

## SUPPLEMENTARY NOTES ON DIRAC NOTATION, QUANTUM STATES, ETC.

©R. L. Jaffe, 1996

These notes were prepared by Prof. Jaffe for the 8.05 course which he taught in 1996. In next couple of weeks we will cover all of this material in lecture, though not in as much detail. I am handing them out early so you have an additional source for the material that you can read as we go along, perhaps also filling in some gaps.

There are three main parts.

1. The “Postulates of Quantum Mechanics”.
2. Completeness and orthonormality.
3. An extended example of the use of Dirac notation — position and momentum.

Note that Prof. Jaffe had not yet introduced spin as a two state system at the time he distributed these notes. I recommend that as you read these notes, at every step of the way you think how to apply them to the two state system.

If you are having difficulty with the concepts we have covered in the first part of 8.05 please study these notes carefully. The notes are written using Dirac Notation throughout. One of the purposes is to give you lots of exposure to this notation.

If, after reading these notes, it all still seems confusing, or overly formal, give yourself time. We will study some simple physical examples (especially the harmonic oscillator and the “two state systems” which we have already introduced) where you will learn by doing.

Here, then, are Prof. Jaffe’s notes:

### 1 The Postulates of Quantum Mechanics

I’m not a lover of “postulates”. In general it is better to develop the ideas gradually and let them sink in and become familiar, rather than to announce a set of postulates and try to derive physics from them. In this case, however, we began 8.05 with a wide range of backgrounds, so it is worthwhile to make sure we have a common understanding of how quantum mechanics works in general. Hence this section.

## 1.1 First Postulate

*At each instant the state of a physical system is represented by a ket  $|\psi\rangle$  in the space of states.*

### Comments

- The space of states is a vector space. This postulate is already radical because it implies that the *superposition* of two states is again a state of the system. If  $|\psi_1\rangle$  and  $|\psi_2\rangle$  are possible states of a system, then so is

$$|\psi\rangle = a_1|\psi_1\rangle + a_2|\psi_2\rangle, \quad (1)$$

where  $a_1$  and  $a_2$  are complex numbers. Imagine that  $|\psi_1\rangle$  is a particle with one value for some property like location and  $|\psi_2\rangle$  is the same particle with a different value. In quantum mechanics we must allow ourselves to consider which superpose a particle in different locations. We were forced to do this by the results of experiments like the double-slit diffraction of electrons.

- The space of states comes equipped with the concept of an *inner product* which we abstract from wave mechanics. The inner product associates a complex number to any two states

$$(|\psi\rangle, |\phi\rangle) \equiv \langle\psi|\phi\rangle = \int dx \psi^*(x)\phi(x). \quad (2)$$

Here we have used two different notations. The first defines the inner product as an operation acting on two states in the ket space. The second introduces another copy of the space of states called the “bra space”, and defines the inner product as an operation involving one element of the bra space and one element of the ket space. Either way, the inner product reduces to the integral overlap of the two states when evaluated in terms of wavefunctions —  $\int \psi^* \phi$ . From (2) we see that

$$\langle\psi|\phi\rangle^* = \langle\phi|\psi\rangle. \quad (3)$$

## 1.2 Second Postulate

*Every observable attribute of a physical system is described by an operator that acts on the kets that describe the system.*

### Comments

- By *convention*, an operator  $\hat{A}$  acting on a ket  $|\psi\rangle$  is denoted by left multiplication,

$$\hat{A} : |\psi\rangle \rightarrow |\psi'\rangle = \hat{A}|\psi\rangle. \quad (4)$$

You are used to this concept in the context of wave-mechanics, where the concept of a *state* is replaced by that of a *wavefunction*. A system (a particle in a potential, for example) is described by a wavefunction  $\psi(x)$  in wave-mechanics. Some simple observable attributes of such a system are its *position*, its *momentum* and its *energy*. These are represented in wave mechanics by *differential operators*,  $\hat{X} = x$ ,  $\hat{P} = -i\hbar\frac{d}{dx}$  and  $\hat{\mathcal{H}} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$  respectively. These operators act on a wavefunction by left-multiplication, like

$$\hat{P}\psi(x) = -i\hbar\frac{d\psi}{dx} \quad (5)$$

- It is important to recognize that acting with an operator on a state in general changes the state. Again think back to wave mechanics. The lowest energy eigenfunction in a square well ( $0 \leq x \leq L$ ) is

$$\psi(x) = \sqrt{2/L} \sin \pi x/L \quad \text{for } 0 \leq x \leq L. \quad (6)$$

