \cdot \mathbf{B})$, the z-component of which is approximately $F_z \simeq \mu_z \cdot \partial B_z/\partial z$. Since the magnetic field gradient $\partial B_z/\partial z$ is negative, atoms with $\mu_z < 0$ are deflected upward and ones with $\mu_z > 0$ are deflected downward.

Now, if we were to capture all these deflected atoms on a plate on the other side of the SG**z** device, we might expect to see a continuum of deflections as $\boldsymbol{\mu}$ can be oriented all sorts of ways in space. But this is not so—instead, the beam simply gets split into two. This means that there are only two possibilities for μ_z, and thus that there are only two possibilities for S_z. Since the magnetic moment of a silver atom is largely due to a single electron, an electron must have exactly two spin states: $\hbar/2$ ("spin up") and $-\hbar/2$ ("spin down"). These deflect upward and downward, respectively.

We can leverage the fact that Stern-Gerlach devices split beams of particles according to their intrinsic spins to devise a few experiments, each of which reveals something important about quantum mechanics. We'll use the "ket vector" shorthand $|\pm\mathbf{z}\rangle$ to denote a particle with $S_z = \pm\hbar/2$.

- A beam of particles in $|+\mathbf{z}\rangle$ enters a SG**z** device. We find that all particles exit still in $|+\mathbf{z}\rangle$.

- A $|+\mathbf{z}\rangle$ beam enters a SG**x** device. Half of the particles exit in $|+\mathbf{x}\rangle$, while the other half exit in $|-\mathbf{x}\rangle$.

- A $|+\mathbf{z}\rangle$ beam passes through a SG**x** device, and the ones that end up in $|+\mathbf{x}\rangle$ enter a SG**z** device. Half of the particles exit in $|+\mathbf{z}\rangle$, while the other half exit in $|-\mathbf{z}\rangle$.

For the fourth experiment we employ a modified Stern-Gerlach device, which involves a set of four "simple" devices, two of which are upside-down. The upside-down devices are squished between the other two so that particles passing through are split according to their spins and then brought back together again.

- A $|+\mathbf{z}\rangle$ beam passes through a modSG**x** device and then enter a SG**z** device. All particles exit in $|+\mathbf{z}\rangle$.

These experiments reveal that spin measurements change the state of the system. In the third experiment none of the particles start with $|-\mathbf{z}\rangle$, but by the end half of them are in this state. We don't see this in the fourth experiment, though, since we never actually make any measurements!

This distinction suggests that probabilities don't tell the whole story here. When we go to develop a formalism for quantum mechanics, we'll need to include some kind of "interference" that cancels out the $|-\mathbf{z}\rangle$ state as the particles traverse the modSG**x** device. In analogy with the relationship between an electromagnetic wave's intensity and its corresponding field amplitude, we'll soon introduce a "probability amplitude" that we can square to get a probability. It will be these amplitudes, not the probabilities themselves, that we do calculations with when we make predictions in the quantum world.

1.2 The Quantum State Vector

We'll continue to constrain our attention primarily to spin for a while, as we begin building the mathematical foundations of quantum mechanics. We will include more degrees of freedom in our quantum states later on, things like position and momentum; the principles here will generalize quite easily.

In our formalism the ket vector $|+\mathbf{z}\rangle$ is more than just a shorthand for a spin-up state. It is actually an element of an abstract vector space with basis $\{|+\mathbf{z}\rangle, |-\mathbf{z}\rangle\}$. So if we're given an arbitrary spin state $|\psi\rangle$, which we could generate by pointing a Stern-Gerlach device in an arbitrary direction, we could express the state as a linear combination of the basis states:

$$|\psi\rangle = c_+|+\mathbf{z}\rangle + c_-|-\mathbf{z}\rangle.$$

The fact that any state can be expressed as a superposition of $|+\mathbf{z}\rangle$ and $|-\mathbf{z}\rangle$ means that these two states comprise a complete set. Now, a state $|\psi\rangle$ can also be represented with a bra vector

$$\langle\psi| = c'_+\langle+\mathbf{z}| + c'_-\langle-\mathbf{z}|.$$

The purpose of a bra $\langle\varphi|$ is to meet with a ket $|\psi\rangle$ to form a bracket $\langle\varphi|\psi\rangle$, which is a (complex) inner product that represents the amplitude that a particle in $|\psi\rangle$ will be measured in $|\varphi\rangle$. We've seen from experiment that $\langle-\mathbf{z}|+\mathbf{z}\rangle = 0$ and that $\langle+\mathbf{z}|+\mathbf{z}\rangle$ is nonzero; we'll soon see that it is also crucial that all states satisfy $\langle\psi|\psi\rangle = 1$ for a probabilistic interpretation of quantum mechanics to make sense.

By the distributivity of the inner product, we can deduce that

$$\langle+\mathbf{z}|\psi\rangle = c_+\langle+\mathbf{z}|+\mathbf{z}\rangle + c_-\langle+\mathbf{z}|-\mathbf{z}\rangle = c_+, \qquad \langle\psi|+\mathbf{z}\rangle = c'_+,$$
$$\langle-\mathbf{z}|\psi\rangle = c_+\langle-\mathbf{z}|+\mathbf{z}\rangle + c_-\langle-\mathbf{z}|-\mathbf{z}\rangle = c_-, \qquad \langle\psi|-\mathbf{z}\rangle = c'_-.$$

Also by distributivity, $\langle\psi|\psi\rangle = \langle\psi|+\mathbf{z}\rangle\langle+\mathbf{z}|\psi\rangle + \langle\psi|-\mathbf{z}\rangle\langle-\mathbf{z}|\psi\rangle = 1$, so we must have $\langle\psi|\pm\mathbf{z}\rangle = \langle\pm\mathbf{z}|\psi\rangle^*$ to guarantee real terms. It follows that $c'_\pm = c^*_\pm$. Using this to rewrite the above gives

$$\langle\psi|\psi\rangle = c^*_+c_+ + c^*_-c_- = 1;$$

we therefore interpret $c^*_\pm c_\pm = |\langle\pm\mathbf{z}|\psi\rangle|^2$ as the probability of measuring $|\pm\mathbf{z}\rangle$ from $|\psi\rangle$. Thus in general a state isn't necessarily spin-up or spin-down at all, but rather is a superposition of these two states until a measurement of S_z is taken. The expectation value and uncertainty of such a measurement are given by

$$\langle S_z\rangle = c^*_+c_+(\hbar/2) + c^*_-c_-(-\hbar/2), \qquad (\Delta S_z)^2 = \langle(S_z - \langle S_z\rangle)^2\rangle = \langle S_z^2\rangle - \langle S_z\rangle^2.$$

Finally, we'll give a quick justification for why we must allow our probability amplitudes ot be complex. Our experiments suggest that we have the kets

$$|+\mathbf{x}\rangle = \frac{e^{i\delta_+}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\delta_-}}{\sqrt{2}}|-\mathbf{z}\rangle, \qquad |+\mathbf{y}\rangle = \frac{e^{i\gamma_+}}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\gamma_+}}{\sqrt{2}}|+\mathbf{z}\rangle;$$

if we define $\delta = \delta_- - \delta_+$ and $\gamma = \gamma_- - \gamma_+$, some algebra reveals that

$$|\langle+\mathbf{y}|+\mathbf{x}\rangle|^2 = \frac{1}{2}\big(1 + \cos(\delta - \gamma)\big).$$

To agree with the observation that $|\langle+\mathbf{y}|+\mathbf{z}\rangle|^2$ we must have $\delta - \gamma = \pm\pi/2$, and no matter which we take we will get complex basis coefficients! The standard is $\delta = 0$, $\gamma = \pi/2$, and $\delta_+ = \gamma_+ = 0$ so that

$$|+\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle, \quad |+\mathbf{y}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle.$$

The spin-down kets are simply the normalized orthogonal states

$$|-\mathbf{x}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle - \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle, \quad |-\mathbf{y}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle - \frac{i}{\sqrt{2}}|-\mathbf{z}\rangle,$$

and we could use our knowledge of $\langle\pm\mathbf{x}|\pm\mathbf{z}\rangle$ and $\langle\pm\mathbf{y}|\pm\mathbf{z}\rangle$ to write $|\pm\mathbf{z}\rangle$ in terms of these kets.

1.3 Rotation and Projection Operators

We'll very often find ourselves in positions where we need to transform one quantum state into another. One way we might do this is by applying a rotation operator $\hat{R}(\phi\mathbf{k})$ to a ket, which rotates its corresponding three-vector through an angle ϕ about the z-axis. For example, $\hat{R}(\frac{\pi}{2}\mathbf{k})|+\mathbf{x}\rangle = |+\mathbf{y}\rangle$. In order to act on a bra we must define a new operator $\hat{R}^\dagger$, called its adjoint; this allows us to satisfy, say,

$$1 = \langle+\mathbf{x}|+\mathbf{x}\rangle = \langle+\mathbf{z}|\hat{R}^\dagger\left(\frac{\pi}{2}\mathbf{j}\right)\hat{R}\left(\frac{\pi}{2}\mathbf{j}\right)|+\mathbf{z}\rangle = \langle+\mathbf{z}|+\mathbf{z}\rangle,$$

so long as $\hat{R}^\dagger\hat{R} = 1$. Such an operator is called a unitary operator, and so $\hat{R}$ must be unitary in order to conserve probability. Now, in giving a more explicit definition for $\hat{R}$, we'll find it useful to define an infinitesimal

$$\hat{R}(d\phi\,\mathbf{k}) = 1 - \frac{i}{\hbar}\hat{J}_z d\phi,$$

where $\hat{J}_z$ is an operator that "generates" rotations about the z-axis. Roughly speaking, it tells the quantum state vector how, exactly, to move through space to reflect a small nudge in three-space. The finite rotation operator is then given by

$$\hat{R}(\phi\mathbf{k}) = \lim_{N\to\infty}\left[1 - \frac{i}{\hbar}\hat{J}_z\left(\frac{\phi}{N}\right)\right]^N = e^{-i\hat{J}_z\phi/\hbar}.$$

The rotation generator $\hat{J}_z$ has a few nice properties that will help us physically identify it. For one,

$$1 = \hat{R}^\dagger(d\phi\,\mathbf{k})\hat{R}(d\phi\,\mathbf{k}) = \left(1 + \frac{i}{\hbar}\hat{J}_z^\dagger d\phi\right)\left(1 - \frac{i}{\hbar}\hat{J}_z d\phi\right) = 1 + \frac{i}{\hbar}(\hat{J}_z^\dagger - \hat{J}_z)d\phi + O(d\phi^2),$$

meaning $\hat{K}_z^\dagger = \hat{J}_z$; we therefore call $\hat{J}_z$ Hermitian (or self-adjoint). We can also deduce its eigen-things—by Taylor-expanding $\hat{R}$,

$$\hat{R}(\phi\mathbf{k}) = 1 - \frac{i\phi\hat{J}_z}{\hbar} + \frac{1}{2!}\left(-\frac{i\phi\hat{J}_z}{\hbar}\right)^2 + \cdots,$$

we can immediately see that its eigenkets are $|\pm\mathbf{z}\rangle$ since no other ket would yield an eigenvalue equation. We can also see that having respective eigenvalues of $\pm\hbar/2$ would yield $\hat{R}(\phi\mathbf{k})|\pm\mathbf{z}\rangle = e^{\mp i\phi/2}|\pm\mathbf{z}\rangle$, meaning $\hat{R}$ changes the states only by an overall phase (so effectively not at all). Comparing this with the result of the Stern-Gerlach experiment allows us to identify $\hat{J}_z$ with the angular momentum in the z-direction!

We'll also find it useful to define a pair of projection operators, $\hat{P}_\pm = |\pm\mathbf{z}\rangle\langle\pm\mathbf{z}|$, which project the components of a ket along $|\pm\mathbf{z}\rangle$. Its eigenkets are $|\pm\mathbf{z}\rangle$ with eigenvalues 1 and 0, depending on which operator we're using, and the identities $\hat{P}_\pm^2 = \hat{P}_\pm$ and $\hat{P}_\pm\hat{P}_\mp = 0$ have nice geometric intuition.

More essential to our future discussion, however, is to define an identity ("do-nothing") operator

$$1_z = |+\mathbf{z}\rangle\langle+\mathbf{z}| + |-\mathbf{z}\rangle\langle-\mathbf{z}|,$$

which reflects how the kets $|+_z\rangle$ and $|-_z\rangle$ form a complete basis of spin states. (We specify 1_z to emphasize that we're writing the identity 1 in the S_z-basis.) We'll soon see how useful it is to insert the identity into different places to illuminate new things about the expressions we write!

Take our fourth Stern-Gerlach experiment, for example. The modSG**x** device there acts as the identity operator as it does not modify the incoming $|+\mathbf{z}\rangle$ state, so the amplitude of measuring $|-\mathbf{z}\rangle$ vanishes—that is, $\langle-\mathbf{z}|+\mathbf{z}\rangle = 0$. But to illuminate more about the situation we can insert the identity 1_x into the bracket:

$$\langle-\mathbf{z}|+\mathbf{z}\rangle = \langle-\mathbf{z}|\Big(|+\mathbf{z}\rangle\langle+\mathbf{z}| + |-\mathbf{z}\rangle\langle-\mathbf{z}|\Big)|+\mathbf{z}\rangle$$
$$= \langle-\mathbf{z}|+\mathbf{x}\rangle\langle+\mathbf{x}|+\mathbf{z}\rangle + \langle-\mathbf{z}|-\mathbf{x}\rangle\langle-\mathbf{x}|+\mathbf{z}\rangle.$$

This hearkens back to the laws of probability we're familiar with. Each term represents the amplitude that a particle takes a particular "path" through the modSG**x** device (and so we add them), while each factor represents a particular "step" in each of these paths (and so we multiply them).

1.4 Matrix Mechanics

We will find it convenient to represent kets $|\psi\rangle = c_+|+\mathbf{z}\rangle + c_-|-\mathbf{z}\rangle$ using column vectors and their corresponding bras using row vectors:

$$|\psi\rangle \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} c_+ \\ c_- \end{bmatrix}, \qquad \langle\psi| \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} c_+^* & c_-^* \end{bmatrix}.$$

In this way, brackets are represented by simple matrix multiplication. Note that this representation of a quantum state is dependent on what basis we choose to write it in—if we were to choose the S_x-basis, for example, the components would be very different.

If kets are represented by column vectors, then operators are represented by matrices that act on those column vectors (and their adjoints are their conjugate transposes). These representations, of course, are also basis-dependent; for example,

$$\hat{A}|\psi\rangle \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} \langle+\mathbf{z}|\hat{A}|\psi\rangle \\ \langle-\mathbf{z}|\hat{A}|\psi\rangle \end{bmatrix} = \begin{bmatrix} \langle+\mathbf{z}|\hat{A}|+\mathbf{z}\rangle & \langle+\mathbf{z}|\hat{A}|-\mathbf{z}\rangle \\ \langle-\mathbf{z}|\hat{A}|+\mathbf{z}\rangle & \langle-\mathbf{z}|\hat{A}|-\mathbf{z}\rangle \end{bmatrix} \begin{bmatrix} \langle+\mathbf{z}|\psi\rangle \\ \langle-\mathbf{z}|\psi\rangle \end{bmatrix}.$$

There's a couple of ways we could've arrived here. Mathematically, we've inserted the identity operator 1_x between $\hat{A}$ and $|\psi\rangle$ in each component. Geometrically, we can see that each column of the matrix is simply the column-vector representation of $\hat{A}|\pm\mathbf{z}\rangle$ in the S_z-basis.

This representation suggests that we can view our operators from a couple of different perspectives. On one hand we might view $\hat{A}$ as operating to the left and literally modifying the quantum state; this is called an active transformation. But on the other, $\hat{A}$ may operate to the left and perform the adjoint (inverse) transformation to the basis states; this is called a passive transformation. For example, for two kets related by $|\psi'\rangle = \hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|\psi\rangle$ we can write

$$|\psi'\rangle \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} \langle+\mathbf{z}|\psi'\rangle \\ \langle-\mathbf{z}|\psi'\rangle \end{bmatrix} = \begin{bmatrix} \langle+\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|\psi\rangle \\ \langle-\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|\psi\rangle \end{bmatrix} = \begin{bmatrix} \langle+\mathbf{x}|\psi\rangle \\ \langle-\mathbf{x}|\psi\rangle \end{bmatrix} \xleftarrow[S_z \text{ basis}]{} |\psi\rangle,$$

where we've temporarily defined $|+\mathbf{x}\rangle = \hat{R}(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle$ and $|-\mathbf{x}\rangle = \hat{R}(\frac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle$ for convenience. We can use these passive transformations to perform changes of basis on kets,

$$\begin{bmatrix} \langle+\mathbf{x}|\psi\rangle \\ \langle-\mathbf{x}|\psi\rangle \end{bmatrix} = \begin{bmatrix} \langle+\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle & \langle+\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle \\ \langle-\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|+\mathbf{z}\rangle & \langle-\mathbf{z}|\hat{R}^\dagger(\frac{\pi}{2}\mathbf{j})|-\mathbf{z}\rangle \end{bmatrix} \begin{bmatrix} \langle+\mathbf{z}|\psi\rangle \\ \langle-\mathbf{z}|\psi\rangle \end{bmatrix},$$

but for computations it may be more convenient to simply insert 1_x between $\langle\pm\mathbf{x}|$ and $|\psi\rangle$:

$$= \begin{bmatrix} \langle+\mathbf{x}|+\mathbf{z}\rangle & \langle+\mathbf{x}|-\mathbf{z}\rangle \\ \langle-\mathbf{x}|+\mathbf{z}\rangle & \langle-\mathbf{x}|-\mathbf{z}\rangle \end{bmatrix} \begin{bmatrix} \langle+\mathbf{z}|\psi\rangle \\ \langle-\mathbf{z}|\psi\rangle \end{bmatrix}.$$

We will denote this matrix by $\mathbb{S}_{x\leftarrow z} = \mathbb{S}_{z\leftarrow x}^\dagger$. Note that $\mathbb{S}$ is unitary for all bases, since it is simply the matrix representation of a rotation operator. We can also do change of basis on operators:

$$\hat{A} \xrightarrow[S_x \text{ basis}]{} \mathbb{S}_{z\leftarrow x}^\dagger \mathbb{A}_z \mathbb{S}_{z\leftarrow x},$$

where $\mathbb{A}_z$ is the representation of $\hat{A}$ in the S_z-basis. Mathematically, we have inserted $1_x 1_z$ and $1_z 1_x$ before and after $\mathbb{A}_z$; geometrically, we are "rotating" into the S_z-basis, performing $\hat{A}$, and then rotating back.

We can use matrix mechanics to determine expectation values, too. For example, we can write

$$\langle S_z \rangle = \langle\psi|\hat{J}_z|\psi\rangle.$$

It's clear why this works when we're working in the $\hat{J}_z$-eigenbasis, since in this case the matrix representation is diagonal with eigenvalue entries. But the great thing about expressing expectation values in this way is that it doesn't matter what representation we use—as long as we're consistent, we'll get the correct value!

Now, what we've done here applies to any two-state system. For example, we can also represent the polarization of a photon using two orthogonal basis states $|x\rangle$ and $|y\rangle$. Experiment suggests that rotating through a counterclockwise angle ϕ gives

$$\begin{aligned} |x'\rangle &= \cos\phi\,|x\rangle + \sin\phi\,|y\rangle, \\ |y'\rangle &= -\sin\phi\,|x\rangle + \cos\phi\,|y\rangle. \end{aligned}$$

Note that we could have derived this from the geometry of projecting the primed states onto the unprimed ones. From here we could get the matrix that transforms from the unprimed basis to the primed one!

Lastly, in order to encode circular polarization we define

$$|R\rangle = \frac{1}{\sqrt{2}}\big(|x\rangle + i|y\rangle\big), \qquad |L\rangle = \frac{1}{\sqrt{2}}\big(|x\rangle - i|y\rangle\big).$$

If we were to rotate these through a counterclockwise angle ϕ about the z-axis, we'd find that $|R'\rangle = e^{-\phi}|R\rangle$ and $|L'\rangle = e^{i\phi}|L\rangle$. This means that these must be eigenstates of $\hat{J}_z$, this time with eigenvalues $\pm\hbar$! This reflects how photons are spin-1 particles (and, because they're massless, how they cannot have spin 0).

At this point we've hinted at all the key postulates of quantum mechanics.

- We will describe the state of a quantum system as a vector in some abstract "Hilbert space";

- each observable A of such a system will correspond to a Hermitian operator $\hat{A}$;

- the eigenstates of this operator are those states with definite A and the corresponding eigenvalues are the results of measuring A; and

- the symmetries of such a system will correspond to unitary operators generated by their respective Noether quantity.