When we act on this wavefunction with  $\hat{P}$ , for example, we get

$$\hat{P}\psi(x) = -i\hbar\pi/L\sqrt{2/L} \cos \pi x/L \quad (7)$$

which is no longer an energy eigenfunction at all. So the operator changed the state of the particle.

- For every operator, there are special states that are not changed (except for being multiplied by a constant) by the action of an operator,

$$\hat{\mathcal{A}}|\psi_a\rangle = a|\psi_a\rangle. \quad (8)$$

These are the *eigenstates* and the numbers  $a$  are the *eigenvalues* of the operator. You have encountered them in wave mechanics, now they show up in the abstract space of states.

### 1.3 Third Postulate

*The only possible result of the measurement of an observable  $\mathcal{A}$  is one of the eigenvalues of the corresponding operator  $\hat{\mathcal{A}}$ .*

#### Comments

- This is, of course, the origin of the word “quantum” in quantum mechanics. If the observable has a continuous spectrum of eigenvalues, like the position  $x$  or the momentum  $p$ , then the statement is not surprising. If it has a discrete spectrum, like the Hamiltonian for an electron bound to a proton (the hydrogen atom), then the statement is shocking. A measurement of the energy of the hydrogen atom will yield

only one of a discrete set of values. Needless to say, this postulate reflects mountains of experimental evidence such as the discrete *spectral lines* observed in the radiation from a tube of hot hydrogen gas.

- Since we measure only real numbers, the eigenvalues of operators corresponding to observables had better be real. Operators with real eigenvalues are *hermitian*.

The eigenstates of a hermitian operator have some important properties.

- They are orthogonal

$$\langle a_j | a_k \rangle \equiv (a_j, a_k) = \int dx \psi_{a_j}^*(x) \psi_{a_k}(x) = \delta_{jk}. \quad (9)$$

- They span the space of states, so they form a *basis*. This means that an arbitrary state can be expanded as a sum (with complex coefficients) of the eigenstates of a hermitian operator. For this reason we say that the set of states is “complete”.

## 1.4 Fourth Postulate

When a measurement of an observable  $\mathcal{A}$  is made on a generic state  $|\psi\rangle$ , the probability of obtaining an eigenvalue  $a_n$  is given by the square of the inner product of  $|\psi\rangle$  with the eigenstate  $|a_n\rangle$ ,  $|\langle a_n | \psi \rangle|^2$ .

### Comments

- The states are assumed to be normalized. Usually we normalize our states to unity,

$$\begin{aligned} \langle \psi | \psi \rangle &= 1 \\ \langle a_j | a_k \rangle &= \delta_{jk}. \end{aligned} \quad (10)$$

Sometimes this is not possible. The case of momentum eigenstates,  $\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} \exp ipx/\hbar$ , is the classic example. In this case we must use “ $\delta$ -function” or “continuum” normalization as discussed in Section 2.

- The complex number,  $\langle a_n | \psi \rangle$  is known as the “probability amplitude” or “amplitude”, for short, to measure  $a_n$  as the value for  $\mathcal{A}$  in the state  $|\psi\rangle$ .
- Here is the algebraic exercise suggested by this postulate. First, any state can be expanded as a superposition of  $\mathcal{A}$ -eigenstates (see Post. 3),

$$|\psi\rangle = \sum_n c_n |a_n\rangle. \quad (11)$$

Next use the orthonormality of the  $\mathcal{A}$  eigenstates to find an expression for the expansion coefficients  $c_n$ ,

$$\begin{aligned}\langle a_j | \psi \rangle &= \sum_n c_n \langle a_j | a_n \rangle \\ &= c_j.\end{aligned}\tag{12}$$