Formalism in hand, it's time to dive into the physics!

2 Quantized Observables

2.1 The Angular Momentum Eigenstates

The first observable we'll discuss is angular momentum, the generator of rotations. We begin with the observation that the rotation operators about different axes don't commute one another—in particular, a simple second-order Taylor expansion of the rotation matrices about the x, y, and z axes reveals that

$$\mathbb{S}(\Delta\phi\,\mathbf{i})\mathbb{S}(\Delta\phi\,\mathbf{j}) - \mathbb{S}(\Delta\phi\,\mathbf{j})\mathbb{S}(\Delta\phi\,\mathbf{i}) = \mathbb{S}(\Delta\phi^2\mathbf{k}) - \mathbb{I}.$$

If we were to replace each matrix here with its corresponding $\hat{R}$, also expanded to second order, we could do some algebra to find that $\hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x = i\hbar\hat{J}_z$. We'll call this the commutator $[\hat{J}_x, \hat{J}_y]$, and every pair of operators has one:

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \quad [\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x, \quad [\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y.$$

So none of the angular momentum operators commute with one another! Interestingly, though, it's easy to show that the total angular momentum $\hat{\mathbf{J}}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ commutes with them all. This has important implications for the operators' eigen-stuff—because $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ commute, they have a complete set of eigenstates $|\lambda, m\rangle$ in common:

$$\hat{\mathbf{J}}^2|\lambda, m\rangle = \lambda\hbar^2|\lambda, m\rangle,$$

$$\hat{J}_z|\lambda, m\rangle = m\hbar|\lambda, m\rangle,$$

where $\hbar$ has been introduced to keep λ and m dimensionless. Our aim, now, is to identify these eigenstates by placing restrictions on their eigenvalues. In doing this we'll find it useful to define the operators

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y,$$

which notably satisfy $\hat{J}_+^\dagger = \hat{J}_-$ and $[\hat{J}_z, \hat{J}_\pm] = \pm\hbar\hat{J}_\pm$. To see the effect that these operators have on a state $|\lambda, m\rangle$ we evaluate $\hat{J}_z\hat{J}_\pm|\lambda, m\rangle$, using the commutation relations to swap the order of the operators:

$$\hat{J}_z\hat{J}_\pm|\lambda, m\rangle = (\hat{J}_\pm\hat{J}_z \pm \hbar\hat{J}_\pm)|\lambda, m\rangle$$

$$= (\hat{J}_\pm m\hbar + \hbar\hat{J}_\pm)|\lambda, m\rangle$$

$$= (m \pm 1)\hbar\hat{J}_\pm|\lambda, m\rangle.$$

We can see that the functions of these operators, then, are to produce a new quantum state with a raised or lowered value of m. We therefore call $\hat{J}_+$ and $\hat{J}_-$ raising and lowering operators. (We are assured that λ is preserved because $\hat{J}_\pm$ commute with $\hat{\mathbf{J}}^2$.)

We'd expect that $\lambda \geq 0$ since the the "magnitude" of $\mathbf{J}^2$ should be non-negative. We can verify this using the expectation value

$$\langle\lambda, m|(\hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2)|\lambda, m\rangle = \lambda\hbar^2 :$$

each term on the left looks like $\langle\lambda, m|\hat{J}_x^2|\lambda, m\rangle = c_x^* c_x \langle\psi_x|\psi_x\rangle$ for some normalized "output" state $|\psi_x\rangle$; it is known that $c_x^* c_x \geq 0$ and $\langle\psi_x|\psi_x\rangle = 1$, so each term on the left is non-negative and so is λ. We thus also have

$$0 \leq \langle\lambda, m|(\hat{J}_x^2 + \hat{J}_y^2)|\lambda, m\rangle$$

$$= \langle\lambda, m|(\hat{\mathbf{J}}^2 - \hat{J}_z^2)|\lambda, m\rangle$$

$$= (\lambda - m^2)\hbar^2\langle\lambda, m|\lambda, m\rangle,$$

meaning $m^2 \leq \lambda$, as we'd expect since $|J_z^2| < |\mathbf{J}^2|$. Every λ therefore has maximum and minimum values of m associated with it; call them j and j'. We must have $\hat{J}_+|\lambda, j\rangle = 0$ and $\hat{J}_-|\lambda, j'\rangle = 0$, and so

$$\hat{J}_-\hat{J}_+|\lambda, j\rangle = (\hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|\lambda, j\rangle \qquad\qquad \hat{J}_+\hat{J}_-|\lambda, j\rangle = (\hat{\mathbf{J}}^2 - \hat{J}_z^2 + \hbar\hat{J}_z)|\lambda, j\rangle$$

$$= (\lambda - j^2 - j)\hbar^2|\lambda, j\rangle = 0, \qquad\qquad\qquad = (\lambda - j'^2 + j')\hbar^2|\lambda, j'\rangle = 0.$$

Thus $j^2 + j = \lambda = j'^2 - j'$ and so $j' = -j$. Finally, because we climb from j' to j in size-1 steps, the distance $2j$ between the two must be an integer and we deduce that

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots, \qquad m = j, j-1, \ldots, -j+1, -j.$$

This motivates a slight change in notation. We'll denote a simultaneous eigenstate of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ by $|j, m\rangle$:

$$\hat{\mathbf{J}}^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle,$$
$$\hat{J}_z|j, m\rangle = m\hbar|j, m\rangle.$$

2.2 Angular Momentum in a Spin-1/2 Particle

We're now in a position to determine the spin states of a spin-$\frac{1}{2}$ particle using some of the linear algebra we're more familiar with. We'll first determine the matrix representations of the raising and lowering operators. It is known that

$$\hat{J}_+|j, m\rangle = c_+\hbar|j, m+1\rangle, \qquad \hat{J}_-|j, m\rangle = c_-\hbar|j, m-1\rangle$$

for some c_+, c_-. To determine c_+ we can take the inner product of the first equation with itself to get

$$
\begin{aligned}
c_+^* c_+ \hbar^2 \langle j,\, m+1 | j,\, m+1 \rangle &= \langle j, m | \hat{J}_- \hat{J}_+ | j, m \rangle \\
&= \langle j, m | (\hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z) | j, m \rangle \\
&= [j(j+1) - m^2 - m]\hbar^2 \langle j, m | j, m \rangle.
\end{aligned}
$$

Assuming that our eigenstates are normalized, we conclude that

$$\hat{J}_+|j, m\rangle = \sqrt{j(j+1) - m(m+1)}\,\hbar|j,\, m+1\rangle.$$

By an analogous line of reasoning,

$$\hat{J}_-|j, m\rangle = \sqrt{j(j+1) - m(m-1)}\,\hbar|j,\, m-1\rangle.$$

The matrix elements of the raising and lowering operators are therefore

$$\langle j, m'|\hat{J}_\pm|j, m\rangle = \sqrt{j(j+1) - m(m \pm 1)}\,\hbar\,\delta_{m',m\pm 1}.$$

We can see that the nonzero elements of the raising operator are just above the main diagonal, while those of the lowering operator are just below. Using the $j = 1/2$ states as a basis gives

$$\hat{J}_+ \to \hbar \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \hat{J}_- \to \hbar \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Now we'll hone in specifically on spin angular momentum, so instead of using J to refer to a generic form of angular momentum we'll use S to refer to spin. We can use the definitions of the raising and lowering operators to solve for the Pauli spin matrices,

$$\hat{S}_x \to \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \frac{\hbar}{2}\sigma_x, \qquad \hat{S}_y \to \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \frac{\hbar}{2}\sigma_y, \qquad \hat{S}_z \to \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \frac{\hbar}{2}\sigma_z,$$

and then use these to get a matrix representation for $\hat{\mathbf{S}} = \hat{S}_x\mathbf{i} + \hat{S}_y\mathbf{j} + \hat{S}_z\mathbf{k}$. We can then determine the eigenstates of $\hat{S}_n = \hat{\mathbf{S}} \cdot \mathbf{n}$—the spin-up and spin-down states along an arbitrary axis $\mathbf{n}$—by solving the eigenvalue equation

$$\hat{S}_n|\mu\rangle = \mu\frac{\hbar}{2}|\mu\rangle$$

in matrix form. Doing this for an $\mathbf{n}$ in the xy-plane yields the familiar

$$|\pm\mathbf{n}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle \pm \frac{e^{i\phi}}{\sqrt{2}}|-\mathbf{z}\rangle.$$

We could run through a very similar process to determine the eigenstates of a spin-1 particle! Doing this along, say, the y-axis involves solving the equation $\hat{S}_y|1, \mu\rangle_y = \mu\hbar|1, \mu\rangle_y$ in matrix form to get, as one of the eigenkets,

$$|1, 1\rangle_y = \frac{1}{2}|1, 1\rangle + i\frac{\sqrt{2}}{2}|1, 0\rangle - \frac{1}{2}|1, -1\rangle.$$

Note that we put subscripts on the kets to denote a basis other than S_z.

2.3 Time Evolution

Having discussed how rotations relate to angular momentum, we'll now look at another operator $\hat{U}(t)$ which translates kets forward in time via $\hat{U}(t)|\psi(0)\rangle = |\psi(t)\rangle$. Like with $\hat{R}$, we require that $\hat{U}$ be unitary in order to conserve probability, and we define it in terms of a Hermitian operator $\hat{H}$ that satisfies

$$\hat{U}(dt) = 1 - \frac{i}{\hbar}\hat{H}\,dt.$$

If $\hat{H}$ is time-independent then we get $\hat{U}(t) = e^{-i\hat{H}t/\hbar}$, also like before. Now we can see that $\hat{H}$ has units of energy, and because $\hat{H}$ commutes with $\hat{U}$ the observable associated with $\hat{H}$ has a time-independent expectation value:

$$\langle\psi(t)|\hat{H}|\psi(t)\rangle = \langle\psi(0)|\hat{U}^\dagger\hat{H}\hat{U}|\psi(0)\rangle = \langle\psi(0)|\hat{H}|\psi(0)\rangle.$$

We therefore identify $\hat{H}$ as the energy operator, more commonly known as the Hamiltonian, and so it has eigenstates $|E\rangle$ satisfying

$$\hat{H}|E\rangle = E|E\rangle, \qquad \langle E\rangle = \langle\psi|\hat{H}|\psi\rangle.$$

So if the initial state of a system is an energy eigenstate $|\psi(0)\rangle = |E\rangle$, we have

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi(0)\rangle = e^{-iEt/\hbar}|E\rangle.$$

Thus time evolution simply causes $|E\rangle$ to pick up an overall phase (i.e., essentially stay the same), so we call $|E\rangle$ a stationary state. We only get interesting behavior when we time-evolve superpositions of these states.

The Hamiltonian has a couple of properties which make it an especially powerful tool in quantum mechanics. Firstly, we can see that the time evolution operator $\hat{U}$ satisfies the differential equation

$$\hat{U}(t+dt) = \left(1 - \frac{i}{\hbar}\hat{H}\,dt\right)\hat{U}(t) \implies i\hbar\frac{d}{dt}\hat{U}(t) = \hat{H}\hat{U}(t),$$

leading to the Schrödinger equation

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

This is the fundamental equation of motion in quantum mechanics—it provides all the information we need to determine how a quantum state will evolve in time. We can leverage this equation to determine whether an arbitrary observable A has a time-dependent expectation value:

$$\begin{aligned}
\frac{d\langle A\rangle}{dt} &= \left(\frac{d}{dt}\langle\psi(t)|\right)\hat{A}|\psi(t)\rangle + \langle\psi(t)|\hat{A}\left(\frac{d}{dt}|\psi(t)\rangle\right) + \langle\psi(t)|\frac{\partial A}{\partial t}|\psi(t)\rangle \\
&= \left(\frac{1}{-i\hbar}\langle\psi(t)|\hat{H}\right)\hat{A}|\psi(t)\rangle + \langle\psi(t)|\hat{A}\left(\frac{1}{i\hbar}\hat{H}|\psi(t)\rangle\right) + \langle\psi(t)|\frac{\partial A}{\partial t}|\psi(t)\rangle \\
&= \frac{i}{\hbar}\langle\psi(t)|[\hat{H},\hat{A}]|\psi(t)\rangle + \langle\psi(t)|\frac{\partial\hat{A}}{\partial t}|\psi(t)\rangle.
\end{aligned}$$

Thus a time-independent operator $\hat{A}$ has a time-independent $\langle A\rangle$ if and only if it commutes with $\hat{H}$.

2.4 Examples of Quantum Dynamics

As a simple example of how all this plays out in actual quantum systems, consider a spin-1/2 particle in a constant magnetic field $\mathbf{B} = B_0\mathbf{k}$. The Hamiltonian of this system is

$$\hat{H} = -\hat{\boldsymbol{\mu}}\cdot\mathbf{B} = \omega_0\hat{S}_z, \qquad \omega_0 = \frac{geB_0}{2mc},$$

so it has energy eigenstates $\hat{H}|\pm\mathbf{z}\rangle = \pm(\hbar\omega_0/2)|\pm\mathbf{z}\rangle$. Since $\hat{H}$ is time-independent, the time evolution operator is

$$\hat{U}(t) = e^{-i\hat{S}_z\phi/\hbar} = \hat{R}(\phi\,\mathbf{k}), \qquad \phi = \omega_0 t.$$

So as time passes, the state gets rotated about the z-axis! For instance, if we take $|\psi(0)\rangle = |+\mathbf{x}\rangle$ then we'd find

$$|\psi(t)\rangle = e^{-i\omega_0 t/2}\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle + \frac{e^{i\omega_0 t}}{\sqrt{2}}|-\mathbf{z}\rangle\right),$$

the spin expectation values of which are what we'd expect based on the symmetry of the system. (Specifically, we get $\langle S_z\rangle = 0$, $\langle S_x\rangle = (\hbar/2)\cos\omega_0 t$, and $\langle S_y\rangle = (\hbar/2)\sin\omega_0 t$.) Note that we could've determined that S_z was a constant of motion ahead of time by leveraging the rotational symmetry of the system about the z-axis—rotations about this axis must leave $\hat{H}$ unchanged, meaning $\hat{S}_z$ commutes with $\hat{H}$ and so $\langle S_z\rangle$ is constant, unlike the other observables.

Suppose we complicate things by letting $\mathbf{B}$ oscillate very slightly in the x-direction, so $\mathbf{B} = B_1(\cos\omega t)\mathbf{i} + B_0\mathbf{k}$. The Hamiltonian is

$$\hat{H} = \omega_0\hat{S}_z + \omega_1(\cos\omega t)\hat{S}_x,$$

where ω_0, ω_1 are defined as before for their respective $\mathbf{B}$-fields. This $\hat{H}$ is time-dependent, so we appeal to the Schrödinger equation

$$i\hbar\begin{bmatrix}\dot{a}(t)\\ \dot{b}(t)\end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix}\omega_0 & \omega_1\cos\omega t\\ \omega_1\cos\omega t & -\omega_0\end{bmatrix}\begin{bmatrix}a(t)\\ b(t)\end{bmatrix}$$

as it is written in the S_z-basis. This system is intractable as it stands, but we can exploit $B_1 \ll B_0$ to obtain an approximate solution. Motivated by the fact that the solution for $\omega_1 = 0$ is $a(t) = a(0)e^{-i\omega_0 t/2}$ and $b(t) = b(0)e^{i\omega_0/2}$, we substitute

$$\begin{bmatrix}a(t)\\ b(t)\end{bmatrix} = \begin{bmatrix}c(t)e^{-i\omega_0 t/2}\\ d(t)e^{i\omega_0 t/2}\end{bmatrix}$$

into the Schrödinger equation and do some algebra to get, after some nice cancellations,

$$i\begin{bmatrix}\dot{c}(t)\\ \dot{d}(t)\end{bmatrix} = \frac{\omega_1}{4}\begin{bmatrix}\left(e^{i(\omega_0+\omega)t} + e^{i(\omega_0-\omega)t}\right)d(t)\\ \left(e^{i(\omega-\omega_0)t} + e^{i(\omega_0+\omega)t}\right)c(t)\end{bmatrix}.$$

For most ω these exponentials will oscillate very quickly compared to $c(t)$ or $d(t)$ and average out to zero. But for $\omega \sim \omega_0$, we can neglect the terms oscillating at $\omega_0 + \omega$ in favor of the $\omega_0 - \omega$ ones.

Solving in the $\omega = \omega_0$ case we'd find, for example, that the probabilities of measuring $|\pm\mathbf{z}\rangle$ for a particle with definite initial spin oscillate with frequency $\omega_1/4$. In fact, there is a probability of measuring a "spin flip" for any $\omega \sim \omega_0$; there's just a resonant peak at ω_0. (The energy difference in the spin-up and spin-down cases comes from interaction with the electromagnetic field.)

As a final example, let the kets $|1\rangle$ and $|2\rangle$ denote configurations of the ammonia molecule in which the nitrogen atom is above and below the plane of the hydrogen atoms, respectively. The symmetry of the system dictates that

$$\hat{H} \longrightarrow \begin{bmatrix}\langle 1|\hat{H}|1\rangle & \langle 1|\hat{H}|2\rangle\\ \langle 2|\hat{H}|1\rangle & \langle 2|\hat{H}|2\rangle\end{bmatrix} = \begin{bmatrix}E_0 & -A\\ -A & E_0\end{bmatrix},$$

where the value $A > 0$ is determined experimentally. (These off-diagonal entries exist because there is a finite potential between $|1\rangle$ and $|2\rangle$.) The eigenvalues of $\hat{H}$ are $E_I = E_0 - A$ and $E_{II} = E_0 + A$, and they have eigenkets

$$|I\rangle = \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle, \qquad |II\rangle = \frac{1}{\sqrt{2}}|1\rangle - \frac{1}{\sqrt{2}}|2\rangle.$$

Using this, we could find that an initial state $|1\rangle$ oscillates above and below the plane with frequency $\nu = 2A/\hbar$. Now, if we once again complicate things by introducing a static electric field $\mathbf{E}$ there is an interaction energy $-\hat{\boldsymbol{\mu}}_e \cdot \mathbf{E}$ and so

$$\hat{H} \longrightarrow \begin{bmatrix}E_0 + \mu_e|\mathbf{E}| & -A\\ -A & E_0 - \mu_e|\mathbf{E}|\end{bmatrix}.$$

The energies are

$$E = E_0 \pm \sqrt{(\mu_e|\mathbf{E}|)^2 + A^2} \simeq E_0 \pm A \pm \frac{1}{2}\frac{(\mu_e|\mathbf{E}|)^2}{A}.$$

We therefore take a measurement of $|I\rangle$ versus $|II\rangle$ by shooting molecules through an inhomogeneous electric field with a high gradient, similar to what we saw with Stern-Gerlach devices, separating the beam of molecule in two according to the sign of $\pm$. In the case of a time-dependent electric field like $\mathbf{E} = \mathbf{E}_0\cos\omega t$ then we can induce resonance in a fashion practically identical to what we saw a bit earlier.

2.5 Uncertainty Relations

We've seen that commuting operators have simultaneous eigenstates, but now we'll make a very important note about the quantitative conclusions we can draw when two operators do not commute. Let $[\hat{A}, \hat{B}] = i\hat{C}$ with all operators Hermitian, and define $|\alpha\rangle = (\hat{A} - \langle A\rangle)|\psi\rangle$ and $|\beta\rangle = (\hat{B} - \langle B\rangle)|\psi\rangle$ so that $\langle\alpha|\alpha\rangle = (\Delta A)^2$ and $\langle\beta|\beta\rangle = (\Delta B)^2$. Now we'd like to compute $\langle\alpha|\beta\rangle$ to apply the Schwarz inequality $\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2$.

To this end, define $\hat{O} = (\hat{A} - \langle A\rangle)(\hat{B} - \langle B\rangle)$ so that

$$\langle\alpha|\beta\rangle = \langle\psi|\hat{O}|\psi\rangle = \langle\psi|\left(\frac{\hat{O} + \hat{O}^\dagger}{2} + \frac{\hat{O} - \hat{O}^\dagger}{2}\right)|\psi\rangle.$$

We can see that $\hat{O} - \hat{O}^\dagger = [\hat{A}, \hat{B}] = i\hat{C}$, so

$$|\langle\alpha|\beta\rangle|^2 = \left|\langle\psi|\left(\frac{\hat{O} + \hat{O}^\dagger}{2}\right)|\psi\rangle + \frac{i}{2}\langle\psi|\hat{C}|\psi\rangle\right|^2 \geq \frac{|\langle C\rangle|^2}{4},$$

and thus

$$(\Delta A)^2(\Delta B)^2 \geq \frac{|\langle C\rangle|^2}{4} \implies \Delta A\,\Delta B \geq \frac{|\langle C\rangle|}{2}.$$

This lines up nicely with the results of our Stern-Gerlach experiments. If we have $\hat{C} = S_z$ and take a measurement of S_z, then ΔJ_z is certainly nonzero since we can get either spin-up or spin-down. Thus ΔJ_x and ΔJ_y must also be nonzero. (This also means angular momentum can't point in any particular direction—otherwise, we'd know all three components at once, a violation of the uncertainty principle!)