So,

$$|\psi\rangle = \sum_n \langle a_n | \psi \rangle \cdot |a_n\rangle.\tag{13}$$

The  $\cdot$  is added just to make clear the separation between the complex number  $\langle a_n | \psi \rangle$  and the state  $|a_n\rangle$ . So, the component of  $|\psi\rangle$  along the “direction” of the  $n^{\text{th}}$  eigenstate of  $\mathcal{A}$  is given by  $\langle a_n | \psi \rangle$ . The measurement operation yields the result  $a_n$  with a probability proportional to the square of this component,  $|\langle a_n | \psi \rangle|^2$ .

- The probability of obtaining *some result* is unity. For states normalized to unity,

$$|\langle \psi | \psi \rangle|^2 = \sum_m \sum_n c_m^* c_n \langle a_m | a_n \rangle.\tag{14}$$

Using  $|\langle \psi | \psi \rangle| = 1$  and  $\langle a_m | a_n \rangle = \delta_{mn}$ , we get

$$\sum_n |c_n|^2 = 1\tag{15}$$

- According to the usual rules of probability, we can compute the “expected value” of the observable  $\mathcal{A}$ . If the probability to observe  $a_n$  is  $|c_n|^2$  then the expected value (denoted  $\langle \mathcal{A} \rangle$ ) is

$$\langle \mathcal{A} \rangle = \sum_n a_n |c_n|^2.\tag{16}$$

- When there is more than one eigenstate with the same eigenvalue, then this discussion needs a little bit of refinement. We’ll let this go until we need to confront it.

## 1.5 Fifth Postulate

*Immediately after the measurement of an observable  $\mathcal{A}$  has yielded a value  $a_n$ , the state of the system is the normalized eigenstate  $|a_n\rangle$ .*

### Comments

- Known picturesquely as the “collapse of the wavepacket”, this is the most controversial of the postulates of quantum mechanics, and the most difficult to get comfortable with. It is motivated by experience with repeated measurements. If an experimental sample is prepared in a state  $|\psi\rangle$  then it is observed that a measurement of  $\mathcal{A}$  can

yield a variety of results  $a_n$  with probabilities  $|\langle a_n | \psi \rangle|^2$ . Identically prepared systems can yield different experimental outcomes. This is encompassed by the fourth postulate. However, if  $\mathcal{A}$  is measured with outcome  $a_n$  on a given system, and then is immediately *remeasured*, the results of the second measurement *are not statistically distributed*, the result is always  $a_n$  again. Hence this postulate.

- The collapse of the wavepacket preserves the normalization of the state. If  $|\psi\rangle$  and  $|a_n\rangle$  are both normalized to unity, then the measurement process replaces  $|\psi\rangle$  by  $|a_n\rangle$ , *not by*  $|\langle a_n | \psi \rangle|^2 \cdot |a_n\rangle$ .

## 1.6 Sixth Postulate

*The time evolution of a quantum system preserves the normalization of the associated ket. The time evolution of the state of a quantum system is described by  $|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle$ , for some unitary operator  $\hat{U}$ .*

### Comments

- We have not got to this subject yet. I include it for completeness.
- Under time evolution, a state  $|\psi\rangle$  moves through the space of states on a trajectory we can define as  $|\psi(t)\rangle$ . The preservation of the norm of the state is associated with conservation of probability. If the observable  $A$  is energy, for example, then the statement (15) says that the probability to find the system with *some value for the energy* is unity when summed over all possible values. For this to remain true as time goes on, it is necessary for the norm of the state to stay unity.
- Soon we will show that this postulate requires  $|\psi\rangle$  to obey a differential equation of the form

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \mathcal{H} |\psi(t)\rangle \quad (17)$$

where  $\mathcal{H}$  is a hermitian operator (we know it as the Hamiltonian). This is Schroedinger's equation written as an operator equation in the space of states (as opposed to a differential equation in the space of wavefunction).