There is one very important uncertainty relation that is similar in form but different in character to the ones described above:
$$\Delta E \Delta t \geq \frac{\hbar}{2}.$$

Here Δt is not an uncertainty—time is a parameter, not an observable. Instead, it denotes a state's "evolutionary time", which is the time it takes for the expectation value of an observable A to change significantly (i.e., by an amount on the order of its uncertainty). Mathematically, we can take advantage of the statements

$$\Delta A \Delta E \geq \frac{1}{2}\left|\langle\psi|[\hat{A}, \hat{H}]|\psi\rangle\right|, \qquad \frac{d\langle A\rangle}{dt} = \frac{i}{\hbar}\langle\psi|[\hat{H}, \hat{A}]|\psi\rangle$$

to write

$$\Delta A \Delta E \geq \frac{\hbar}{2}\left|\frac{d\langle A\rangle}{dt}\right| \implies \Delta t = \frac{\Delta A}{|d\langle A\rangle/dt|}.$$

This is where the natural linewidth comes from. The excited states of an atom are not stationary states, and so they decay on some characteristic timescale Δt with a corresponding uncertainty in energy ΔE.

2.6 Multiparticle Systems

Now we'll take a step up and work with quantum states of multiple particles, specifically of spin-1/2 particles. Such a state can be written as, for example,

$$|+\mathbf{z}, +\mathbf{z}\rangle = |+\mathbf{z}\rangle_1 \otimes |+\mathbf{z}\rangle_2 = |+\mathbf{z}\rangle_1|+\mathbf{z}\rangle_2,$$

where $\otimes$ denotes a direct product. (In practice we omit the $\otimes$ because the alternatives are unambiguous.) Conventionally we use the basis

$$|1\rangle = |+\mathbf{z}\rangle_1|+\mathbf{z}\rangle_2, \quad |2\rangle = |+\mathbf{z}\rangle_1|-\mathbf{z}\rangle_2, \quad |3\rangle = |-\mathbf{z}\rangle_1|+\mathbf{z}\rangle_2, \quad |4\rangle = |-\mathbf{z}\rangle_1|-\mathbf{z}\rangle_2.$$

The generators of rotations for particle 1 satisfy $\hat{\mathbf{S}}_1 = \hat{S}_{1x}\mathbf{i} + \hat{S}_{1y}\mathbf{j} + \hat{S}_{1z}\mathbf{k}$. Since rotations of particle 1 and particle 2 are independent of one another, $[\hat{\mathbf{S}}_1, \hat{\mathbf{S}}_2] = 0$, and likewise for the generators about each axis. We can also define the total spin operator $\hat{\mathbf{S}}$ via

$$\hat{R}(d\theta\mathbf{n}) = 1 - \frac{i}{\hbar}\hat{\mathbf{S}}\cdot\mathbf{n}\,d\theta = \left(1 - \frac{i}{\hbar}\hat{\mathbf{S}}_1\cdot\mathbf{n}\,d\theta\right) \otimes \left(1 - \frac{i}{\hbar}\hat{\mathbf{S}}_2\cdot\mathbf{n}\,d\theta\right),$$

from which we can see that, to first order, $\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 \otimes 1 + 1 \otimes \hat{\mathbf{S}}_2 = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2$. At this point we could show that the components of $\hat{\mathbf{S}}$ obey the expected commutation relations.

As an example, we'll look at one of the most important spin-1/2 systems: the hydrogen atom. We'll discuss it in more detail later on, but for now we'll focus on the "hyperfine" energy structure due to the magnetic spin-spin interaction between the proton and electron, specifically the case in which there is no orbital angular momentum:

$$\hat{H} = \frac{2A}{\hbar^2}\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 = \frac{A}{\hbar^2}\left(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} + 2\hat{S}_{1z}\hat{S}_{2z}\right)$$

for some experimentally-determined A with dimensions of energy. We've expressed $\hat{H}$ in terms of raising and lowering operators because that makes it easier to determine the matrix elements, especially because most of the elements are zero:

$$\hat{H} \longrightarrow \begin{bmatrix} A/2 & 0 & 0 & 0 \\ 0 & -A/2 & A & 0 \\ 0 & A & -A/2 & 0 \\ 0 & 0 & 0 & A/2 \end{bmatrix}.$$

This produces one "singlet" eigenket with $E = -3A/2$ and three "triplet" eigenkets with $E = A/2$; these kets are given below. We could show that each is also an eigenstate of both $\mathbf{S}_1^2$ and $\mathbf{S}_2^2$, so the energy eigenstates are also states of definite total spin! It's easy to show that the singlet has $s = 0$ and the triplets have $s = 1$, and so we label the eigenkets

$$|0,0\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z},-\mathbf{z}\rangle - \frac{1}{\sqrt{2}}|-\mathbf{z},+\mathbf{z}\rangle, \qquad\qquad |1,1\rangle = |+\mathbf{z},+\mathbf{z}\rangle,$$

$$|1,0\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z},-\mathbf{z}\rangle + \frac{1}{\sqrt{2}}|-\mathbf{z},+\mathbf{z}\rangle, \qquad\qquad |1,-1\rangle = |-\mathbf{z},-\mathbf{z}\rangle.$$

There's another way we could've come to this conclusion. We could have immediately recognized that $|1,1\rangle = |+\mathbf{z},+\mathbf{z}\rangle$ and $|1,-1\rangle = |-\mathbf{z},-\mathbf{z}\rangle$ because they are eigenstates of $\hat{S}_z$, and thus of $\hat{\mathbf{S}}^2$ with the proper eigenvalue. For the other two, we apply

$$\hat{S}_-|1,1\rangle = (\hat{S}_{1-} + \hat{S}_{2-})|+\mathbf{z},+\mathbf{z}\rangle,$$

$$\sqrt{2}\,\hbar|1,0\rangle = \hbar(|-\mathbf{z},+\mathbf{z}\rangle + |+\mathbf{z},-\mathbf{z}\rangle)$$

and normalize to get our expression for $|1,0\rangle$. For $|1,1\rangle$ we take the normalized ket that's orthogonal to $|1,0\rangle$.

Notice, however, that having anti-aligned spins in the z-basis isn't a sufficient condition for having total spin zero because $|1,0\rangle$ and $|0,0\rangle$ both exist. It turns out that $|1,0\rangle$ corresponds to a spin configuration that is always aligned on the x-axis, while $|0,0\rangle$ has anti-aligned spins in every basis.

2.7 Quantum Entanglement

Consider a spin-0 particle at rest that decays into two spin-1/2 particles. The particles are moving in opposite directions and comprise a spin state $|0,0\rangle$, so if a measurement of particle 1's spin gives $S_{1z} = \hbar/2$ then we must get $S_{2z} = -\hbar/2$ for the other particle, and vice versa (and along any axis). Thus measuring particle 1 instantaneously determines particle 2's state, even though the particles are completely non-interacting and may be separated by a very large distance.

Einstein and his colleagues Podolsky and Rosen (EPR) were not happy with this idea. Perhaps, they argued, upon decay each particle is assigned a particular spin state (a "hidden variable") that is simply unknown until we measure it. In particular, suppose we take spin measurements along three coplanar spin axes $\mathbf{a}, \mathbf{b}, \mathbf{c}$ so that in our hidden-variable model each particle belongs to one of eight populations $N_1, \cdots, N_8$. To name a few,

$$\begin{aligned} N_1: & \quad \{+\mathbf{a},+\mathbf{b},+\mathbf{c}\}, \ \{-\mathbf{a},-\mathbf{b},-\mathbf{c}\}, \\ N_2: & \quad \{+\mathbf{a},+\mathbf{b},-\mathbf{c}\}, \ \{-\mathbf{a},-\mathbf{b},+\mathbf{c}\}, \\ & \vdots \\ N_8: & \quad \{-\mathbf{a},-\mathbf{b},-\mathbf{c}\}, \ \{+\mathbf{a},+\mathbf{b},+\mathbf{c}\}. \end{aligned}$$

Suppose we take measurements of the two particles along different axes. For populations N_1 and N_8 this will always result in opposite spin states, while for all other populations this occurs $1/3$ of the time. We could use the $\mathbf{a}$-basis representation of $|0,0\rangle$ along with $\langle+\mathbf{z}|+\mathbf{n}\rangle = \cos(\theta/2)$ to get

$$|\langle+\mathbf{a},-\mathbf{b}|0,0\rangle|^2 + |\langle-\mathbf{a},+\mathbf{b}|0,0\rangle|^2 = \cos^2\frac{\theta_{ab}}{2},$$

where θ_{ab} is the angle between $\mathbf{a}$ and $\mathbf{b}$. We have symmetric results for the other pairs of axes. But if we take $\theta_{ab} = \theta_{bc} = \theta_{ac} = 120°$ then the probability of measuring opposite spins is $1/4$—below $1/3$! So at least one of the two models we've considered thus far is incorrect. If the hidden-variable model is correct then we must have

$$N_3 + N_4 \leq (N_2 + N_4) + (N_3 + N_7),$$

which we can divide by N_{tot} to get the Bell's inequality

$$P(+\mathbf{a};+\mathbf{b}) \leq P(+\mathbf{a};+\mathbf{c}) + P(+\mathbf{c};+\mathbf{b}).$$

This inequality is not consistent with the predictions of quantum mechanics for certain θ-values, so if quantum mechanics is valid and Bell's inequality is violated then no (local) hidden-variable theory can be valid. Experiments have shown, to around 250 standard deviations, that this is actually the case.

As an application, consider the problem of cloning (or copying) a quantum state $|\psi\rangle$. For this to be possible there must exist a unitary operator $\hat{U}$ satisfying $\hat{U}|\psi\rangle_1|e\rangle_2 = |\psi\rangle_1|\psi\rangle_2$ for some initial state $|e\rangle$ independent of $|\psi\rangle$; consequently, we must have the same for some other state $|\varphi\rangle$. Since $\langle e|e\rangle = 1$ we can write

$$\langle\varphi|\psi\rangle = \left(_1\langle\varphi|\,_2\langle e|\right)\left(|\psi\rangle_1|e\rangle_2\right)$$
$$= {_1\langle\varphi|\,_2\langle e|\hat{U}^\dagger\hat{U}|\psi\rangle_1|e\rangle_2}$$
$$= \langle\varphi|\psi\rangle^2,$$

which is not true for arbitrary $|\psi\rangle, |\varphi\rangle$. Thus it is impossible to clone quantum states. We can, however, teleport states with the power of quantum entanglement.

Suppose we'd like to teleport the state $|\psi_1\rangle = a|+\mathbf{z}\rangle_1 + b|-\mathbf{z}\rangle_1$, and further that there is another pair of particles in a total spin-0 state. Before we take any measurements, the three-particle state is

$$|\psi_{123}\rangle = (a|+\mathbf{z}\rangle_1 + b|-\mathbf{z}\rangle_1)\left(\frac{1}{\sqrt{2}}|+\mathbf{z}\rangle_2|-\mathbf{z}\rangle_3 - \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle_2|+\mathbf{z}\rangle_3\right).$$

We say that particles 2 and 3 are entangled because it is impossible factor their two-particle state into independent one-particle states. The idea, now, is to take a very specific measurement in order to entangle particles 1 and 2; this measurement occurs in the Bell basis defined by

$$|\Psi_1^{(\pm)}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle_1|-\mathbf{z}\rangle_2 \pm \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle_1|+\mathbf{z}\rangle_2,$$
$$|\Phi_1^{(\pm)}\rangle = \frac{1}{\sqrt{2}}|+\mathbf{z}\rangle_1|+\mathbf{z}\rangle_2 \pm \frac{1}{\sqrt{2}}|-\mathbf{z}\rangle_1|-\mathbf{z}\rangle_2.$$

In terms of these states,

$$|\psi_{123}\rangle = \frac{1}{2}|\Psi_{12}^{(-)}\rangle\left(-a|+\mathbf{z}\rangle_3 - b|-\mathbf{z}\rangle_3\right) + \frac{1}{2}|\Psi_{12}^{(+)}\rangle\left(-a|+\mathbf{z}\rangle_3 + b|-\mathbf{z}\rangle_3\right)$$
$$+ \frac{1}{2}|\Phi_{12}^{(-)}\rangle\left(b|+\mathbf{z}\rangle_3 + a|-\mathbf{z}\rangle_3\right) + \frac{1}{2}|\Phi_{12}^{(+)}\rangle\left(-b|+\mathbf{z}\rangle + a|-\mathbf{z}\rangle_3\right).$$

Thus by taking a Bell-state measurement we disentangle particle 3 from particle 2, and we can use the result of our measurement to determine precisely how to modify particle 3 to put it in the state we desire. Notably, the original state of particle 1 is destroyed in the process and so the no-cloning theorem remains intact.

2.8   The Density Operator

So far, whenever we've dealt with ensembles of particles we've worked primarily with pure states—that is, ensembles of particles that are all in exactly the same state. Such states, however, are merely an idealization

of the physical world, and so we should develop a systemic way to handle mixed states as well. We do so using the density operator, which is defined by

$$\hat{\rho} = \sum_k p_k |\psi^{(k)}\rangle\langle\psi^{(k)}|$$

where p_k is the probability that a particle is in the state $|\psi^{(k)}\rangle$. The trace of such an operator is always

$$\operatorname{tr}\hat{\rho} = \sum_{i,k} p_k \langle i|\psi^{(k)}\rangle\langle\psi^{(k)}|i\rangle = \sum_k p_k \sum_i \langle\psi^{(k)}|i\rangle\langle i|\psi^{(k)}\rangle = 1.$$

However, since we can always diagonalize the density matrix with diagonal entries p_k, we have

$$\operatorname{tr}\hat{\rho}^2 = \sum_k p_k^2 \leq 1$$

with equality only for a pure state. Thus we can test for a mixed state by computing $\operatorname{tr}\hat{\rho}^2$ and seeing if it's less than one. The expected value of an observable A is the weighted average

$$\langle A \rangle = \sum_k p_k \langle\psi^{(k)}|\hat{A}|\psi^{(k)}\rangle = \sum_{i,j,k} p_k \langle\psi^{(k)}|i\rangle\langle i|\hat{A}|j\rangle\langle j|\psi^{(k)}\rangle = \sum_{i,j} A_{ij}\rho_{ji} = \operatorname{tr}(\hat{A}\hat{\rho}).$$

If we imagine $\hat{A}$ and $\hat{\rho}$ as represented in the basis in which $\hat{\rho}$ is diagonalized, this trace is simply the weighted sum of all the different states' expectation values. Notice that we can take $\hat{A} = \hat{P}_{|\varphi\rangle} = |\varphi\rangle\langle\varphi|$ for some other state $|\varphi\rangle$ to get

$$|\langle\varphi|\psi\rangle|^2 = \operatorname{tr}\left(\hat{P}_{|\varphi\rangle}\hat{\rho}\right).$$

Also, we can do some product rule work on $d\hat{\rho}/dt$ (with the Schrödinger equation) to get the time evolution:

$$\frac{d}{dt}\hat{\rho}(t) = \sum_k \left[\left(\frac{d}{dt}|\psi^{(k)}(t)\rangle\right)\langle\psi^{(k)}(t)| + |\psi^{(k)}(t)\rangle\left(\frac{d}{dt}|\psi^{(k)}(t)\rangle\right)\right]$$

$$= \sum_k \left[\frac{1}{i\hbar}\hat{H}|\psi^{(k)}(t)\rangle\langle\psi^{(k)}(t)| + \frac{1}{-i\hbar}|\psi^{(k)}(t)\rangle\langle\psi^{(k)}(t)|\hat{H}\right] = \frac{1}{i\hbar}[\hat{H},\hat{\rho}(t)].$$

The density matrix for a two-particle system is defined in precisely the same way. To compute the expectation value of, say, $\langle S_{1z}\rangle$, we could either compute $\operatorname{tr}\left((\hat{S}_{1z}\otimes 1)\hat{\rho}\right)$ or we can define the reduced density operator

$$\hat{\rho}^{(1)} = \sum_k {}_2\langle k|\hat{\rho}|k\rangle_2$$

and compute $\operatorname{tr}(\hat{S}_{1z}\hat{\rho}^{(1)})$. The matrix elements of $\hat{\rho}^{(1)}$ are ${}_1\langle i|\hat{\rho}^{(1)}|j\rangle_1 = \sum_k\langle i,k|\hat{\rho}|j,k\rangle$. We can see that here we're essentially computing a "partial trace"—we imagine the matrix for $\hat{\rho}$ as being broken into four 2×2 chunks, computing the trace of each chunk, and taking these as our matrix elements. We can employ similar methods for arbitrarily large systems of particles.

Importantly, if a particle is entangled with another particle then its reduced density operator is a mixed state. So if we look at our three-particle state $|\psi_{123}\rangle$ from the previous section, we could compute the reduced density operator for particle 3:

$$\hat{\rho}^{(3)} = \frac{1}{2}\left(|a|^2 + |b|^2\right)(|+\mathbf{z}\rangle_3)\,({}_3\langle+\mathbf{z}|) + \frac{1}{2}\left(|a|^2 + |b|^2\right)(|-\mathbf{z}\rangle_3)\,({}_3\langle-\mathbf{z}|)$$

$$= \frac{1}{2}(|+\mathbf{z}\rangle_3)\,({}_3\langle+\mathbf{z}|) + \frac{1}{2}(|-\mathbf{z}\rangle_3)\,({}_3\langle-\mathbf{z}|).$$

This matrix describes what would happen if the recipient of a teleported state were to make a measurement before learning how to maneuver their particle into the proper state. The particle is in a completely unpolarized state, so the recipient cannot get any information from such a measurement.

As an important final note, it is possible that seemingly different mixed states can correspond to the same density operator (for example, $(|+\mathbf{z}\rangle\langle+\mathbf{z}| + |-\mathbf{z}\rangle\langle-\mathbf{z}|)/2$ and $(|+\mathbf{x}\rangle\langle+\mathbf{x}| + |-\mathbf{x}\rangle\langle-\mathbf{x}|)/2$). In this case the two mixed states are actually the same quantum state, since all of their expectation values match.

3 Wave Mechanics

3.1 Position States and Translations

Now we'll leave the world of two-state systems to discuss a particle's position and momentum in one dimension. We begin by introducing the position states $|x\rangle$, which satisfy $\hat{x}|x\rangle = x|x\rangle$ for $x \in \mathbb{R}$. Of course, it isn't possible to obtain a single, definite value when we go to measure a particle's positions (a measuring device cannot have infinite precision). Any physical ket is a superposition of position states with coefficients $dx\langle x|\psi\rangle$:

$$|\psi\rangle \int dx |x\rangle\langle x|\psi\rangle.$$

(Integrals are over all space unless otherwise stated.) Notice that if $|\psi\rangle$ is itself a position state $|x'\rangle$ then

$$|x'\rangle = \int dx |x\rangle\langle x|x'\rangle,$$

meaning $\langle x|x'\rangle = \delta(x - x')$. This δ-function appears non-physical in isolation, but in all physical scenarios it appears as part of an integral. For example, in normalizing $|\psi\rangle$ we write

$$1 = \langle\psi|\psi\rangle = \left(\int dx\langle\psi|x\rangle\langle x|\right)\left(\int dx|x\rangle\langle x|\psi\rangle\right)$$
$$= \iint dx' dx \langle\psi|x\rangle\langle x|x'\rangle\langle x'|\psi\rangle,$$

and since the $\langle x|x'\rangle = \delta(x - x') \neq 0$ only for $x = x'$,

$$= \int dx\langle\psi|x\rangle\langle x|\psi\rangle.$$

We therefore identify $dx = |\langle x|\psi\rangle|^2$ with the probability of finding a particle in the interval $[x, x + dx]$. The amplitude density function $\langle x|\psi\rangle$, often denoted $\psi(x)$, is called the wave function; we can use it to rewrite the normalization condition:
$$1 = \int dx\, \psi^*(x)\psi(x) = \int dx\, |\psi(x)|^2.$$

The expected position of a particle in a state $|\psi\rangle$ is given by

$$\langle x\rangle = \langle\psi|\hat{x}|\psi\rangle = \int dx\langle\psi|\hat{x}|x\rangle\langle x|\psi\rangle,$$
$$= \int dx\langle\psi|x\rangle x\langle x|\psi\rangle$$
$$= \int dx\, |\psi(x)|^2 x,$$

where in the second step we wrote $|\psi\rangle$ as a superposition (or, alternatively, inserted the identity operator $1 = \int dx|x\rangle\langle x|$). This matches with our understanding of continuous expectation values from probability!

Now we introduce a translation operator $\hat{T}(a)$ which has the effect $\hat{T}(a)|x\rangle = |x + a\rangle$. For an arbitrary state the action is

$$\hat{T}(a)|\psi\rangle = \hat{T}(a)\int dx'|x'\rangle\langle x'|\psi\rangle \int dx'|x' + a\rangle\langle x'|\psi\rangle,$$

so to shift a state is to shift all the basis kets it's comprised of by an amount a. Note that $\hat{T}(a)$ must be unitary to conserve probability.

Using the above, the amplitude of finding the translated state $|\psi'\rangle$ at a position x is

$$\psi'(x) = \langle x|\psi'\rangle = \int dx' \langle x|x' + a\rangle\langle x'|\psi\rangle$$
$$= \int dx' \, \delta[x - (x' + a)]\langle x'|\psi\rangle$$
$$= \langle x - a|\psi\rangle = \psi(x - a),$$

which is, of course, equivalent to shifting the wave function rightward by a. Alternatively, we could take advantage of the fact that $\hat{T}(a)$ is unitary to write

$$\langle x|\hat{T}(a)|\psi\rangle = \langle x - a|\psi\rangle = \psi(x - a),$$

as before. We might view this through the lens of active and passive transformations—we may literally translate the ket rightward by a, or we can find the amplitude in a space that's been shifted leftward by a.

3.2 The Momentum Operator

Now we define a generator of translations $\hat{p}_x$ using an infinitesimal translation operator:

$$\hat{T}(dx) = 1 - \frac{i}{\hbar}\hat{p}_x dx \implies \hat{T}(a) = e^{-i\hat{p}_x a/\hbar}.$$

Given that $\hat{p}_x$ is Hermitian (since $\hat{T}(dx)$ is unitary) and that it has units of linear momentum, we identify $\hat{p}_x$ as the operator corresponding to the x-component of linear momentum.

It is easy to show, using the definition of $\hat{T}(\delta x)$, that

$$\left(\hat{x}\,\hat{T}(\delta x) - \hat{T}(\delta x)\hat{x}\right)|\psi\rangle = \left(\frac{-i\,\delta x}{\hbar}\right)[\hat{x}, \hat{p}_x]|\psi\rangle.$$

We can rewrite the left-hand side using the superposition expansion for ψ:

$$\left(\hat{x}\,\hat{T}(\delta x) - \hat{T}(\delta x)\hat{x}\right)\int dx|x\rangle\langle x|\psi\rangle$$
$$= \hat{x}\int dx|x + \delta x\rangle\langle x|\psi\rangle - \hat{T}(\delta x)\int dx\, x|x\rangle\langle x|\psi\rangle$$
$$= \int dx(x + \delta x)|x + \delta x\rangle\langle x|\psi\rangle - \int dx\, x|x + \delta x\rangle\langle x|\psi\rangle$$
$$= \delta x\int dx|x + \delta x\rangle\langle x|\psi\rangle = \delta x\int dx|x\rangle\langle x|\psi\rangle = \delta x|\psi\rangle,$$

where in the penultimate step we've kept only the leading order in δx. (Alternatively, we could substitute $x' = x + \delta x$ and Taylor-expand the wave function to leading order.) Comparing this expression with $(-i\,\delta x/\hbar)[\hat{x}, \hat{p}_x]|\psi\rangle$ gives

$$[\hat{x}, \hat{p}_x] = i\hbar.$$

The Heisenberg uncertainty principle $\Delta x \Delta p_x \geq \hbar/2$ follows. With this commutation relation in mind, let's now look at time evolution—if our particle has mass m then the Hamiltonian is

$$\hat{H} - \frac{\hat{p}^2}{2m} + V(\hat{x})$$

for some potential V. So the position expectation value changes like

$$\frac{d\langle x\rangle}{dt} = \frac{i}{2m\hbar}[\hat{p}_x^2, \hat{x}]|\psi\rangle$$
$$= \frac{i}{2m\hbar}\langle\psi|\left(\hat{p}_x[\hat{p}_x, \hat{x}] + [\hat{p}_x, \hat{x}]\hat{p}_x\right)|\psi\rangle$$
$$= \frac{\langle\psi|\hat{p}_x|\psi\rangle}{m} = \frac{\langle p_x\rangle}{m}.$$

This is the familiar expression for classical linear momentum! Further, we could show that

$$\frac{d\langle p_x \rangle}{dt} = \frac{i}{\hbar}\langle \psi | [\hat{H}, \hat{p}_x] | \psi \rangle = \langle -\frac{dV}{dx} \rangle,$$

which is just Newton's second law! Note that the momentum is a constant of motion only when $[\hat{H}, \hat{p}_x]$, that is, when the Hamiltonian is translationally invariant (when $V(x)$ is independent of x).

Now we'll determine the representation of the momentum operator in position space. The action of the infinitesimal translation operator is

$$\hat{T}(\delta x)|\psi\rangle = \int dx |x + \delta x\rangle \langle x|\psi\rangle$$
$$= \int dx' |x'\rangle \langle x' - \delta x|\psi\rangle$$

Taylor-expanding the wave function to first order,

$$= \int dx' |x'\rangle \left(\langle x'|\psi\rangle - \delta x \frac{\partial}{\partial x'}\langle x'|\psi\rangle \right)$$
$$= |\psi\rangle - \delta x \int dx' |x'\rangle \frac{\partial}{\partial x'}\langle x'|\psi\rangle.$$

Comparing with the definition of $\hat{T}(\delta x)$ gives

$$\hat{p}_x|\psi\rangle = \frac{\hbar}{i} \int dx' |x'\rangle \frac{\partial}{\partial x'}\langle x'|\psi\rangle,$$

and so

$$\langle x|\hat{p}_x|\psi\rangle = \frac{\hbar}{i} \frac{\partial}{\partial x}\langle x|\psi\rangle.$$

For an arbitrary state the expectation value is

$$\langle \psi|\hat{p}_x|\psi\rangle = \int dx' \langle \psi|x'\rangle \frac{\hbar}{i} \frac{\partial}{\partial x'}\langle x'|\psi\rangle.$$

All of this suggests that

$$\hat{p}_x \xrightarrow[x \text{ basis}]{} \frac{\hbar}{i} \frac{\partial}{\partial x}.$$

3.3   Momentum Space

We now introduce a new set of momentum eigenstates satisfying $\hat{p}_x|p\rangle = p|p\rangle$. Arbitrary states look like

$$|\psi\rangle = \int dp |p\rangle \langle p|\psi\rangle,$$

and of course $\langle p'|p\rangle = \delta(p' - p)$. The normalization condition suggests that we identify $dp \, |\langle p|\psi\rangle|^2$ as the probability that a particle in $|\psi\rangle$ has a momentum in $[p, p + dp]$; the amplitude density $\langle p|\psi\rangle$ is called the wave function in momentum space.

It will be useful to determine the momentum eigenstates' position-space wave functions $\langle x|p\rangle$. First we write

$$\langle x|\hat{p}_x|p\rangle = p\langle x|p\rangle = \frac{\hbar}{i} \frac{\partial}{\partial x}\langle x|p\rangle,$$

a differential equation which has solution $\langle x|p\rangle = Ne^{ipx/\hbar}$. We could evaluate (using $\delta(x) = \frac{1}{2\pi}\int dk \, e^{ikx}$)

$$\langle p'|p\rangle = \int dx \langle p'|x\rangle \langle x|p\rangle$$

to find that $N = 1/\sqrt{2\pi\hbar}$, and so

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}.$$

Inserting the identity operator here allows us to transform back and forth between position and momentum space:

$$\langle p|\psi\rangle = \int dx\langle p|x\rangle\langle x|\psi\rangle = \int dx \frac{1}{\sqrt{2\pi\hbar}}e^{-ipx/\hbar}\langle x|\psi\rangle$$

$$\langle x|\psi\rangle = \int dp\langle x|p\rangle\langle p|\psi\rangle = \int dp \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}\langle p|\psi\rangle$$

These together show that $\langle p|\psi\rangle$ and $\langle x|\psi\rangle$ form a Fourier transform pair.

As an application, consider the Gaussian wave packet

$$\langle x|\psi\rangle = \psi(x) = Ne^{-x^2/2a^2}, \quad N = \frac{1}{\sqrt{a\sqrt{\pi}}}.$$

By symmetry $\langle x\rangle = 0$, and by integration $\langle x^2\rangle = a^2/2$ and so $\Delta x = a/\sqrt{2}$, which reflects the fact that the Gaussian narrows as $a \to 0$. We could show, using the above integrals, that

$$\langle p|\psi\rangle = \sqrt{\frac{a}{\hbar\sqrt{\pi}}}e^{-p^2a^2/2\hbar^2}.$$

Once again, by symmetry $\langle p_x\rangle = 0$, by integration $\langle p_x^2\rangle = \hbar^2/2a^2$, and $\Delta p_x = \hbar/a\sqrt{2}$. Note that $\Delta x \Delta p_x = \hbar/2$, so this Gaussian wave packet is a minimum uncertainty state. (We could prove that it is *the* minimum uncertainty state by requiring $|\beta\rangle = c|\alpha\rangle$ and $\langle\psi|(\hat{O} + \hat{O}^\dagger)|\psi\rangle = 0$ in the proof of the uncertainty principle.)

One advantage of expressing states in momentum space is that free particles have $\hat{H} = \hat{p}_x^2/2m$, meaning momentum states are also energy eigenstates! Thus

$$|\psi(t)\rangle = e^{i\hat{H}t/\hbar}\int dp|p\rangle\langle p|\psi(0)\rangle = \int dp\, e^{-ip^2t/2m\hbar}|p\rangle\langle p|\psi(0)\rangle.$$

We put this into position space by prepending $\langle x|$ which, after integrating, gives

$$\psi(x,t) = \frac{1}{\sqrt{[a + (i\hbar t/ma)]\sqrt{\pi}}}\exp\left(-\frac{x^2}{2a^2[1 + (i\hbar t/ma^2)]}\right).$$

If we compare $\psi^*(x,t)\psi(x,t)$ with the time-independent form from before, we see that

$$\Delta x = \frac{a}{\sqrt{2}}\left(1 + \frac{\hbar^2 t^2}{m^2 a^4}\right)^{1/2}.$$

3.4 Energy in a Harmonic Oscillator

Given how ubiquitous harmonic oscillators are in physics, in this section we will "solve" these oscillators—that is, we will determine the eigenvalues and eigenkets of the Hamiltonian

$$\hat{H} = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$$

given the commutation relation $[\hat{x}, \hat{p}_x] = i\hbar$. The operator methods employed here will be very similar in spirit to the ones we used to solve the angular momentum problem.

We start by defining two non-Hermitian operators that are reminiscent of our $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$, noting that we can't simply add $\hat{x} \pm i\hat{p}_x$ because of the unit discrepancy:

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i}{m\omega}\hat{p}_x\right), \qquad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i}{m\omega}\hat{p}_x\right).$$

The factor out front is just there to make the operators dimensionless; this produces $[\hat{a}, \hat{a}^\dagger] = 1$. We also get

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^\dagger), \qquad \hat{p}_x = -i\sqrt{\frac{m\omega\hbar}{2}}(\hat{a} - \hat{a}^\dagger),$$

and so the Hamiltonian can be expressed as

$$\hat{H} = \frac{\hbar\omega}{2}(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger) = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right).$$

Thus finding the eigenstates of $\hat{H}$ is equivalent to finding the eigenstates of the operator $\hat{N} = \hat{a}^\dagger\hat{a}$. We'll do this now! Denote the eigenstates by $|\eta\rangle$. Note that the eigenvalues are non-negative—we have the expectation value

$$\langle\eta|\hat{N}|\eta\rangle = (|\eta\rangle\hat{a}^\dagger)(\hat{a}|\eta\rangle) = \eta\langle\eta|\eta\rangle,$$

and since both of these inner products are zero we must have $\eta \geq 0$. To see the effects of $\hat{a}$ and $\hat{a}^\dagger$ on the $|\eta\rangle$, we first determine the commutation relations

$$
\begin{aligned}
[\hat{N}, \hat{a}] &= [\hat{a}^\dagger\hat{a}, \hat{a}] & \qquad [\hat{N}, \hat{a}^\dagger] &= [\hat{a}^\dagger\hat{a}, \hat{a}^\dagger] \\
&= [\hat{a}^\dagger, \hat{a}]\hat{a} & &= \hat{a}^\dagger[\hat{a}, \hat{a}^\dagger] \\
&= -\hat{a}, & &= \hat{a}^\dagger,
\end{aligned}
$$

and evaluate $\hat{N}\hat{a}^\dagger|\eta\rangle$ and $\hat{N}\hat{a}|\eta\rangle$:

$$
\begin{aligned}
\hat{N}\hat{a}^\dagger|\eta\rangle &= (\hat{a}^\dagger\hat{N} + \hat{a}^\dagger)|\eta\rangle & \qquad \hat{N}\hat{a}|\eta\rangle &= (\hat{a}\hat{N} - \hat{a})|\eta\rangle \\
&= (\hat{a}^\dagger\eta + \hat{a}^\dagger)|\eta\rangle & &= (\hat{a}\eta - \hat{a})|\eta\rangle, \\
&= (\eta + 1)\hat{a}^\dagger|\eta\rangle & &= (\eta - 1)\hat{a}|\eta\rangle,
\end{aligned}
$$

where we've used the commutation relations above to swap the order of the operations. We can see, then, that $\hat{a}^\dagger|\eta\rangle$ is an eigenket of $\hat{N}$ with eigenvalue $\eta + 1$, and that $\hat{a}|\eta\rangle$ has eigenvalue $\eta - 1$, meaning these are raising and lowering operators.

The fact that $\eta \geq 0$ puts a restriction on how far we can lower our eigenvalues. The minimum-eigenvalue ket must satisfy $\hat{a}|\eta_{\min}\rangle = 0$, and if we apply the raising operator to this we get

$$\hat{a}^\dagger\hat{a}|\eta_{\min}\rangle = \eta_{\min}|\eta_{\min}\rangle = 0,$$

we can see that we must have $\eta_{\min} = 0$. So we label the lowest state as $|0\rangle$, and applying the $\hat{a}^\dagger$ n times yields the state satisfying

$$\hat{N}|n\rangle = n|n\rangle, \qquad n = 0, 1, 2, \ldots.$$

Thus the eigenvalues of $\hat{N}$—now called the number operator—are the non-negative integers, and those of the Hamiltonian are

$$\hat{H}|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle, \qquad n = 0, 1, 2, \ldots.$$

So we've quantized the energy of the harmonic oscillator! The spacing is uniform with step size $\hbar\omega$.

Like before, it will be useful to determine the matrix representations of our raising and lowering operators. We'd like to solve for the c_+, c_- satisfying $\hat{a}^\dagger|n\rangle = c_+|n+1\rangle$ and $\hat{a}|n\rangle = c_-|n-1\rangle$. If we take the inner product of the two equations

$$\hat{a}^\dagger|n\rangle = c_+|n+1\rangle, \qquad \langle n|\hat{a} = c_+^*\langle n+1|$$

we get $(n+1)\langle n|n\rangle = c_+^* c_+\langle n+1|n+1\rangle$. (Note that on the left we get an $\hat{a}\hat{a}^\dagger$, and so we use the commutation relation to swap the order of the operators.) If the eigenstates are normalized then we must have $c_+ = \sqrt{n+1}$. Similarly, $c_- = \sqrt{n}$, meaning

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \qquad \hat{a}|n\rangle = \sqrt{n}|n-1\rangle,$$

and the matrix elements of these operators are

$$\langle n'|\hat{a}^\dagger|n\rangle = \sqrt{n+1}\,\delta_{n',n+1}, \qquad \langle n'|\hat{a}|n\rangle = \sqrt{n}\,\delta_{n',n-1}.$$

The matrix representations in the energy eigenbasis, therefore, look like

$$\hat{a}^\dagger \to \begin{bmatrix} 0 & 0 & \cdots \\ \sqrt{1} & 0 & \cdots \\ 0 & \sqrt{2} & \\ \vdots & \vdots & \ddots \end{bmatrix}, \qquad \hat{a} \to \begin{bmatrix} 0 & \sqrt{1} & 0 & \cdots \\ 0 & 0 & \sqrt{2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

We could construct the matrix representations of the position and momentum operators from here (again, in the energy eigenbasis). At this point we can also establish. via induction, that a normalized ket $|n\rangle$ can be expressed as

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle.$$

3.5 Analysis of the Energy Eigenvalues and Eigenfunctions

Now let's connect this discussion back to wave mechanics by determining the position-space eigenfunctions for the harmonic oscillator. Projecting the ground-state $\hat{a}|0\rangle = 0$ into position space gives

$$0 = \langle x|\hat{a}|0\rangle = \sqrt{\frac{m\omega}{2\hbar}}\langle x|\left(\hat{x} + \frac{i}{m\omega}\hat{p}_x\right)|0\rangle = \sqrt{\frac{m\omega}{2\hbar}}\left(x\langle x|0\rangle + \frac{i}{m\omega}\cdot\frac{\hbar}{i}\frac{\partial\langle x|0\rangle}{\partial x}\right),$$

a differential equation which has the (normalized) solution

$$\langle x|0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}.$$

We can use this to generate all of the position-space energy eigenfunctions:

$$\langle x|n\rangle = \frac{1}{\sqrt{n!}}\langle x|(a^\dagger)^n|0\rangle = \frac{1}{\sqrt{n!}}\left(\frac{m\omega}{2\hbar}\right)^{n/2}\left(x - \frac{\hbar}{m\omega}\frac{d}{dx}\right)^n\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x/2\hbar}.$$

Note that both of the energies' expectation values

$$\frac{\langle p_x^2\rangle}{2m} = -\frac{\hbar^2}{2m}\int dx\langle n|x\rangle\frac{d^2}{dx^2}\langle x|n\rangle,$$

$$\langle V(x)\rangle = \frac{1}{2}m\omega^2\int dx\langle n|x\rangle x^2\langle x|n\rangle.$$

get larger with increasing n due to the wave function's increasing curvature (for the kinetic energy) and widening domain of applicability (for the potential).

Let's hone in on the ground-state energy. In a classical oscillator the lowest energy occurs when the particle is at rest, but for us this is disallowed by the Heisenberg uncertainty principle. These uncertainties influence the energy via

$$\langle E\rangle = \frac{(\Delta p_x)^2 + \langle p_x\rangle^2}{2m} + \frac{1}{2}m\omega^2[(\Delta x)^2 + \langle x\rangle^2].$$

It's easy to show, using the raising and lowering operators, that $\langle x\rangle = \langle p_x\rangle = 0$ in an energy eigenstate, and so in such a state

$$\langle E\rangle = \frac{(\Delta p_x)^2}{2m} + \frac{1}{2}m\omega^2(\Delta x)^2.$$

To minimize the energy nature picks the combination of $(\Delta x)^2 = \langle x^2\rangle$ and $(\Delta p_x)^2 = \langle p_x^2\rangle$ that minimizes this expression. In the ground state these quantities are

$$(\Delta x)^2 = \frac{\hbar}{2m\omega}\langle 0|(\hat{a} + \hat{a}^\dagger)^2|0\rangle \qquad\qquad (\Delta p_x)^2 = -\frac{m\omega\hbar}{2}\langle 0|(\hat{a} - \hat{a}^\dagger)^2|0\rangle$$

$$= \frac{\hbar}{2m\omega}\langle 0|(\hat{a}^2 + (\hat{a}^\dagger)^2 + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})|0\rangle \qquad = -\frac{m\omega\hbar}{2}\langle 0|(\hat{a}^2 + (\hat{a}^\dagger)^2 - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a})|0\rangle$$

$$= \frac{\hbar}{2m\omega}\langle 0|\hat{a}\hat{a}^\dagger|0\rangle = \frac{\hbar}{2m\omega}, \qquad\qquad = \frac{m\omega\hbar}{2}\langle 0|\hat{a}\hat{a}^\dagger|0\rangle = \frac{m\omega\hbar}{2},$$

which notably produces $\Delta x\Delta p_x = \hbar/2$. So the ground state, as we found before—our $\langle x|0\rangle$ is a Gaussian, after all—is a minimum-uncertainty state! For the excited states we similarly have

$$\Delta x = \sqrt{\left(n + \frac{1}{2}\right)\frac{\hbar}{m\omega}}, \qquad \Delta p_x = \sqrt{\left(n + \frac{1}{2}\right)m\omega\hbar},$$

which produce $\Delta x\Delta p_x = (n + 1/2)\hbar$.