## 2 Bases and Operators

In this section I want to go over two topics not treated in sufficient detail in lecture. First, more about orthonormality and completeness, and second, some of the properties of hermitian operators.

## 2.1 Orthonormality and Completeness

We are working in the space of states  $\mathcal{S}$  of a quantum system. The states are denoted by kets,  $|\psi\rangle$ , or equivalently by bras  $\langle\psi|$ . We will assume from the start that it is possible to find a basis for this space. The process for finding a basis is discussed in courses in linear algebra. Simply put, one chooses a state, and includes it in the basis. Next one chooses another. If it is proportional to the first, one discards it. If not, one puts it in the basis. On to another. If it is a linear combination of the first two, it is not “new” and is discarded. If it is not a linear combination of the first two, it must be added to the basis. This continues until one has the largest possible set of linearly independent states, which is the basis.

We will have to deal with two different types of bases:

- Denumerable

*Denumerable*  $\equiv$  *countable*, meaning the states can be put into one-to-one correspondence with the integers. We can denote them by  $\{|n\rangle\}$ , where  $n = 1, 2, 3, \dots$ . The energy eigenstates of the one-dimensional harmonic oscillator are a good example.

- Non-denumerable

Here the elements of the basis cannot be counted. Instead they are denoted  $|z\rangle$ , labelled by a continuous variable,  $z$ , that ranges from  $-\infty < z < \infty$ . The eigenstates of position and momentum are the classic examples.

Of course there are other cases, where the basis requires two integer labels, or two continuous labels, or one of each. These are handled by simple generalizations of the analysis given here.

Our space of states has an inner product,  $\langle\psi|\phi\rangle$ . Using the inner product it is always possible to generate an *orthonormal* basis from one originally possessing no special properties with respect to the inner product. The process of making a basis orthonormal is called the “Schmidt orthogonalization process”. It is explained in courses on linear algebra. I am going to skip it because we will see that the eigenstates of hermitian operators are automatically orthogonal. All we need to do to make them into an orthonormal basis is to divide out their length:  $|\psi\rangle \rightarrow \sqrt{\langle\psi|\psi\rangle}|\psi\rangle$ , so  $\langle\psi|\psi\rangle = 1$ .

I need to say how “orthonormality” is defined. If we were dealing with a finite dimensional vector space (like Euclidean 3-space), then orthonormality means that the inner product of two different basis elements is zero and the inner product of a basis vector with itself is unity:  $\vec{v}_j \cdot \vec{v}_k = \delta_{jk}$ . This generalizes directly to the case of a *denumerable basis*

$$\langle m|n\rangle = \delta_{mn} \tag{18}$$

but not to a non-denumerable basis. The symbol  $\delta_{zz'}$  doesn't make sense — how closely equal do  $z$  and  $z'$  have to be before  $\delta$  takes on the value 1? For a *non-denumerable* basis, we *postulate* an orthonormality relation more suited to continuous variables,

$$\langle z|z'\rangle = \delta(z - z') \tag{19}$$

where  $\delta(z)$  is Dirac's  $\delta$ -function. This is known as "continuum normalization". My hope is that you'll accept this as a candidate for an orthonormality condition and pursue its consequences with me.

Anyone who has studied Fourier Series and Fourier Integrals will recognize these two different orthonormality conditions. The independent *sine* functions,  $\sqrt{\frac{2}{L}} \sin n\pi x/L$ , that enter into Fourier Series satisfy an orthonormality condition like (18), where the inner product is defined as the integral from 0 to  $L$ . On the other hand, the independent functions that appear in Fourier Integrals,  $\frac{1}{\sqrt{2\pi}} \exp iqx$ , obey an orthonormality condition with a  $\delta$ -function like (19).

There is a fancy way to state orthonormality that is very useful in quantum mechanics. First let's develop it in the case of denumerable bases. Take an arbitrary ket and expand it in the orthonormal basis  $\{|n\rangle\}$ ,

$$|\psi\rangle = \sum_n c_n |n\rangle. \tag{20}$$

Using the orthonormality of the basis states we obtain an equation for the expansion coefficients,

$$c_n = \langle n|\psi\rangle, \tag{21}$$

and substitute back into (20), to obtain

$$|\psi\rangle = \sum_n \langle n|\psi\rangle \cdot |n\rangle. \tag{22}$$

The "dot" is again added to the equation explicitly to remind us that each term in the sum is the product of a complex number (*c-number*)  $\langle n|\psi\rangle$  and a ket  $|n\rangle$ .

Now we do some Dirac trickery, by rearranging the terms in (22) and separating them in a suggestive manner,

$$\begin{aligned} |\psi\rangle &= \sum_n |n\rangle \cdot \langle n|\psi\rangle \\ &= \left[ \sum_n |n\rangle \langle n| \right] |\psi\rangle. \end{aligned} \tag{23}$$

In the last version the quantity in square-brackets plays the role of *the identity operator*,

$$1 = \sum_n |n\rangle \langle n| \tag{24}$$

Acting on the state  $|\psi\rangle$ , it gives  $|\psi\rangle$  back again. Also it is placed properly (acting by left multiplication) for an operator in the ket space.

Perhaps it will help to make the analogy to ordinary vectors in 3-space. Expanding an arbitrary vector in a Cartesian basis (call the basis vectors  $\hat{e}_j$ ,  $j = 1, 2, 3$  and keep track



of transposes) we get

$$\begin{aligned}\vec{v} &= \sum_j (\hat{e}_j^T \cdot \vec{v}) \hat{e}_j \\ &= \left[ \sum_j \hat{e}_j \hat{e}_j^T \right] \cdot \vec{v}.\end{aligned}\tag{25}$$

So the quantity  $\sum_j \hat{e}_j \hat{e}_j^T$  seems to play the role of the identity. Well, if we write

$$\hat{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \hat{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \hat{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\tag{26}$$

and explicitly construct the  $3 \times 3$  matrix  $\sum_j \hat{e}_j \hat{e}_j^T$ , it is easy to see it is the *unit matrix*. Our result, (24) is the extension of this result to a denumerably infinite dimensional space.

Finally I want to generalize (24) to non-denumerable bases. Since the label for the basis states is continuous, *we must integrate over it* in the analog of (20)

$$|\psi\rangle = \int_{-\infty}^{\infty} dz c(z) |z\rangle.\tag{27}$$

Note that the *function*  $c(z)$  has taken the place of the expansion coefficients  $c_n$ . To find the analog of (21), we take the inner product with  $\langle z'|$ ,

$$\langle z'|\psi\rangle = \int_{-\infty}^{\infty} dz c(z) \langle z'|z\rangle,\tag{28}$$

use  $\langle z'|z\rangle = \delta(z - z')$  and perform the integral over  $z$ , leaving

$$c(z) = \langle z|\psi\rangle\tag{29}$$

just like (21). Now substitute back into (27) and rearrange the terms as we did in the denumerable case,

$$|\psi\rangle = \left[ \int_{-\infty}^{\infty} dz |z\rangle \langle z| \right] |\psi\rangle\tag{30}$$

from which we extract

$$1 = \int_{-\infty}^{\infty} dz |z\rangle \langle z|\tag{31}$$

## 2.2 Operators and Hermitian Conjugation

Here I want to derive some of the properties of hermitian operators. First, however, lets review the derivation of hermitian conjugation.

### 2.2.1 Hermitian Conjugation

If  $\hat{\mathcal{A}}$  is an operator defined by its action on the kets,

$$\hat{\mathcal{A}} : |\psi\rangle \rightarrow |\psi'\rangle = \hat{\mathcal{A}}|\psi\rangle, \quad (32)$$

then the hermitian conjugate of  $\hat{\mathcal{A}}$ , denoted  $\hat{\mathcal{A}}^\dagger$ , is defined to be the operator that has the same action on the bras,

$$\hat{\mathcal{A}}^\dagger : \langle\psi| \rightarrow \langle\psi'| = \langle\psi|\hat{\mathcal{A}}^\dagger, \quad (33)$$