Now let's zoom back out. The correspondence principle states that we should expect that the predictions of quantum mechanics to agree with those of classical physics on macroscopic scales—that is, scales on which our quantum number n is very large. On such a scale, the classical motion of a particle with energy E is restricted to lie within the classical turning points at which the kinetic energy is zero:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega = \frac{1}{2}m\omega^2 x_n^2.$$

This agrees with the (visual) observation that the eigenfunctions' excursions beyond these turning points becomes less pronounced as n increases.

For a classical oscillator, the probability of finding an energy-E particle in $[x, x + dx]$ is proportional to the amount of time dx/v it spends in this interval. We can express this probability as

$$P_{\text{cl}}\, dx \propto \frac{dx}{(2/m)(E_n - V(x))} = \frac{dx}{\sqrt{2E_n/m - \omega^2 x^2}} = \frac{dx}{\omega\sqrt{x_n^2 - x^2}},$$

and normalizing gives a factor of $1/\sqrt{\pi}$ out front. Plotting this over a a $|\langle x|n\rangle|^2$ (for reasonably large n) reveals that the P_{cl} we've found approximates the mean value of $|\langle x|n\rangle|^2$ at any given position!

3.6 Time Dependence and Coherent States

The energy eigenstates of the harmonic oscillator are stationary states. So time dependence, as we've come to expect, results from superpositions of eigenstates with different energies. If this superposition is of two adjacent energy states $|\psi(0)\rangle = c_n|n\rangle + c_{n+1}|n+1\rangle$ then the time evolution is

$$|\psi(t)\rangle = e^{-i(n+1/2)\omega t}\left(c_n|n\rangle + c_{n+1}e^{-i\omega t}|n+1\rangle\right),$$

and we could use the raising and lowering operators to compute $\langle x\rangle = A\cos(\omega t + \delta)$.

Another example of a state with interesting time evolution is the coherent state $|\alpha\rangle$, which satisfies $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$. If we start by writing such a state as a superposition of the energy eigenstates with coefficients c_n, the eigenvalue equation becomes

$$\alpha \sum_{n=0}^{\infty} c_n|n\rangle = \sum_{n=1}^{\infty} \sqrt{n}\, c_n|n-1\rangle$$
$$= \sum_{n'=0}^{\infty} \sqrt{n'+1}\, c_{n'+1}|n'\rangle.$$

So all of the coefficients must satisfy the recurrence $\sqrt{n+1}\, c_{n+1} = \alpha c_n$, the solution of which is $c_n = (\alpha^n/\sqrt{n!})c_0$ and thus

$$|\alpha\rangle = c_0 \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle.$$

To normalize we could evaluate $\langle\alpha|\alpha\rangle = |c_0|^2 e^{|\alpha|^2}$, meaning $c_0 = e^{-|\alpha|^2/2}$ up to an overall phase and so

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle.$$

The time evolution is given by

$$|\alpha(t)\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n e^{-i(n+1/2)\omega t}}{\sqrt{n!}}|n\rangle$$
$$= e^{i\omega t/2} e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{(e^{-i\omega t})^n}{\sqrt{n!}}|n\rangle$$
$$= e^{-i\omega t/2}|\alpha e^{-i\omega t}\rangle,$$

so the eigenvalue α of the lowering operator becomes $\alpha e^{i\omega t}$ as time progresses.

As a first step in investigating the behavior of this coherent state, let's compute the expectation values:

$$\langle x \rangle = \sqrt{\frac{\hbar}{2m\omega}} \langle \alpha(t)|(\hat{a}+\hat{a}^\dagger)|\alpha(t)\rangle \qquad \langle p_x \rangle = -i\sqrt{\frac{m\omega\hbar}{2}} \langle \alpha(t)|(\hat{a}-\hat{a}^\dagger)|\alpha(t)\rangle$$

$$= \sqrt{\frac{\hbar}{2m\omega}} \left(\alpha e^{-i\omega t} + \alpha^* e^{i\omega t}\right) \qquad\qquad = -i\sqrt{\frac{m\omega\hbar}{2}} \left(\alpha e^{-i\omega t} - \alpha^* e^{i\omega t}\right)$$

$$= \sqrt{\frac{\hbar}{2m\omega}} \, 2|\alpha|\cos(\omega t + \delta), \qquad\qquad = -\sqrt{\frac{m\omega\hbar}{2}} \, 2|\alpha|\sin(\omega t + \delta),$$

where $\alpha = |\alpha|e^{-i\delta}$. We can see that the position and momentum are oscillating back and forth, as we'd expect from motion in a harmonic oscillator. Note that the phase δ of the eigenvalue determines the phase of these expectation values. Now we'll compute the uncertainties, using the commutator $[\hat{a}, \hat{a}^\dagger] = 1$ to swap the order of operators as needed:

$$\langle x^2 \rangle = \frac{\hbar}{2m\omega} \langle \alpha(t)|(\hat{a}+\hat{a}^\dagger)^2|\alpha(t)\rangle \qquad \langle p_x^2 \rangle = -\frac{m\omega\hbar}{2} \langle \alpha(t)|(\hat{a}-\hat{a}^\dagger)^2|\alpha(t)\rangle$$

$$= \frac{\hbar}{2m\omega} \langle \alpha(t)|(\hat{a}+(\hat{a}^\dagger)^2+2\hat{a}^\dagger\hat{a}+1)|\alpha(t)\rangle \qquad = -\frac{m\omega\hbar}{2} \langle \alpha(t)|(\hat{a}^2+(\hat{a}^\dagger)^2-2\hat{a}^\dagger\hat{a}-1)|\alpha(t)\rangle$$

$$= \frac{\hbar}{2m\omega} \left(\alpha(t) + \alpha^*(t)^2 + 2|\alpha(t)|^2 + 1\right), \qquad = \frac{m\omega\hbar}{2} \left(2|\alpha(t)|^2 + 1 - \alpha(t)^2 - \alpha^*(t)^2\right),$$

which combined with the previous expectation values gives

$$(\Delta x)^2 = \frac{\hbar}{2m\omega}, \qquad (\Delta p_x)^2 = \frac{m\omega\hbar}{2}.$$

The product of the uncertainties is $\Delta x \Delta p_x = \hbar/2$. Thus the coherent state is a minimum uncertainty state, and unlike the Gaussian wave packet for the free particle, this coherent Gaussian does not spread with time. This is just about as close as we can get to a quantum state that might represent, say, a classical pendulum!

3.7   The Schrödinger Equation and Inversion Symmetry

As an alternative way of solving the harmonic oscillator, we can write down the energy eigenvalue equation as a differential equation:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\langle x|E\rangle + \frac{1}{2}m\omega^2 x^2 \langle x|E\rangle = E\langle x|E\rangle.$$

If we substitute $y = x\sqrt{m\omega/\hbar}$ we get

$$\frac{d^2\psi}{dy^2} + (\varepsilon - y^2)\psi = 0, \qquad \varepsilon = \frac{2E}{\hbar\omega}.$$

As a first step toward finding physical solutions, we note that the solutions in the $|y| \to \infty$ limit look like $\psi(y) = Ae^{-y^2/2} + Be^{y^2/2}$. So to "factor out" this limiting behavior we can express the wave function as

$$\psi(y) = h(y)e^{-y^2/2},$$

where we have dropped the non-normalizable term. Substituting this into the Schrödinger equation reveals that h must satisfy

$$\frac{d^2h}{dy^2} - 2y\frac{dh}{dy} + (\varepsilon - 1)h = 0,$$

which has the power series solution

$$h(y) = \sum_{k=0}^{\infty} a_k y^k, \qquad \frac{a_{k+2}}{a_k} = \frac{2k+1-\varepsilon}{(k+2)(k+1)}.$$

But there's a problem—this solution exhibits the same behavior as e^{y^2} in the large-k limit, so it diverges. To reconcile this we note that $h(y)$ is finite so long as $\varepsilon = 2n+1$ for some integer n, so $\psi(y)$ becomes a finite polynomial multiplied by a decreasing exponential. Since $\varepsilon = 2E/\hbar\omega$ the corresponding energies are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \qquad n = 0, 1, 2, \ldots.$$

Each $h(y)$ is called the Hermite polynomial of order n, and we may express it either in power series form or using our previous eigenfunctions for the harmonic oscillator.

One of the most striking features of the energy eigenfunctions is that they're all either even or odd. This behavior is rooted in a symmetry in the Hamiltonian. To formalize this symmetry we introduce the parity operator $\hat{\Pi}$, which acts on position states via

$$\hat{\Pi}|x\rangle = |-x\rangle.$$

An eigenstate of this operator satisfies $\hat{\Pi}|\psi_A\rangle = \lambda|\psi_A\rangle$, and since $\hat{\Pi}^2$ is the identity operator, the eigenvalues are $\lambda = \pm 1$. For the corresponding eigenstates note that for an arbitrary state $|\psi\rangle$ we have $\langle x|\hat{\Pi}^2|\psi\rangle = \langle -x|\psi\rangle = \psi(-x)$, so the parity eigenfunctions satisfy

$$\psi_\lambda(-x) = \lambda\psi_\lambda(x).$$

Thus $\lambda = 1$ corresponds to the even eigenfunctions and $\lambda = -1$ the odd ones.

Importantly, the harmonic oscillator Hamiltonian and the parity operator commute:

$$
\begin{aligned}
\langle x|\hat{\Pi}\hat{H}|\psi\rangle &= \langle -x|\hat{H}|\psi\rangle \\
&= \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(-x)\right)\psi(-x) \\
&= \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi(-x) \\
&= \langle x|\hat{H}\hat{\Pi}|\psi\rangle.
\end{aligned}
$$

In fact, this relation holds for every even choice V! By observing the symmetry of the Hamiltonian under this inversion (among others), we can deduce some of the properties of the eigenfunctions before we've even solved the eigenvalue equation. This fact will come in handy in the next chapter.

4 Central Potentials

4.1 Wave Mechanics in Three Dimensions

Now w'll extend our discussion of wave mechanics to three dimensions. The position eigenstates are specified by three coordinates $|\mathbf{r}\rangle = |x, y, z\rangle$, each dimension with its own position operator:

$$\hat{x}|\mathbf{r}\rangle = x|\mathbf{r}\rangle, \quad \hat{y}|\mathbf{r}\rangle = y|\mathbf{r}\rangle, \quad \hat{z}|\mathbf{r}\rangle = z|\mathbf{r}\rangle.$$

An arbitrary state is a superposition of position states,

$$|\psi\rangle = \int d^3\mathbf{r} |\mathbf{r}\rangle\langle\mathbf{r}|\psi\rangle,$$

and like before $\langle\mathbf{r}|\mathbf{r}'\rangle = \delta^3(\mathbf{r} - \mathbf{r}')$, where δ^3 is the product of three deltas (one for each dimension). The normalization condition $\langle\psi|\psi\rangle = 1$ suggests that we identify $d^3\mathbf{r} |\langle\mathbf{r}|\psi\rangle|^2$ as the probability of finding the particle in some $d^3\mathbf{r}$ of space.

Now we introduce a three-dimension translation operator $\hat{T}(\mathbf{a})|\mathbf{r}\rangle = |\mathbf{r} + \mathbf{a}\rangle$ which is generated by three operators:

$$\hat{T}(a_x\mathbf{i}) = e^{-i\hat{p}_x a_x/\hbar}, \qquad \hat{T}(a_y\mathbf{j}) = e^{-i\hat{p}_y a_y/\hbar}, \qquad \hat{T}(a_z\mathbf{k}) = e^{-i\hat{p}_z a_z/\hbar}.$$

In contrast with rotations, translations in different directions certainly commute with one another. Their generators do, too—we could Taylor-expand the elements of $[\hat{T}(a_x\mathbf{i}), \hat{T}(a_y\mathbf{j})] = 0$ to second order, for example, to find that $[\hat{p}_x, \hat{p}_y] = 0$. We can therefore express the translation operator as

$$\hat{T}(\mathbf{a}) = e^{-i\hat{p}_x a_x/\hbar} e^{-i\hat{p}_y a_y/\hbar} e^{-i\hat{p}_z a_z/\hbar} = e^{-i\mathbf{p}\cdot\mathbf{a}/\hbar}.$$

In each dimension we have the same commutation relation $[\hat{x}, \hat{p}_x] = i\hbar$ as before, but across dimensions the position and momentum operators do commute:

$$\hat{T}(a_x\mathbf{i})\,\hat{y}|x, y, z\rangle = y|x + a_x, y, z\rangle = \hat{y}\,\hat{T}(a_x\mathbf{i})|x, y, z\rangle.$$

As shorthand for this, we write

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}.$$

Now, because the momentum operators all commute with one another, we can form three-dimensional momentum states $|\mathbf{p}\rangle = |p_x, p_y, p_z\rangle$ normalized by $\langle\mathbf{p}'|\mathbf{p}\rangle = \delta^3(\mathbf{p}' - \mathbf{p})$ with $d^3\mathbf{p} |\langle\mathbf{p}|\psi\rangle|^2$ as the probability of finding a particle in some $d^3\mathbf{p}$ of momentum space.

The generalization of the momentum operator in the position basis is

$$\langle\mathbf{r}|\hat{\mathbf{p}}|\psi\rangle = \frac{\hbar}{i}\nabla\langle\mathbf{r}|\psi\rangle.$$

If we take $\psi = \mathbf{p}$ then this becomes a differential equation with solution

$$\langle\mathbf{r}|\mathbf{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}}e^{i\mathbf{p}\cdot r/\hbar},$$

just the product of three of the one-dimensional eigenfunctions we found before.

4.2 Translational Symmetry and Linear Momentum

We'll take a special interest in systems with central potentials, which have Hamiltonians of the form

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m_1} + \frac{\hat{\mathbf{p}}^2}{2m_2} + V(|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|).$$

The most important example of such a two-body system is the one where V is the Coulomb potential, but for now we'll keep our analysis general. We introduce the two-body position basis states $|\mathbf{r}_1, \mathbf{r}_2\rangle = |\mathbf{r}_1\rangle_1 \otimes |\mathbf{r}_2\rangle_2$. The translation operators $\hat{T}_1(\mathbf{a}), \hat{T}_2(\mathbf{a})$ commute since we can translate the particles independent of one another; it follows that $[\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2] = 0$ and that the "joint" translation operator is

$$\hat{T}_1(\mathbf{a})\hat{T}_2(\mathbf{a}) = e^{-i\hat{\mathbf{p}}_1 \cdot \mathbf{a}/\hbar}e^{-i\hat{\mathbf{p}}_2 \cdot \mathbf{a}/\hbar} = e^{-i\hat{\mathbf{P}} \cdot \mathbf{a}/\hbar}, \qquad \hat{\mathbf{P}} = \hat{\mathbf{p}}_1 + \hat{\mathbf{p}}_2.$$

Applying this operator to a pair of particles doesn't affect the distance between them, so we expect that it commutes with the Hamiltonian. We can immediately see that $\hat{T}_1(\mathbf{a})\hat{T}_2(\mathbf{a})$ commutes with the kinetic energy bit; as for the potential term, the process is a bit mechanical:

$$\begin{aligned}
\hat{T}_1(\mathbf{a})\hat{T}_2(\mathbf{a})V(|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|)|\mathbf{r}_1, \mathbf{r}_2\rangle &= \hat{T}_1(\mathbf{a})\hat{T}_2(\mathbf{a})V(|\mathbf{r}_1 - \mathbf{r}_2|)|\mathbf{r}_1.\mathbf{r}_2\rangle \\
&= V(|\mathbf{r}_1 - \mathbf{r}_2|)|\mathbf{r}_1 + \mathbf{a}.\mathbf{r}_2 + \mathbf{a}\rangle \\
&= V(|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|)|\mathbf{r}_1 + \mathbf{a}.\mathbf{r}_2 + \mathbf{a}\rangle \\
&= V(|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|)\hat{T}_1(\mathbf{a})\hat{T}_2(\mathbf{a})|\mathbf{r}_1, \mathbf{r}_2\rangle.
\end{aligned}$$

Thus $[\hat{H}, \hat{T}_1(\mathbf{a})\hat{T}_2(\mathbf{a})] = 0$, as expected. It follows that time-evolving the translated state $\hat{T}(\mathbf{a})|\psi(0)\rangle$ is equivalent to translating the time-evolved state $|\psi(t)\rangle$. This is precisely what we'd expect from (Galilean) relativity. Also, because translations are generated by $\hat{\mathbf{P}}$ we can write

$$[\hat{H}, \hat{\mathbf{P}}] = 0,$$

meaning $d\langle\mathbf{P}\rangle/dt = 0$—translational invariance implies conservation of momentum! This is another example (following parity) of the deep connection between symmetries of the Hamiltonian and conservation laws. Notice that we can turn the argument around, too: conservation of momentum implies $[\hat{H}, \hat{\mathbf{P}}] = 0$ and so $\hat{H}$ exhibits translational symmetry. Such an argument holds whenever there are no interactions external to the system we're working with.

We can use this discussion to simplify our Hamiltonian considerably. The natural coordinates for our problem aren't the positions standard Cartesian coordinates, but rather the center-of-mass position $\mathbf{R}$ and the relative position $\mathbf{r}$ defined by

$$\hat{\mathbf{r}} = \hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2, \qquad \hat{\mathbf{R}} = \frac{m_1\hat{\mathbf{r}}_1 + m_2\hat{\mathbf{r}}_2}{m_1 + m_2}.$$

It's easy to see that the relative position and total momentum operators satisfy $[\hat{x}_i, \hat{P}_j] = 0$. We could also show that $[\hat{X}_i, \hat{P}_j] = i\hbar\delta_{ij}$, as we'd expect from the usual canonical commutation relations. If we introduce the relative momentum

$$\hat{\mathbf{p}} = \frac{m_2\hat{\mathbf{p}}_1 - m_1\hat{\mathbf{p}}_2}{m_1 + m_2},$$

we could show that it satisfies $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$ and $[\hat{X}_i, \hat{p}_j] = 0$ in a symmetric fashion. Thus relative operators obey the usual commutation relations when paired with other relative operators (as do the center-of-mass operators), but operators from different groups commute.

In these coordinates we could show that the Hamiltonian is

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2M} + \frac{\hat{\mathbf{p}}^2}{2\mu} + V(|\hat{\mathbf{r}}|), \qquad M = m_1 + m_2, \quad \mu = \frac{m_1 m_2}{m_1 + m_2}.$$

We view this as the sum of the center-of-mass kinetic energy $\hat{H}_{\mathrm{cm}} = \hat{\mathbf{P}}^2/2M$ and the energy due to the relative motion $\hat{H}_{\mathrm{rel}} = \hat{\mathbf{p}}^2/2\mu + V(|\hat{\mathbf{r}}|)$. These operators commute, so they have simultaneous eigenstates $|E_{\mathrm{cm}}, E_{\mathrm{rel}}\rangle$; in particular, the eigenstates of $\hat{H}_{\mathrm{cm}}$ are those of $\hat{\mathbf{P}}$, so in momentum space

$$\langle\mathbf{R}|\mathbf{P}\rangle = \frac{1}{(2\pi\hbar)^{3/2}}e^{i\mathbf{P}\cdot\mathbf{R}/\hbar},$$

as we saw before. To simplify things we'll usually analyze the two-body problem in the center-of-mass frame, where $\mathbf{P} = 0$:

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2\mu} + V(|\hat{\mathbf{r}}|).$$

We've therefore reduced the problem to a single body in a central potential $V(r)$, as seen in classical mechanics.