The quantity  $\hat{\mathcal{A}}|\psi\rangle$  is again a state in  $\mathcal{S}$ . To emphasize that fact we could put it into a ket notation,

$$|\hat{\mathcal{A}}\psi\rangle = |\hat{\mathcal{A}}\psi\rangle \quad (34)$$

The equivalent statement for the bra space is

$$\langle\psi|\hat{\mathcal{A}}^\dagger = \langle\hat{\mathcal{A}}\psi| = \langle\psi'|. \quad (35)$$

Note that we *do not* denote the bra by  $\langle\hat{\mathcal{A}}^\dagger\psi|$ , because it is the state with the same attributes as  $|\psi'\rangle = |\hat{\mathcal{A}}\psi\rangle$ .

Consider the inner product of  $|\psi'\rangle$  defined in (32) with some arbitrary state  $|\phi\rangle$

$$\langle\phi|\psi'\rangle \equiv \langle\phi|\hat{\mathcal{A}}|\psi\rangle \quad (36)$$

This defines a new kind of expression, with an operator sandwiched between a bra and a ket. Think of it as follows: when  $\hat{\mathcal{A}}$  operates on  $|\psi\rangle$ , it creates some ket which one can overlap with  $|\phi\rangle$ . Completely equivalently, one can view  $\hat{\mathcal{A}}$  as an operator on the bra space, transforming the bra  $\langle\phi|$  to a new element of the bra space,  $\langle\phi|\hat{\mathcal{A}}$ , which then overlaps with  $|\psi\rangle$ . The notation defined in (36) with  $\hat{\mathcal{A}}$  between bra and ket includes both points of view and is the one we'll generally use.

Quantities of the form  $\langle\phi|\hat{\mathcal{A}}|\psi\rangle$  are called “matrix elements”. They are *c-numbers* and measure the capacity of the operator  $\hat{\mathcal{A}}$  to provide overlap between the two states.

Complex conjugation of matrix elements involves hermitian conjugation of operators. Consider

$$\langle\phi|\hat{\mathcal{A}}|\psi\rangle^* = \langle\phi|\psi'\rangle^* \quad (37)$$

Using the complex conjugation property of the inner product,  $\langle\phi|\psi'\rangle^* = \langle\psi'|\phi\rangle$ , and (35) we find

$$\langle\phi|\hat{\mathcal{A}}|\psi\rangle^* = \langle\psi|\hat{\mathcal{A}}^\dagger|\phi\rangle \quad (38)$$

If the operator  $\hat{\mathcal{A}}$  is hermitian, then its matrix elements have a simple behavior: complex conjugation is equivalent to exchanging the bra and the ket.

### 2.2.2 Eigenvalues and eigenstates of a hermitian operator

Here we show that the eigenvalues of a hermitian operator are real and that the eigenstates of a hermitian operator form an orthonormal set. Let's define the eigenkets and eigenvalues of some operator  $\hat{\mathcal{A}}$  in the usual fashion

$$\hat{\mathcal{A}}|a_k\rangle = a_k|a_k\rangle \quad (39)$$

To be definite, we have assumed the eigenvalues are countable, and we've labelled the states by their eigenvalues.

From the definition of  $\hat{\mathcal{A}}^\dagger$  we see that  $\langle a_k|$  is an eigenbra of  $\hat{\mathcal{A}}^\dagger$  with eigenvalue  $a_k^*$ ,

$$\langle a_k|\hat{\mathcal{A}}^\dagger = a_k^*\langle a_k| \quad (40)$$

Let us restrict ourselves to hermitian operators,  $\hat{\mathcal{A}}^\dagger = \hat{\mathcal{A}}$ . Then take the inner product of (39) with the state  $|a_k\rangle$  to obtain,

$$\langle a_k|\hat{\mathcal{A}}|a_k\rangle = a_k\langle a_k|a_k\rangle. \quad (41)$$

If  $\hat{\mathcal{A}}$  is hermitian, then the matrix element  $\langle a_k|\hat{\mathcal{A}}|a_k\rangle$  is real, as is the norm  $\langle a_k|a_k\rangle$ . So if we take the complex conjugate of (41) we conclude that *the eigenvalues of a hermitian operator are real*.

Next take the inner product of (39) with an eigenstate belonging to a different eigenvalue of  $\hat{\mathcal{A}}$ , say  $a_j$ ,

$$\langle a_j|\hat{\mathcal{A}}|a_k\rangle = a_k\langle a_j|a_k\rangle. \quad (42)$$

Next write the eigenvalue equation for the eigenbra  $\langle a_j|$ ,

$$\langle a_j|\hat{\mathcal{A}} = a_j\langle a_j|, \quad (43)$$

where we used both that  $\hat{\mathcal{A}}$  is hermitian, and that  $a_j$  is real. Finally take the inner product of (43) with the eigenket  $|a_k\rangle$ ,

$$\langle a_j|\hat{\mathcal{A}}|a_k\rangle = a_j\langle a_j|a_k\rangle. \quad (44)$$

If we compare (42) with (44) (subtract the two equations), we see that

$$\langle a_j|a_k\rangle = 0 \quad \text{if} \quad a_j \neq a_k. \quad (45)$$

So the eigenstates belonging to different eigenvalues of a hermitian operator are necessarily orthogonal. Next, choose the norm of the eigenstates to be unity and we have an orthonormal set of eigenstates for every hermitian operator.

There is one lacuna in this derivation. If two eigenstates share the same eigenvalue (the eigenvalue is said to be *degenerate*), we have not shown that the eigenstates are orthogonal. The flaw is fairly simple to repair: it is always possible to choose linear combinations of the

set of states belonging to the same eigenvalue so that they are orthogonal to one another, and of course any linear combination is still orthogonal to the eigenstates corresponding to different eigenvalues.

It is more difficult to show that the set of orthonormal eigenvectors of a hermitian operator is complete. It mirrors the completeness proof for Fourier Series or Fourier Integrals, which you may have encountered in other courses. Accepting completeness, we see that hermitian operators generate complete orthonormal bases for our space of states.

### 2.2.3 Operators as matrices in an orthonormal basis

The objects of interest in quantum mechanics are matrix elements of the form

$$\langle \phi | \hat{A} | \psi \rangle. \quad (46)$$

If we make use of an orthonormal basis we can reduce the manipulation of matrix elements to matrix algebra. We consider denumerable and non-denumerable bases in turn.

Suppose the space of states has a denumerable basis,  $\{|n\rangle\}$ . We may expand both  $|\psi\rangle$  and  $|\phi\rangle$  in this basis,

$$\begin{aligned} |\psi\rangle &= \sum_n \langle n | \psi \rangle \cdot |n\rangle \\ |\phi\rangle &= \sum_m \langle m | \phi \rangle \cdot |m\rangle \end{aligned} \quad (47)$$

and substitute in (46).

$$\langle \phi | \hat{A} | \psi \rangle = \sum_n \sum_m \langle \phi | m \rangle \cdot \langle m | \hat{A} | n \rangle \cdot \langle n | \psi \rangle \quad (48)$$

a double sum over products of three complex numbers. To make (48) simpler in appearance, we define

$$\begin{aligned} \psi_n &\equiv \langle n | \psi \rangle \\ \phi_m &\equiv \langle m | \phi \rangle \\ A_{mn} &\equiv \langle m | \hat{A} | n \rangle \end{aligned} \quad (49)$$

so (48) now reads

$$\langle \phi | \hat{A} | \psi \rangle = \sum_m \sum_n \phi_m^* A_{mn} \psi_n \quad (50)$$

or

$$\langle \phi | \hat{A} | \psi \rangle = \phi^\dagger A \psi \quad (51)$$

where  $A$  is the matrix of complex numbers whose  $mn^{th}$  component is  $A_{mn}$ ,  $\psi$  is the column vector whose  $n^{th}$  component is  $\psi_n$ , and  $\phi^\dagger$  is the row vector whose  $m^{th}$  component is  $\phi_m^*$ . This is the origin of the term ‘‘Matrix Mechanics’’ that was originally applied to this

formulation of quantum mechanics. States are represented by vectors and operators by matrices.