4.3 Rotational Symmetry and Angular Momentum

Another observation we can make about this $\hat{H}$ is that it is invariant under rotations; we'll once again verify this using operators. We first introduce

$$\hat{R}(d\phi\,\mathbf{k}) = \left(1 - \frac{i}{\hbar}\hat{L}_z\,d\phi\right),$$

which rotates a position state through an angle $d\phi$ about the z-axis. Its effect can be seen by

$$\hat{R}(d\phi\,\mathbf{k})|x, y, z\rangle = |x - y\,d\phi,\ y + x\,d\phi,\ z\rangle$$
$$= \left[1 - \frac{i}{\hbar}\hat{p}_x(-y\,d\phi)\right]\left[1 - \frac{i}{\hbar}\hat{p}_y(x\,d\phi)\right]|x, y, z\rangle$$
$$= \left[1 - \frac{i}{\hbar}(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x)d\phi\right]|x, y, z\rangle,$$

and so we identify

$$\hat{L}_z = \hat{x}\hat{p}_x - \hat{y}\hat{p}_x$$

as the z-component of the orbital angular momentum operator $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$. (Orbital angular momentum appears here because we're now generating rotations of position states.) Using this form of $\hat{L}_z$ it's easy to directly verify the intuitive relations

$$[\hat{L}_z, \hat{p}_x] = i\hbar\hat{p}_y, \qquad [\hat{L}_z, \hat{p}_y] = -i\hbar\hat{p}_x, \qquad [\hat{L}_z, \hat{p}_z] = 0,$$
$$[\hat{L}_z\hat{x}] = i\hbar\hat{y}, \qquad [\hat{L}_z, \hat{y}] = -i\hbar\hat{x}, \qquad [\hat{L}_z, \hat{z}] = 0,$$

meaning $[\hat{L}_z, \hat{\mathbf{p}}^2] = 0$ and $[\hat{L}_z, \hat{\mathbf{r}}^2] = 0$. Thus $\hat{L}_z$ commutes with our Hamiltonian. Alternatively, we may have used known eigenvalue equations:

$$\hat{R}(d\phi\,\mathbf{k})V(|\hat{\mathbf{r}}|)|r, \theta, \phi\rangle = V(r)|r, \theta,\ \phi + d\phi\rangle$$
$$= V(|\hat{\mathbf{r}}|)\hat{R}(d\phi\,\mathbf{k})|r, \theta, \phi\rangle.$$

Since $\hat{H}$ commutes with $\hat{L}_z$ and our axis of the z-axis was arbitrary, our Hamiltonian must commute with the generator of rotations about any axis! Thus the system is invariant under rotations; it follows that angular momentum is conserved.

We'll once again leverage this discussion in our analysis of $\hat{H}$, this time to specify the energy eigenstates and an eigenvalue equation that these states must satisfy. Just like we saw with spin, we can form simultaneous eigenstates of $\hat{\mathbf{L}}^2$ and $\hat{L}_z$, and the Hamiltonian commutes with both of these. So we can write the following.

$$\hat{H}|E, l, m\rangle = E|E, l, m\rangle,$$
$$\hat{\mathbf{L}}^2|E, l, m\rangle = l(l+1)\hbar^2|E, l, m\rangle,$$
$$\hat{L}_z|E, l, m\rangle = m\hbar|E, l, m\rangle.$$

We wish to write down the energy eigenvalue equation in position space. The only nontrivial bit here is the kinetic energy term, so really we seek the position-space representation of $\hat{\mathbf{p}}^2$. To this end, we note the identity

$$(\hat{\mathbf{r}} \times \hat{\mathbf{p}}) \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{p}}) = \hat{\mathbf{r}}^2\hat{\mathbf{p}}^2 - (\hat{\mathbf{r}} \cdot \hat{\mathbf{p}})^2 + i\hbar\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}$$

and compute

$$\langle\mathbf{r}|\hat{\mathbf{L}}^2|\psi\rangle = \langle\mathbf{r}|[\hat{\mathbf{r}}^2\hat{\mathbf{p}}^2 - (\hat{\mathbf{r}} \cdot \hat{\mathbf{p}})^2 + i\hbar\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}]|\psi\rangle$$
$$= r^2\langle\mathbf{r}|\hat{\mathbf{p}}^2|\psi\rangle + \hbar^2 r\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)\langle\mathbf{r}|\psi\rangle + \hbar^2\frac{\partial}{\partial r}\langle\mathbf{r}|\psi\rangle.$$

Solving for $\langle\mathbf{r}|\hat{\mathbf{p}}^2|\psi\rangle$ and substituting into the position-space representation of the energy eigenvalue equation gives

$$-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)\langle\mathbf{r}|\psi\rangle + \frac{\langle\mathbf{r}|\hat{\mathbf{L}}^2|\psi\rangle}{2\mu r^2} + V(r)\langle\mathbf{r}|\psi\rangle = E\langle\mathbf{r}|\psi\rangle.$$

We recognize the second term as the rotational kinetic energy, so the first term must be the radial part. Now, if $|\psi\rangle$ is a simultaneous eigenstate $|E, l, m\rangle$ the equation becomes

$$\left[-\frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r) \right] \langle \mathbf{r}|E, l, m\rangle = E\langle \mathbf{r}|E, l, m\rangle.$$

If we write $\langle \mathbf{r}|E, l, m\rangle = R(r)\Theta(\theta)\Phi(\phi)$, we can divide out the angular part of the wave function to get the radial equation

$$\left[-\frac{\hbar^2}{2\mu} \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} \right) + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r) \right] R(r) = ER(r),$$

and we can simplify it using the substitution $R(r) = u(r)/r$:

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

This has the exact same form as the one-dimensional Schrödinger equation, only this time with an effective potential $V_{\text{eff}}(r) = l(l+1)\hbar^2/2\mu r^2 + V(r)$. Notably, the radial equation has no dependence on the quantum number m; this is a consequence of the rotational invariance of the Hamiltonian, as there is no preferred axis upon which we should measure L_z. So solutions to this equation do not necessarily specify states uniquely—we must specify m (and the spin) to do so.

The set of commuting operators that are necessary to label states uniquely is called a complete set of commuting observables. For any given system there may exist several such sets—for example, we might have taken $\{\hat{H}, \hat{\mathbf{L}}^2, \hat{L}_x\}$ instead of $\{\hat{H}, \hat{\mathbf{L}}^2, \hat{L}_z\}$. (Again, in reality neither of these are complete set since we haven't yet seen how spin enters into $\hat{H}$.) For a system with cylindrical symmetry we might take $\{\hat{H}, \hat{L}_z, \hat{p}_z\}$.

4.4 The Total Angular Momentum Operator

We seek a position-space representation of $\hat{\mathbf{L}}^2$. We'll go component-by-component—for $\hat{L}_z$ we'll start by Taylor-expanding

$$\langle r, \theta, \phi|\hat{R}(d\phi\,\mathbf{k})|\psi\rangle = \langle r, \theta, \phi - d\phi|\psi\rangle$$
$$= \langle r, \theta, \phi|\psi\rangle - d\phi \frac{\partial \langle r, \theta, \phi|\psi\rangle}{d\phi},$$

which we can compare wth the definition of $\hat{R}(d\phi\,\mathbf{k})$ to write

$$\hat{L}_z \longrightarrow \frac{\hbar}{i} \frac{\partial}{\partial \phi}.$$

Note the similarity with the linear momentum operator. Now, for an eigenstate of $\hat{L}_z$ we have the differential equation

$$\frac{\hbar}{i} \frac{\partial}{\partial \phi} \langle r, \theta, \phi|l, m\rangle = m\hbar\langle r, \theta, \phi|l, m\rangle,$$

meaning the ϕ-dependence of the eigenfunction is $e^{im\phi}$ for $m \in \mathbb{Z}$ (so that the function is single-valued with period 2π). It follows, from our much earlier discussion of angular momentum, that $l = 0, 1, 2, \dots$.

To find the other components of angular momentum we use the gradient in spherical coordinates:

$$\hat{\mathbf{L}} \to \mathbf{r} \times \frac{\hbar}{i}\nabla = r\mathbf{u}_r \times \frac{\hbar}{i} \left(\mathbf{u}_r \frac{\partial}{\partial r} + \mathbf{u}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \mathbf{u}_\phi \frac{1}{r\sin\theta} \frac{\partial}{\partial \phi} \right)$$
$$= \frac{\hbar}{i} \left(\mathbf{u}_\phi \frac{\partial}{\partial \theta} - \mathbf{u}_\theta \frac{1}{\sin\theta} \frac{\partial}{\partial \phi} \right).$$

If we split the x and y components of these unit vectors we find that

$$\hat{L}_x \to \frac{\hbar}{i} \left(-\sin\phi \frac{\partial}{\partial \theta} - \cot\theta \cos\phi \frac{\partial}{\partial \phi} \right),$$
$$\hat{L}_y \to \frac{\hbar}{i} \left(\cos\phi \frac{\partial}{\partial \theta} - \cot\theta \sin\phi \frac{\partial}{\partial \phi} \right).$$

The total angular momentum operator is therefore represented by

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

This agrees entirely with the Laplacian in spherical coordinates! In fact, we can write the Schrödinger equation as

$$E\langle\mathbf{r}|\psi\rangle = \left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(r) \right] \langle\mathbf{r}|\psi\rangle$$

which, after expanding, agrees entirely with the energy eigenvalue equation we found before, provided we identify $\hat{\mathbf{L}}^2$ as we have.

The position-space representations of the angular momentum operators depend only on θ and ϕ, meaning we can isolate the angular dependence of our eigenfunctions and determine the amplitudes $\langle\theta,\phi|l,m\rangle = Y_{l,m}(\theta,\phi)$ called the spherical harmonics. Expressed this way, the energy eigenfunctions are

$$\langle r,\theta,\phi|E,l,m\rangle = R(r)Y_{l,m}(\theta,\phi).$$

The normalization condition for these eigenfunctions is

$$1 = \int d^3\mathbf{r}\, |R(r)|^2 |Y_{l,m}(\theta,\phi)|^2$$

$$= \left[\int_0^\infty r^2 dr\, |R(r)|^2 \right] \left[\int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\phi\, |Y_{l,m}(\theta,\phi)|^2 \right],$$

where we have substituted $d^3r = r^2 dr\, d\Omega = r^2 dr\, \sin\theta\, d\theta\, d\phi$. We normalize the radial and angular parts separately, so we can interpret

$$|\langle\theta,\phi|l,m\rangle|^2 d\Omega = |Y_{m,l}(\theta,\phi)|^2 d\Omega$$

as the probability of finding a particle in $|l,m\rangle$ within the solid angle $d\Omega$.

To determine the spherical harmonics we start by using the position-space representations of the angular momentum operators to write

$$\hat{L}_\pm \to \frac{\hbar}{i} e^{\pm i\phi} \left(\pm i\frac{\partial}{\partial\theta} - \cot\theta \frac{\partial}{\partial\phi} \right),$$

and so when we raise the highest angular momentum state we get

$$0 = \langle\theta,\phi|\hat{L}_+|l,l\rangle = \frac{\hbar}{i} e^{i\phi} \left(i\frac{\partial}{\partial\theta} - \cot\theta \frac{\partial}{\partial\phi} \right) \langle\theta,\phi|l,l\rangle.$$

With the known dependence on $e^{il\phi}$ the equation becomes $(\partial/\partial\theta - l\cot\theta)\langle 0,\phi|l,l\rangle = 0$, and our solution is

$$Y_{l,l}(\theta,\phi) = \langle\theta,\phi|l,l\rangle = c_l e^{il\phi} \sin^l\theta, \qquad c_l = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)!}{4\pi}}.$$

Successive applications of the lowering operator reveal that

$$Y_{l,m}(\theta,\phi) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}} e^{im\phi} \frac{1}{\sin^m\theta} \frac{d^{l-m}}{d(\cos\theta)^{l-m}} \sin^{2l}\theta.$$

This choice of phase factor ensures that $Y_{1,0}(\theta,\phi)$ has a real positive value for $\theta = 0$. For negative m, note that

$$Y_{l,-m}(\theta,\phi) = (-1)^m [Y_{l,m}(\theta,\phi)]^*.$$

The $l = 0$ and $l = 1$ states are called s and p states, respectively. The s state is spherically symmetric; the p states are a bit smashed on the xy-plane for $m = \pm 1$, but appears oriented on the z-axis for $m = 0$. A similar pattern exists for higher values of l. We can create harmonics oriented along different axes by taking superpositions of the "z-harmonics".

4.5 Miscellaneous Results in Atoms and Molecules

Now we'll look at a couple of ideas that will be useful in our future discussion of central potentials. First we estimate the energy scale for the hydrogen atom without actually solving the energy eigenvalue equation. The Hamiltonian is

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2\mu} - \frac{q_e^2}{|\hat{\mathbf{r}}|},$$

the ground-state expectation value is $E_1 = \langle \mathbf{p}^2/2\mu - q_e^2/r \rangle$, and using dimensional analysis we express $\langle q_e^2/r \rangle = q_e^2/a$, where a is a length characteristic of the atom's size. The uncertainty in r is at most of order a, so from the Heisenberg uncertainty principle $|\Delta p| \gtrsim \hbar/a$.

Now, the expectation value of the kinetic energy is

$$\frac{\langle \mathbf{p}^2 \rangle}{2\mu} = \frac{\Delta \mathbf{p}^2 + \langle \mathbf{p} \rangle^2}{2\mu} = \frac{\Delta \mathbf{p}^2}{2\mu},$$

since $\langle \mathbf{p} \rangle$ is time-independent in an eigenstate and $\langle \mathbf{p} \rangle \neq 0$ would give an unbounded system (it wouldn't stay within a particular region of space). The total energy estimate is thus

$$E_1 \gtrsim \frac{\hbar^2}{2\mu a^2} - \frac{q_e^2}{a},$$

and it is minimized at

$$a = \frac{\hbar^2}{m_e q_e^2} \implies E_1 \simeq -\frac{m_e q_e^4}{2\hbar^2},$$

where we've taken $\mu \simeq m_e$. This gives a on the order of angstroms and E_1 on the order of 10 eV.

Now we'll look at diatomic molecules. The potential energy $V(r)$ of such a molecule can be approximated as a harmonic oscillator with $\mu\omega^2 = d^2V/dr^2$. This energy is on the order of q_e^2/a, so by dimensional analysis

$$\frac{d^2V}{dr^2} \sim \frac{e^2}{a^3} \implies \hbar\omega = \hbar \left[\frac{1}{\mu} \left(\frac{d^2V}{dr^2} \right)_{r=r_0} \right]^{1/2} \sim \left(\frac{m_e}{M_N} \right)^{1/2} \left(\frac{m_e e^4}{\hbar^2} \right),$$

where μ (the reduced mass of the nucleus-nucleus system) is on the order of the nuclear mass M_N. We recognize the second factor as the electronic energy scale we found earlier, and it's been scaled down by roughly $(m_e/M_N)^{1/2} \approx 1/40$. Thus the wavelengths of photons absorbed or emitted when the system moves between vibrational energy levels are about 40 times longer than they are for a typical atomic transition. In particular, the purely vibrational energies are $E_{n_\nu} = (n_\nu + 1/2)\hbar\omega$.

When we look at the harmonic oscillator energy eigenfunctions we see that, at low excitations, the molecule vibrates on a distance scale

$$\frac{\sqrt{\hbar}}{\mu\omega} \sim \left(\frac{m_e}{M_N} \right)^{1/4} \frac{\hbar^2}{m_e e^2} = a \left(\frac{m_e}{M_N} \right)^{1/4},$$

which is a small fraction of the equilibrium separation r_0 between the nuclei. We therefore approximate the molecule as rigid and separate the rotational motion from the vibrational motion. The rigid rotator Hamiltonian is

$$\hat{H} = \frac{\hat{\mathbf{L}}^2}{2I}, \qquad I = \mu r_0^2,$$

and so its eigenstates are just the angular momentum eigenstates $|l, m\rangle$ with $E_l = l(l+1)\hbar^2/2I$. The spacing between energy levels is

$$E_l - E_{l-1} = \frac{l\hbar^2}{I} \sim l \frac{\hbar^2}{M_N a^2} = l \frac{m_e}{M_N} \left(\frac{m_e e^4}{\hbar^2} \right),$$

which corresponds to wavelength a factor of $(M_n/m_e)^{1/2}$ longer, for low l, than for vibrational transitions. This is only slightly below order-$k_B T$ at room temperature, so many levels are excited here. We can take advantage of this to experimentally determine r_0; setting $E_l - E_{l-1} = hc/\lambda$ gives the observable

$$\lambda l = 2\pi I c/\hbar,$$

which we've found to be constant in experiment in alignment with our treatment of the molecule as a rotator!

In practice, it's difficult to produce the wavelengths necessary for observations of purely rotational transitions. But putting rotational energy alongside vibrational energy gives

$$E_{n_\nu, l} = \left(n_\nu + \frac{1}{2} \right) \hbar\omega + \frac{l(l+1)\hbar^2}{2I},$$

and since it turns out that most transitions satisfy $\Delta n_\nu = \pm 1$ and $\Delta l = \pm 1$, the frequencies resulting from vibrational-rotational transitions are more accessible than we may believe.

4.6   Solutions in the Coulomb Potential

Now we will solve the radial equation for a variety of potentials, but first we'll do a more general analysis of the behavior of our solutions near the origin. If $V(r)$ is not more singular than $1/r^2$ at the origin then we are guaranteed a power series solution, so to determine the leading behavior of $u_{E,l}(r)$ we substitute r^s into the radial equation to get

$$-\frac{\hbar^2}{2\mu}s(s-1)r^{s-2} + \frac{l(l+1)\hbar^2}{2\mu}r^{s-2} + V(r)r^s = Er^s.$$

The r^{s-2} terms dominate here, meaning $-s(s-1) + l(l+1) = 0$ and so $s = l+1$ or $s = -l$. But we must discard the r^{-l} solutions for normalizability, so we must have

$$u_{E,l}(r) \longrightarrow r^{l+1} \implies R_{E,l}(r) \longrightarrow r^l.$$

Our wave functions therefore vanish at the origin for all states except for $l = 0$. Notice how this dependence on l is in line with the "centrifugal term" in the effective potential $V_{\text{eff}}(r)$ from earlier!

Now let's solve the radial equation for the Coulomb potential $V(r) = -Zq_e^2/r$. All bound states of such a system have negative energy, so it is convenient to write $E = -|E|$ and substitute $\rho = \sqrt{8\mu|E|/\hbar^2}\, r$ to get

$$\frac{d^2 u}{d\rho^2} - \frac{l(l+1)}{\rho^2}u + \left(\frac{\lambda}{\rho} - \frac{1}{4} \right) u = 0, \qquad \lambda = \frac{Ze^2}{\hbar}\sqrt{\frac{\mu}{2|E|}}.$$

Solving this using power series methods gives a nasty recurrence relation. But we can observe that the solutions look like $u(\rho) = Ae^{-\rho/2}$ in the $\rho \to \infty$ limit, so we make the substitution $u(\rho) = \rho^{l+1}e^{-\rho/2}F(\rho)$ to get

$$\frac{d^2 F}{d\rho^2} = \left(\frac{2l+2}{\rho} - 1 \right) \frac{dF}{d\rho} - \left(\frac{\lambda}{\rho} - \frac{l+1}{\rho} \right) F = 0,$$

which leads to the first-order

$$\frac{c_{k+1}}{c_k} = \frac{k+l+1-\lambda}{(k+1)(k+2l+2)}.$$

This has exponential limiting behavior, so we truncate the series by requiring $\lambda = 1 + l + n_r$ for any $n_r = 0, 1, 2, \ldots$. From our definition of λ we therefore have the quantized energies

$$E_n = -\frac{\mu Z^2 q_e^4}{2\hbar^2 n^2}, \quad n = 1, 2, 3, \ldots,$$

where $n = l+1+n_r$ is called the principal quantum number. (We often write E_n in terms of the dimensionless fine-structure constant $\alpha = q_e^2/\hbar c$.) This leads to a shocking amount of degeneracy—not only does each l have $2l + 1$ values of m all with the same energy, but the $l = 0, 1, \ldots, n-1$ also all have the same energy, despite the different centrifugal terms! The total degeneracy is therefore $\sum_{l=0}^{n-1}(2l+1) = n^2$. (There is no "obvious" symmetry causing this degeneracy in l, so we call it an "accidental degeneracy".)

In the ground state the power series F is constant, and so

$$u_{1,0}(\rho) = c_0\rho e^{-\rho/2} \implies R_{1,0}(r) = 2\left(\frac{Z}{a_0} \right)^{3/2} e^{-Zr/a_0},$$

where we've used our expressions for ρ and $|E|$ to write $\rho = (2Z/n)(r/a_0)$. The length $a_0 = \hbar/\mu c\alpha$, called the Bohr length, is a convenient length scale for expressing the wave functions. Note that the eigenfunctions $R_{n,n-1}$ only have one "bump" in their probability distributions $r^2|R_{n,n-1}|^2$, adding one more bump for each drop in l. These bumps will be key in determining the order in which states fill up in multielectron atoms.

4.7 Solutions in the Finite Spherical Well

Now we'll move into the realm of nuclear physics by investigating the bound state of a proton and neutron (called a deuteron). The finite spherical well

$$V(r) = \begin{cases} -V_0 & r < a, \\ 0 & r > a. \end{cases}$$

provides a crude approximation of the strong interaction between the two particles. For the ground state we have $l = 0$, so the radial equation is

$$\frac{d^2u}{dr^2} = -k_0^2 u, \quad r < a, \qquad k_0^2 = \frac{2\mu(V_0 + E)}{\hbar^2},$$

$$\frac{d^2u}{dr^2} = q^2 u, \quad r > a, \qquad q^2 = -\frac{2\mu E}{\hbar^2}.$$

Solving and imposing the boundary condition $u(0) = 0$ yields

$$u(r) = A \sin k_0 r, \quad r < a,$$
$$u(r) = C e^{-qr}, \qquad r > a.$$

Continuity of $u(r)$ and $u'(r)$ at $r = a$ yields

$$A \sin k_0 a = C e^{-qa},$$
$$A k_0 \cos k_0 a = -q C e^{-qa},$$

and dividing these gives $\tan k_0 a = -k_0/q$. If we define $k_0 a = \zeta$ and $qa = \eta$ we can therefore write

$$\zeta \cot \zeta = -\eta, \qquad \zeta^2 + \eta^2 = r_0^2, \qquad r_0^2 = \frac{2\mu V_0 a^2}{\hbar^2}$$

Each solution to this system corresponds to a different bound state. A graphical analysis reveals that there are no bound states unless $r_0 > \pi/2$, that there is exactly one bound state for $\pi/2 < r_0 < 3\pi/2$, and so on. This is in contrast with the one-dimensional finite square well, which always has at least one bound state.

The finite spherical well obviously isn't a very realistic model for the deuteron's interaction potential—experiment has shown, for example, that the deuteron has exactly one bound state. It can still tell us some useful things, though! For example, the deuteron has an intrinsic spin of one, and the proton and neutron do not bind in a spin-0 state, meaning the strong nuclear force is spin-dependent.

4.8 Solutions in the Infinite Spherical Well

For heavier nuclei, we imagine each nucleon as being confined to its own spherical potential well. This well gets wider as we add more nucleons, but given that (for $l = 0$) the effect of increasing a is the same as that for increasing V_0, we'll just examine the infinite spherical well

$$V(r) = \begin{cases} 0 & r < a, \\ \infty & r > a. \end{cases}$$

We begin with the radial equation in $r < a$:

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} - \frac{l(l+1)}{r^2}R + k^2 R = 0, \qquad k^2 = \frac{2\mu E}{\hbar^2}.$$

We could solve this equation using power series, but we can define the dimensionless variable $\rho = kr$ to get

$$\frac{d^2R}{d\rho^2} + \frac{2}{\rho}\frac{dR}{d\rho} + \left[1 - \frac{l(l+1)}{\rho^2}\right]R = 0,$$

which we can solve in terms of simple functions (applying the proper boundary conditions):

$$j_l(\rho) = (-\rho)^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \left(\frac{\sin \rho}{\rho} \right),$$

$$\eta_l(\rho) = -(-\rho)^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \left(\frac{\cos \rho}{\rho} \right).$$

These are called the spherical Bessel and Neumann functions, respectively, but since $u(r) = rR(r)$ must vanish at $r = 0$ we can discard the Neumann function. We also require that $R(a) = 0$ for continuity, so the energy eigenvalues are determined by

$$j_l(ka) = 0.$$

The $l = 0$ condition, $j_0(ka) = (\sin ka)/ka = 0$, is satisfied when

$$E_{n_r,l=0} = \frac{\hbar^2 k^2}{2\mu} = \frac{\hbar^2}{2\mu} \left(\frac{n_r \pi}{a} \right)^2, \qquad n_r = 1, 2, 3, \ldots,$$

which agrees with the finite well in the limit $V_0 \to \infty$. For the higher-order Bessel functions we can look up the zeroes in a table. The most stable nuclei are the ones for which the protons and neutrons fill full energy levels n_r (as we will later see, we can place at most two protons and two neutrons in the same state).

4.9 Solutions in the Isotropic Harmonic Oscillator

As our final example, we consider the three-dimensional isotropic (i.e., symmetric in all directions) harmonic oscillator:

$$V(r) = \frac{1}{2} \mu \omega^2 r^2.$$

This problem is easily solved in Cartesian coordinates by leveraging our previous results—the Hamiltonian is simply the sum of three (commuting) one-dimensional Hamiltonians in the x, y, and z directions, so the eigenstates are $|E\rangle = |E_x, E_y, E_z\rangle$ and the energies are

$$E = \left(n + \frac{3}{2} \right) \hbar\omega, \qquad n = 0, 1, 2, \ldots,$$

where $n = n_x + n_y + n_z$. The eigenfunctions are found by performing separation of variables on the Schrödinger equation; the result is a set of functions of the form $X_{n_x}(x) Y_{n_y}(y) Z_{n_z}(z)$, where each piece is an eigenfunction of the one-dimensional harmonic oscillator.

Alternatively, we can use spherical coordinates to write the radial equation

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

Alternatively,

$$\frac{d^2 u}{d\rho^2} - \frac{l(l+1)}{\rho^2} u - \rho^2 u = -\lambda u, \qquad \rho = \sqrt{\frac{\mu\omega}{\hbar}} r, \quad \lambda = \frac{2E}{\hbar\omega}.$$

We once again search for a solution of the form $u(r)\rho^{l+1} e^{-\rho^2/2} f(\rho)$ using power series methods, arriving at a solution with exponential with limiting behavior unless the series terminates. The quantization condition

$$E = \left(2n_r + l + \frac{3}{2} \right) \hbar\omega, \qquad n_r = 0, 1, 2, \ldots,$$

which matches with our previous energies provided we take $n = 2n_r + l$.

The high degree of degeneracy for the harmonic oscillator suggests that there is a "hidden" symmetry at play that has not yet been accounted for, just like we saw with the hydrogen atom. The relevant constant of motion can be seen from classical mechanics—the $1/r$ and r^2 potentials turn out to be the only ones for which orbits close on themselves and don't precess, so the direction of the vector pointing from the orbit's apocenter to its pericenter is preserved.

5 More Complex Systems

5.1 Time-Independent Perturbation Theory

Lots of the systems we care about don't have Hamiltonians that we can solve exactly. But in many cases they are sufficiently close to ones that we can solve, in which case we can treat the differences as small "perturbations" and real with them in a systematic way. We begin by expressing such a Hamiltonian in the form

$$\hat{H} = \hat{H}_0 + \hat{H}_1,$$

where we say the "big" $\hat{H}_0$ is the unperturbed Hamiltonian and the "small" $\hat{H}_1$ is the perturbing Hamiltonian. The unperturbed part has known eigenstates and eigenvalues that satisfy

$$\hat{H}_0|\varphi_n^{(0)}\rangle = E_n^{(0)}|\varphi_n^{(0)}\rangle,$$

and we seek an approximate solution to $\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle$. To this end we introduce a "tuning" parameter λ via $\hat{H} = \hat{H}_0 + \lambda\hat{H}_1$ and assume that we can express the exact eigenstates and eigenvalues as power-series expansions in λ:

$$|\psi_n\rangle = |\varphi_n^{(0)}\rangle + \lambda|\varphi_n^{(1)}\rangle + \lambda^2|\varphi_n^{(2)}\rangle + \cdots,$$
$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots.$$

We require that as $\lambda \to 0$, the energies and eigenstates go smoothly to the exact values for the unperturbed Hamiltonian. Now, if we substitute these expressions into the energy eigenvalue equation we get

$$\left(\hat{H}_0 + \lambda\hat{H}_1\right)\left(|\varphi_n^{(0)}\rangle + \lambda|\varphi_n^{(1)}\rangle + \lambda^2|\varphi_n^{(2)}\rangle + \cdots\right)$$
$$= \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots\right)\left(|\varphi_n^{(0)}\rangle + \lambda|\varphi_n^{(1)}\rangle + \lambda^2|\varphi_n^{(2)}\rangle + \cdots\right).$$

Expanding and collecting like terms yields

$$\hat{H}_0|\varphi_n^{(0)}\rangle = E_n^{(0)}|\varphi_n^{(0)}\rangle,$$
$$\hat{H}_0|\varphi_n^{(1)}\rangle + \hat{H}_1|\varphi_n^{(0)}\rangle = E_n^{(0)}|\varphi_n^{(1)}\rangle + E_n^{(1)}|\varphi_n^{(0)}\rangle,$$
$$\hat{H}_0|\varphi_n^{(2)}\rangle + \hat{H}_1|\varphi_n^{(1)}\rangle = E_n^{(0)}|\varphi_n^{(2)}\rangle + E_n^{(1)}|\varphi_n^{(1)}\rangle + E_n^{(2)}|\varphi_n^{(0)}\rangle$$

for the λ^0, λ^1, and λ^2 terms, respectively. A pattern emerges in higher-order terms, but these suffice for us. These equations encode information about the corrections we must make to the eigenstates and energies!

The first-order correction to the energy is found by applying $\langle\varphi_n^{(0)}|$ to the λ^1 equation:

$$\langle\varphi_n^{(0)}|\hat{H}_0|\varphi_n^{(1)}\rangle + \langle\varphi_n^{(0)}|\hat{H}_1|\varphi_n^{(0)}\rangle = \langle\varphi_n^{(0)}|E_n^{(0)}|\varphi_n^{(1)}\rangle + \langle\varphi_n^{(0)}|E_n^{(1)}|\varphi_n^{(0)}\rangle,$$
$$\langle\varphi_n^{(0)}|\hat{H}_1|\varphi_n^{(0)}\rangle = E_n^{(1)}.$$

Thus the first-order shift in the nth energy level is simply the perturbing $\langle E\rangle$ for the nth unperturbed state. If we were to instead apply some other $\langle\varphi_k^{(0)}|$ with $k \neq n$ we'd get

$$\langle\varphi_k^{(0)}|\hat{H}_0|\varphi_n^{(1)}\rangle + \langle\varphi_k^{(0)}|\hat{H}_1|\varphi_n^{(0)}\rangle = E_n^{(0)}\langle\varphi_k^{(0)}|\varphi_n^{(1)}\rangle \implies \langle\varphi_k^{(0)}|\varphi_n^{(1)}\rangle = \frac{\langle\varphi_k^{(0)}|\hat{H}_1|\varphi_n^{(0)}\rangle}{E_n^{(0)} - E_k^{(0)}},$$

meaning $|\varphi_n^{(1)}\rangle$ generally has a nonzero component along $|\varphi_k^{(0)}\rangle$ for $k \neq n$. For the $k = n$ case we use the expansion from before to write

$$1 = \langle\psi_n|\psi_n\rangle = \langle\varphi_n^{(0)}|\varphi_n^{(0)}\rangle + \lambda\langle\varphi_n^{(0)}|\varphi_n^{(1)}\rangle + \lambda\langle\varphi_n^{(1)}|\varphi_n^{(0)}\rangle + \cdots,$$

meaning $\langle \varphi_n^{(0)} | \varphi_n^{(1)} \rangle$ is purely imaginary. Inserting the identity operator into the $|\psi_n\rangle$ expansion gives

$$|\psi_n\rangle = |\varphi_n^{(0)}\rangle + ia\,\lambda|\varphi_n^{(0)}\rangle + \lambda \sum_{k \neq n} |\varphi_k^{(0)}\rangle \langle \varphi_k^{(0)} | \varphi_n^{(1)} \rangle + \mathcal{O}(\lambda^2)$$

$$= e^{ia\lambda}|\varphi_n^{(0)}\rangle + \lambda \sum_{k \neq n} |\varphi_k^{(0)}\rangle \langle \varphi_k^{(0)} | \varphi_n^{(1)} \rangle + \mathcal{O}(\lambda^2).$$

Even after normalizing $|\psi_n\rangle$ it turns out that a is not determined, so we can set $a = 0$ so that the first-order correction is orthogonal to the unperturbed state and

$$= |\varphi_n^{(0)}\rangle + \lambda \sum_{k \neq n} |\varphi_k^{(0)}\rangle \frac{\langle \varphi_k^{(0)} | \hat{H}_1 | \varphi_n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}} + \mathcal{O}(\lambda^2).$$

This summation is the $|\varphi_n^{(1)}\rangle$ we sought. For the second-order energy shift we take the λ^2 equation and apply $\langle \varphi_n^{(0)} |$; after some cancellation we can substitute this expression for $|\varphi_n^{(1)}\rangle$ to get

$$E_n^{(2)} = \langle \varphi_n^{(0)} | \hat{H}_1 | \varphi_n^{(1)} \rangle$$

$$= \sum_{k \neq n} \frac{\langle \varphi_n^{(0)} | \hat{H}_1 | \varphi_k^{(0)} \rangle \langle \varphi_k^{(0)} | \hat{H}_1 | \varphi_n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}} = \sum_{k \neq n} \frac{\left| \langle \varphi_k^{(0)} | \hat{H}_1 | \varphi_n^{(0)} \rangle \right|^2}{E_n^{(0)} - E_k^{(0)}}.$$

Everything we've done so far has assumed that there is no degeneracy in $E_n^{(0)}$. But if there is, we then face the additional problem of figuring out which combinations of the degenerate states will converge to the correct eigenstates and energies of $\hat{H}$. To solve this problem we first allow for degeneracy in writing

$$|\psi_n\rangle = \sum_{i=1}^{N} c_i |\varphi_{n,i}^{(0)}\rangle + \lambda|\varphi_n^{(1)}\rangle + \cdots,$$

and if we substitute this into the energy eigenvalue equation we obtain, for the λ^1 term,

$$\hat{H}_0 |\varphi_n^{(1)}\rangle + \hat{H}_1 \sum_{i=1}^{N} c_i |\varphi_{n,i}^{(0)}\rangle = E_n^{(0)} |\varphi_n^{(1)}\rangle + E_n^{(1)} \sum_{i=1}^{N} c_i |\varphi_{n,i}^{(0)}\rangle.$$

Note the similarities to the λ^1 equation from before. Now we apply each of the degenerate $\langle \varphi_{n,j}^{(0)} |$ to get

$$\sum_{i=1}^{N} \langle \varphi_{n,j}^{(0)} | \hat{H}_1 | \varphi_{n,i}^{(0)} \rangle c_i = E_n^{(1)} \sum_{i=1}^{N} \langle \varphi_{n,j}^{(0)} | \varphi_{n,i}^{(0)} \rangle = E_n^{(1)} \sum_{i=1}^{N} \delta_{ji} c_i = E_n^{(1)} c_j,$$

where we have assumed that the degenerate states are orthonormal. We can recognize this as the jth component of an energy eigenvalue equation! In the $N = 2$ case the equation looks like

$$\begin{bmatrix} \langle \varphi_{n,1}^{(0)} | \hat{H}_1 | \varphi_{n,1}^{(0)} \rangle & \langle \varphi_{n,1}^{(0)} | \hat{H}_1 | \varphi_{n,2}^{(0)} \rangle \\ \langle \varphi_{n,2}^{(0)} | \hat{H}_1 | \varphi_{n,1}^{(0)} \rangle & \langle \varphi_{n,2}^{(0)} | \hat{H}_1 | \varphi_{n,2}^{(0)} \rangle \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = E_n^{(1)} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Notice howe we're looking only at how the Hamiltonian behaves in the subspace of degenerate states The eigenvalues are the first-order shifts, and the eigenstates are the "correct" linear combinations of the degenerate states that we sought.

5.2 Perturbing Electric Fields

As a first example we will apply an external electric field $\mathbf{E} = |\mathbf{E}|\mathbf{z}$ to the hydrogen atom. The relevant Hamiltonians are

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2\mu} - \frac{q_e^2}{|\hat{\mathbf{r}}|}, \qquad \hat{H}_1 = q_e|\mathbf{E}|\hat{z}.$$

The ground-state first-order shift can be easily computed using the nondegenerate theory: we could show that the eigenstates $|n, l, m\rangle$ have definite parity $(-1)^l$, meaning the position expectation value

$$E_1^{(1)} = q_e |\mathbf{E}| \langle 1, 0, 0 | \hat{z} | 1, 0, 0 \rangle = 0.$$

The second-order shift is

$$E_1^{(2)} = \sum_{\substack{n,l,m \\ n \neq 1}} \frac{q_e^2 |\mathbf{E}|^2 \, |\langle n, l, m | \hat{z} | 1, 0, 0 \rangle|^2}{E_1^{(0)} - E_n^{(0)}}.$$

As for the first-excited states, the degenerate subspace is spanned by $\{|2, 0, 0\rangle, |2, 1, 0\rangle, |2, 1, 1\rangle, |2, 1, -1\rangle\}$. In principle this gives us sixteen matrix elements to compute, but we can use symmetry to simplify things considerably. For one, all of the elements in which the bra and ket have the same l-value vanish immediately, since the parity $(-1)^l$ depends only on l.

For most of the remaining states, we can also take advantage of the fact that $[\hat{H}_1, \hat{L}_z] = 0$ to write

$$m' \hbar \langle n, l', m' | \hat{z} | n, l, m \rangle = \langle n, l', m' | \hat{L}_z \hat{z} | n, l, m \rangle$$
$$= \langle n, l', m' | \hat{z} \hat{L}_z | n, l, m \rangle$$
$$= m \hbar \langle n, l', m' | \hat{z} | n, l, m \rangle.$$

Thus the matrix elements with different m-values vanish. (This is true in general—if an operator $\hat{A}$ commutes with $\hat{H}_1$ then matrix elements with different eigenvalues a vanish.) Thus the only entry we must evaluate is

$$\langle 2, 0, 0 | \hat{H}_1 | 2, 1, 0 \rangle = q_e |\mathbf{E}| \int_0^\infty r^2 dr \int_0^\pi \sin \theta \, d\theta \int_0^{2\pi} d\phi \, R_{2,0}^* Y_{0,0}^* r \cos \theta \, R_{2,1} Y_{1,0} = -3 q_e |\mathbf{E}| a_0,$$

where a_0 is the Bohr radius. We can use this to construct the matrix representation of $\hat{H}_1$, and some computations give the first-order shifts $E_2^{(1)} = 0, \, 0, \, 3 q_e |\mathbf{E}| a_0, \, -3 q_e |\mathbf{E}| a_0$ with corresponding kets

$$|2, 1, 1\rangle, \quad |2, 1, -1\rangle, \quad \frac{1}{\sqrt{2}} \big(|2, 0, 0\rangle - |2, 1, 0\rangle \big), \quad \frac{1}{\sqrt{2}} \big(|2, 0, 0\rangle + |2, 1, 0\rangle \big),$$

respectively. This splitting is known as the Stark effect. Notice that, as a consequence of $[\hat{H}_1, \hat{L}_z] = 0$, eigenkets with different m-values are not mixed by the perturbation. Commutativity therefore would have allowed us to focus only on the $m = 0$ states, analyzing the others using the nondegenerate theory.

As a second example, consider again the ammonia molecule in a weak external electric field. We have, in the energy eigenbasis,

$$\hat{H}_0 \to \begin{bmatrix} E_0 - A & 0 \\ 0 & E_0 + A \end{bmatrix}, \qquad \hat{H}_1 \to \begin{bmatrix} 0 & \mu_e |\mathbf{E}| \\ \mu_e |\mathbf{E}| & 0 \end{bmatrix}.$$

The vanishing first-order shift is explained by $\hat{\Pi} |I\rangle = |I\rangle$ and $\hat{\Pi} |II\rangle = -|II\rangle$, as the electric dipole moment operator has a vanishing expectation value in a state of definite parity. Notice how this conclusion is in agreement with the Taylor expansion

$$E = E_0 \pm A \pm A \left[\frac{1}{2} \left(\frac{\mu_e |\mathbf{E}|^2}{A} \right)^2 - \frac{1}{8} \left(\frac{\mu_e |\mathbf{E}|^2}{A} \right)^4 + \cdots \right].$$

But in the case of a strong electric field $\mu_e |\mathbf{E}| \gg A$ the expansion is

$$E = E_0 \pm \mu_e |\mathbf{E}| \pm \frac{A^2}{2 \mu_e |\mathbf{E}|} \mp \cdots,$$

which has a term that is linear in $|\mathbf{E}|$. We therefore have a nonzero first-order shift, suggesting that we cannot use the nondegenerate theory here. Roughly speaking, this is because the states $|I\rangle$ and $|II\rangle$ are very close together in energy, so the factor

$$\frac{\langle II | \hat{H}_1 | I \rangle}{E_{II}^{(0)} - E_I^{(0)}} = \frac{\mu_e |\mathbf{E}|}{2A}$$

diverges. Thus we use the degenerate theory on the $|1\rangle$-$|2\rangle$ representations of $\hat{H}_0$ and $\hat{H}_1$ to determine the relevant information.

5.3   Relativistic Perturbations to the Hydrogen Atom

We've already studied the hydrogen atom extensively, but now we'll discuss three relativistic perturbations to the Hamiltonian that together create a fine structure to our previously-derived energy levels. First, we write the relativistically correct expression for the kinetic energy in the system (using $m_p \gg m_e$):

$$\hat{K} = m_e c^2 \left(\sqrt{1 + (\hat{\mathbf{p}}_e^2/m_e^2 c^2)} - 1 \right) + \frac{\hat{\mathbf{p}}_p^2}{2m_p}$$

$$= \frac{\hat{\mathbf{p}}_e^2}{2m_e} - \frac{(\hat{\mathbf{p}}_e^2)^2}{8m_e^3 c^2} + \cdots + \frac{\hat{\mathbf{p}}_p^2}{2m_p} \;=\; \frac{\hat{\mathbf{p}}^2}{2\mu} - \frac{(\hat{\mathbf{p}}^2)^2}{8m_e^3 c^2} + \cdots .$$

Our perturbing Hamiltonians is therefore $\hat{H}_K = -(\hat{\mathbf{p}}^2)^2/8m_e^3 c^2$. Conveniently, $[\hat{H}_K, \hat{\mathbf{L}}^2] = 0$, so we can use the nondegenerate theory to compute the first-order shift

$$E_{n,l}^{(1)} = -\langle n, l, m| \frac{(\hat{\mathbf{p}}^2)^2}{8m_e^3 c^2} |n, l, m\rangle$$

$$= -\frac{1}{2m_e c^2} \langle n, l, m| \left(\hat{H}_0 + \frac{Zq_e^2}{|\hat{\mathbf{r}}|} \right)^2 |n, l, m\rangle.$$

Going through the motions here reveals that

$$E_K^{(1)} = -\frac{1}{2} m_e c^2 Z^4 \alpha^4 \left(-\frac{3}{4n^4} + \frac{1}{n^3(l + 1/2)} \right).$$

Now we'll look at the interaction between the electron's spin and its orbital angular momentum. In the electron's rest frame the proton generates a magnetic field $\mathbf{B}$, and so the interaction energy is

$$-\boldsymbol{\mu} \cdot \mathbf{B} = -\left(-\frac{ge}{2m_e c} \mathbf{S} \cdot \frac{-Zq_e \mathbf{v} \times \mathbf{r}}{cr^3} \right) = \frac{Zq_e^2}{m_e^2 c^2 r^3} \mathbf{S} \cdot \mathbf{L},$$

where $g = 2$ for the electron, and $\mathbf{L}$ and $\mathbf{v}$ are the electron's angular momentum and velocity (in the proton's frame). An effect called Thomas precession reduces this energy by a factor of two, so

$$\hat{H}_{\text{S-O}} = \frac{Zq_e^2}{2m_e^2 c^2 |\hat{\mathbf{r}}|^3} \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}.$$

The work we'll do here is very similar to what we did much earlier for the spin-spin interaction between the proton and electron. Rather than work with the total spin $\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2$, we will use the total angular momentum $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$:

$$\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}},$$

$$\hat{J}_z = \hat{L}_z + \hat{S}_z,$$

We can see that $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ commutes with $\hat{J}_z$—the latter commutes with $\hat{\mathbf{J}}^2$, $\hat{\mathbf{L}}^2$, and $\hat{\mathbf{S}}^2$, so it must also commute with $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$. Also notice that each of these three "standalone" operators commute with one another, and so they must have simultaneous eigenstates with $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$; in particular, $[\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{\mathbf{L}}^2] = 0$.

This simplifies things considerably. We now know that only states with the same values of l and m_j can be mixed in the spin-orbit perturbation. For fixed l and $m_j = (m_l + 1/2)\hbar$, the only two such states are

$$|l, m_l, +\mathbf{z}\rangle, \qquad |l, m_l + 1, -\mathbf{z}\rangle,$$

unless $m_l = l$ (in which case only the former state exists). We'll now determine the two linear combinations of these states that are eigenstates of $2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \hat{L}_+ \hat{S}_- + \hat{L}_- \hat{S}_+ + 2\hat{L}_z \hat{S}_z$. The matrix representation in the above basis is

$$2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \rightarrow \hbar^2 \begin{bmatrix} m & \sqrt{l(l+1) - m(m+1)} \\ \sqrt{l(l+1) - m(m+1)} & -(m+1) \end{bmatrix},$$

and the eigenvalues are $\lambda = l$ and $\lambda = -(l+1)$. So the total angular momentum j satisfies

$$j(j+1) = l(l+1) + \tfrac{1}{2}(\tfrac{1}{2} + 1) + \begin{cases} l, \\ -(l+1), \end{cases}$$

meaning $j = l \pm 1/2$. We therefore label the eigenstates

$$|j = l + \tfrac{1}{2}, m_j\rangle = \sqrt{\frac{l + m_l + 1}{2l + 1}} \, |l, m, +\mathbf{z}\rangle + \sqrt{\frac{l - m_l}{2l + 1}} \, |l, m + 1, -\mathbf{z}\rangle,$$

$$|j = l - \tfrac{1}{2}, m_j\rangle = \sqrt{\frac{l - m_l}{2l + 1}} \, |l, m, +\mathbf{z}\rangle - \sqrt{\frac{l + m_l + 1}{2l + 1}} \, |l, m + 1, -\mathbf{z}\rangle.$$

Note that we can use $m_j = m_l + 1/2$ to write the above directly in terms of m_j. The first order shift for each of these is

$$E_{\text{S-O}}^{(1)} = \frac{m_e c^2 Z^4 \alpha^4}{4 n^3 l (l + 1/2)(l + 1)} \begin{cases} l & j = l + 1/2, \\ -(l + 1) & j = l - 1/2. \end{cases}$$

To label these states we use the spectroscopic notation $1s_{1/2}$, $2s_{1/2}$, $2p_{1/2}$, $2p_{3/2}$, and so on—the coefficient denotes n, the letter l, and the subscript j.

But there's a problem here: evaluating the above for $l = 0$ gives a nonzero result, but a state with zero orbital angular momentum should have no spin-orbit coupling! Solving using the more exact Dirac equation gives a result consistent with this, but it also spits out an additional perturbation

$$\hat{H}_D = -\frac{1}{8 m_e^2 c^2} \Big[\hat{\mathbf{p}} \cdot, \big[\hat{\mathbf{p}}, V(|\hat{\mathbf{r}}|)\big]\Big],$$

where the momentum operators are dotted with one another in the expansion. When we go to compute the first-order shift due to this Darwin term, we find that it contributes only for s states—in fact, its magnitude

$$E_D^{(1)} = \langle n, 0, 0|\hat{H}_D|n, 0, 0\rangle = \frac{m_e c^2 Z^4 \alpha^4}{2 n^3}$$

is exactly the same as the seemingly nonphysical $l = 0$ contribution from the spin-orbit Hamiltonian! Finally, if we were to add together all of these first-order shifts we would get

$$E_K^{(1)} + E_{\text{S-O}}^{(1)} + E_D^{(1)} = E_{n,j}^{(1)} = -\frac{m_e c^2 (Z\alpha)^4}{2 n^3} \left(\frac{1}{j + 1/2} - \frac{3}{4n}\right),$$

which is $(Z\alpha)^2$ times the unperturbed energy (hence the term "fine structure"). The surprising degeneracy in l is, as it turns out, maintained to all orders! (We could show this by finding an exact solution the Dirac equation with a Coulomb potential.)

There's one more phenomenon here that we haven't yet considered: a very small but observable energy difference between the $2s_{1/2}$ and $2p_{1/2}$ levels. This Lamb shift, which is on the order of $m_e c^2 (Z\alpha)^4 \alpha \log \alpha$, is explained by quantum electrodynamics. Taken with the hyperfine structure discussed much earlier, all this gives us a much more complete look at the energy structure of the hydrogen atom.

5.4 The Zeeman Effect

We've seen what happens when the hydrogen atom is put in an external electric field. For a magnetic field $\mathbf{B} = B\mathbf{z}$,

$$\hat{H}_B = -\boldsymbol{\mu} \cdot \mathbf{B}$$
$$= \frac{q_e B}{2 m_e c}(\hat{L}_z + 2\hat{S}_z) = \frac{q_e B}{2 m_e c}(\hat{J}_z + \hat{S}_z),$$

which clearly commutes with $\hat{J}_z$. We can therefore compute the shifts in the (nondegenerate) $j = l \pm 1/2$ states:

$$E_B^{(1)} = \frac{q_e B}{2 m_e c} \langle j = l \pm \tfrac{1}{2}, m_j|(\hat{J}_Z + \hat{S}_z)|j = l \pm \tfrac{1}{2}, m_j\rangle$$
$$= \frac{q_e \hbar B}{2 m_e c} m_j \left(1 \pm \frac{1}{2l + 1}\right) = \frac{g(j, l) q_e \hbar B}{2 m_e c} m_j, \qquad g(j = l \pm \tfrac{1}{2}, l) = 1 \pm \frac{1}{2l + 1},$$

where $g(j, l)$ is called the Landé g factor.

5.5   Identical Particles

For the remainder of the chapter we will focus on systems of identical particles. The key here is that, unlike in the macroscopic world, identical particles in quantum mechanics are entirely indistinguishable. There is no measurement we can take to distinguish, say, one electron from any other.

To formalize this we introduce the exchange operator $\hat{P}_{12}$, which "exchanges" the roles of particles 1 and 2 in a system. In other words, we swap how the kets are labeled:

$$\hat{P}_{12}|a, b\rangle = |b, a\rangle \quad \text{or} \quad \hat{P}_{12}(|a\rangle_1 \otimes |b\rangle_2) = |b\rangle_1 \otimes |a\rangle_2.$$

The "exchanged" must be the same overall physical state as the original state, meaning the states must differ by at most an overall phase; further, since $\hat{P}_{12}^2 = 1$, it follows that the exchange operator has eigenvalues ± 1. A state with $\lambda = 1$ is called symmetric, while one with $\lambda = -1$ is called antisymmetric.

Since we are doing nonrelativistic quantum mechanics we must state the spin-statistics theorem as a law of nature:

- particles with integral spin, called bosons, are found only in symmetric states and obey Bose-Einstein statistics; and

- those with half-integral spin, called fermions, are found only in antisymmetric states and obey Fermi-Dirac statistics.

If two identical particles are in the same state then clearly the state is symmetric under exchange; it follows that no two fermions can be in the same state. This is known as the Pauli exclusion principle.

5.6   The Helium Atom

As an example of a system with two identical fermions, consider an atom with two electrons and a nuclear charge Z. We will take the unperturbed and perturbed Hamiltonians to be

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}_1^2}{2m_e} - \frac{Zq_e^2}{|\hat{\mathbf{r}}_1|} + \frac{\hat{\mathbf{p}}_2^2}{2m_e} - \frac{Zq_e^2}{|\hat{\mathbf{r}}_2|}, \qquad \hat{H}_1 \frac{q_e^2}{|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|}.$$

In the atom's ground state the identical electrons occupy the same spatial state, forcing their spin state to be antisymmetric:

$$|1s, 1s\rangle = |1, 0, 0\rangle_1 |1, 0, 0\rangle_2 \frac{1}{\sqrt{2}} \big(|+\mathbf{z}\rangle_1 |-\mathbf{z}\rangle_2 - |-\mathbf{z}\rangle_1 |+\mathbf{z}\rangle_2 \big).$$

The unperturbed energy is known to be $E_{1s, 1s}^{(0)} = 2(-\frac{1}{2}m_e c^2 Z^2 \alpha^2)$. The first-order shift is the integral

$$E_{1s, 1s}^{(1)} = \langle 1s, 1s| \frac{q_e^2}{|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|} |1s, 1s\rangle = \iint d^3 r_1 d^3 r_2 \, |\langle \mathbf{r}_1|1, 0, 0\rangle|^2 \, |\langle \mathbf{r}_2|1, 0, 0\rangle|^2 \, \frac{q_e^2}{|\mathbf{r}_1 - \mathbf{r}_2|},$$

where we have inserted the symmetric spatial state

$$\begin{aligned}
|\psi_S\rangle &= \frac{1}{2} \iint d^3 r_1 d^3 r_2 \left(\frac{1}{\sqrt{2}} |\mathbf{r}_1, \mathbf{r}_2\rangle + \frac{1}{\sqrt{2}} |\mathbf{r}_2, \mathbf{r}_1\rangle \right) \left(\frac{1}{\sqrt{2}} \langle \mathbf{r}_1, \mathbf{r}_2|\psi_S\rangle + \frac{1}{\sqrt{2}} \langle \mathbf{r}_2, \mathbf{r}_1|\psi_S\rangle \right) \\
&= \frac{1}{2} \iint d^3 r_1 d^3 r_2 \big(|\mathbf{r}_1.\mathbf{r}_2\rangle \langle \mathbf{r}_1, \mathbf{r}_2|\psi_S\rangle + |\mathbf{r}_2, \mathbf{r}_1\rangle \langle \mathbf{r}_1, \mathbf{r}_2|\psi_S\rangle \big) \\
&= \iint d^3 r_1 d^3 r_2 |\mathbf{r}_1, \mathbf{r}_2\rangle \langle \mathbf{r}_1, \mathbf{r}_2|\psi_S\rangle.
\end{aligned}$$

We could get a similar integral for $|\psi_A\rangle$. To evaluate, we'd substitute the known $R_{1,0}(r)Y_{0,0}(\theta, \phi)$ from our previous discussion of the hydrogen atom; after doing this we'd find a sizable discrepancy with the experimental value due to the fact that the Coulomb repulsion between the electrons isn't actually that small. We'll look at a more accurate way of determining the energies shortly.

As for the first-excited states, one of the particles is in $|1, 0, 0\rangle$ while the other has $|2, l, m\rangle$; there are four states that fit this description. The unperturbed energy is $E_{1s, 2s \text{ or } 2p} = -\frac{1}{2}m_e c^2 Z^2 \alpha^2(1 + 1/2^2)$. Now, taking advantage of the fact that $\hat{H}_1$ commutes with the exchange operator in both space and spin, we can diagonalize $\hat{H}_1$ by choosing states that are (overall antisymmetric) eigenstates of both $\hat{P}_{12}^{\text{spin}}$ and $\hat{P}_{12}^{\text{space}}$.

We'd end up with a first-order shift of the form $E^{(1)} = J \pm K$, where the $\pm$ is correlated with the value of the total spin so that the overall state is antisymmetric: we have $J - K$ for the triplet states, while the singlet has $J + K$. (This reflects how the antisymmetric wave functions vanish at $\mathbf{r}_1 = \mathbf{r}_2$, causing increased separation due to Coulomb repulsion.)

For atoms with more electrons the Hamiltonian is much less wieldy. We can hone in on what the potential looks like by taking a guess (based on what we know about the limiting values), numerically solving the Schrödinger equation for this guess, determining the charge distribution $q_e |\langle \mathbf{r}_1 | n, l, m \rangle|^2$ corresponding to the resulting wave function, and using that to determine a new guess potential, all in succession until a self-consistent solution materializes. The states fill according to the Pauli exclusion principle.

- The $1s$, $2s$, $2p$, $3s$, and $3p$ states fill like normal.

- The $4s$ level is very slightly below the $3d$ level, so those fill next.

- By the time the $3d$ level is filled with four electrons, the $4s$ level has been raised enough to be slightly above the $3d$ level; chromium therefore has $4s^1\,3d^5$.

- Manganese, iron, cobalt, and nickel have $4s^2\,3d^5$ through $4s^2\,3d^8$, but then again the pattern changes and copper gets $4s^1\,3d^{10}$. (This configuration is so close in energy to $4s^2\,3d^9$, however, that copper may be one or the other depending on its environment.)

- The rest of the row fills like normal.

5.7 The Variational Method

In practice, we estimate the energy levels of helium (and heavier) atoms using the variational method. The idea is to exploit the fact that every state $|\psi\rangle$ is a linear combination of energy eigenstates $|E_n\rangle$, each of which satisfies $E_n \geq E_0$, meaning the expectation value $\langle \psi | \hat{H} | \psi \rangle \geq E_0$.

The method itself is also pretty simple. We begin by picking a trial state $|\psi(\alpha_1, \alpha_2, \dots)\rangle$ which depends on parameters $\alpha_1, \alpha_2, \dots$. We then compute $\langle E \rangle$ for this state and then minimize it with respect to the free parameters to get as close as possible to the ground-state E_0.

When we go to pick a trial state, we generally choose one that closely resembles what we think the actual state should look like. Two common wave functions are the (radial) ground state of the hydrogen atom

$$\psi(r) = \frac{b^{3/2}}{\sqrt{\pi}} e^{-br}$$

and that of the one-dimensional harmonic oscillator

$$\psi(x) = \left(\frac{2a}{\pi} \right)^{1/4} e^{-ax^2}.$$

Note that we could use a similar method to determine higher energy eigenvalues E_n simply by picking a trial state that is orthogonal every eigenstate $|E_0\rangle, \dots, |E_{n-1}\rangle$ below it. We may do this either by simply making a clever choice, or we might construct one; for example, if $|\psi\rangle$ is the trial ground state then we can build an excited trial state via

$$|\varphi\rangle \longrightarrow \frac{|\varphi\rangle - |\psi\rangle\langle\psi|\varphi\rangle}{\sqrt{1 - |\langle\psi|\varphi\rangle|^2}}.$$

However, since $|\psi\rangle$ is only a trial state rather than the true ground state, minimizing $\langle \varphi | \hat{H} | \varphi \rangle$ will not necessarily yield an upper bound for E_1.

5.8 Covalent Bonding

Ionic bonding can be explained classically as the exchange of electrons, but covalent bonding is pure quantum mechanics. Suppose two nuclei are spaced a fixed distance R apart; if we allow this distance to appear as a parameter in our wave functions, we can minimize $\langle E \rangle$ with respect to R to approximate both the energy and size of a molecule held together by a covalent bond.

Consider the hydrogen molecule ion H_2^+. If the protons are spaced far apart, there are clearly at least two possible states: the electron is bound to one proton or the other. The corresponding wave functions are

$$\langle \mathbf{r}|1\rangle = \frac{1}{\sqrt{\pi a_0^3}} e^{-|\mathbf{r}-\mathbf{R}/2|/a_0}, \qquad \langle \mathbf{r}|2\rangle = \frac{1}{\sqrt{\pi a_0^3}} e^{-|\mathbf{r}+\mathbf{R}/2|/a_0},$$

where $\mathbf{r}$ is the position of the electron and $\mathbf{R}/2$ is the position of each proton with respect to their midpoint. The linear combinations of $|1\rangle$ and $|2\rangle$ that diagonalize the Hamiltonian turn out to be

$$|\pm\rangle = \frac{1}{\sqrt{2 \pm 2\langle 1|2\rangle}} (|1\rangle \pm |2\rangle), \qquad E_\pm \frac{1}{1 \pm \langle 1|2\rangle} (H_{11} \pm H_{12}),$$

where H_{11} and H_{12} are the matrix elements of $\hat{H}$ (expressed as integrals). Notice that $\langle \mathbf{r}|+\rangle$ and $\langle \mathbf{r}|-\rangle$ have even and odd parity, respectively. Of course, only the even parity state has a minimum, and so we call $|+\rangle$ a bonding molecular orbital; the state $|-\rangle$ is called an antibonding molecular orbital. Note that these molecular orbitals are linear combinations of atomic orbitals. (Also note that our simple trial wave functions, despite producing pretty good qualitative results, undershoot the binding energy and overshoot the separation.)

Adding another electron into the mix produces H_2, which has a higher binding energy and a smaller inter-nuclear separation. The electrons' spatial state, of course, is symmetric, meaning their spin state has to be antisymmetric in a spin-0 state. Bonding *only* occurs in such a state—in a spin-1 state the spatial state would be antisymmetric, meaning the electrons wouldn't exist in the space between the protons.

This explains why, say, hydrogen and helium do not bind together to form HHe. The electrons in He are already paired up in the spatial state $1s$ and the total-spin-0 state, so if the H electron were to form a covalent bond by pairing up with one of the He electrons it would expend a great deal of energy in kicking the other one up to the $2s$. If the He electrons were instead in a total-spin-1 state we would see repulsion between the two atoms. Helium is therefore inert, along with all the other noble gases.