Now consider the case of a non-denumerable basis  $|z\rangle$ . We start with (46). The expansions given in (47) are replaced by integrals

$$\begin{aligned} |\psi\rangle &= \int_{-\infty}^{\infty} dz \langle z|\psi\rangle \cdot |z\rangle \\ |\phi\rangle &= \int_{-\infty}^{\infty} dz \langle z|\phi\rangle \cdot |z\rangle. \end{aligned} \tag{52}$$

The *c-number*  $\langle z|\psi\rangle$  is a *function* of the parameter  $z$ , so we denote it

$$\langle z|\psi\rangle = \psi(z). \tag{53}$$

When we substitute these decompositions of the states into (46) we will encounter a matrix element of the form  $\langle z|\hat{\mathcal{A}}|z'\rangle$ , which is a *c-number* function of the two variables  $z$  and  $z'$ , so we denote it,

$$\langle z|\hat{\mathcal{A}}|z'\rangle = A(z, z'). \tag{54}$$

Putting this all together in the case of a non-denumerable basis we obtain,

$$\langle \phi|\hat{\mathcal{A}}|\psi\rangle = \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} dz' \phi^*(z) A(z, z') \psi(z') \tag{55}$$

Vectors generalize to functions, matrices generalize to functions of two variables, and matrix multiplication generalizes to integration. We still talk about the states as vectors and the operators as “matrix elements” even in the case of non-denumerable bases. Next we’ll go through the example of position and momentum eigenstates in an attempt to put all this together.

### 3 Position and Momentum in Dirac Notation

We are now ready to apply all this formalism to some physically interesting cases. The simplest, and most important, place to start is with the operators that dominate wave mechanics: position,  $x$ , and momentum  $p$ . [For simplicity we consider only one space dimension, the generalization to three dimensions is obvious.]

On the basis of everyday experience with the mechanics of classical systems, we *assume* the existence of two observables, position and momentum. We assume that measurement of position or momentum yield values from the continuum of real numbers.<sup>1</sup> Given the discussion of the previous two sections, we can assume the existence of

---

<sup>1</sup>Note: we could be wrong! Sometimes the quantum world is fundamentally unlike the classical one. There are cases where, for example, the momentum of a quantum particle is quantized. But until we encounter a disagreement with experiment, we will follow the natural path.

- Two hermitian operators,  $\hat{X}$  and  $\hat{P}$
- Two complete, orthonormal sets of eigenkets,  $|x\rangle$ , and  $|p\rangle$ , obeying

$$\begin{aligned}
\hat{X}|x\rangle &= x|x\rangle \\
\hat{P}|p\rangle &= p|p\rangle \\
\langle x|y\rangle &= \delta(x-y) \\
\langle p|q\rangle &= \delta(p-q)
\end{aligned}
\tag{56}$$

with  $x$  and  $p$  both real.

Notice that the existence of the eigenstates, and the reality of the eigenvalues follow from the postulates of quantum mechanics. The orthonormality condition is the one we choose for non-denumerable bases.

So far, we have not given any information defining a relation between  $x$  and  $p$ . We would have written down the same statements about two coordinates,  $x_1$  and  $x_2$ , for example. The standard thing to do to connect  $x$  and  $p$  would be to postulate the *commutator* between  $X_{op}$  and  $P_{op}$ , but the reason for postulating  $[\hat{X}, \hat{P}] = i\hbar$  won't be clear until we talk about "Canonical Quantization" in a few weeks. Instead I propose we rely on our knowledge of wave mechanics to tell us the *position space wavefunction for a momentum eigenstate*, namely,

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar},
\tag{57}$$

which embodies the observation that momentum eigenstates propagate like plane waves.

First let's check the normalization of (57). This will give us an opportunity to use the completeness relation for non-denumerable states. We demand  $\langle q|p\rangle = \delta(q-p)$ . We want to convert this to an expression in terms of coordinate space wavefunctions, so we insert a complete set of coordinate eigenstates,

$$\begin{aligned}
\langle q|p\rangle &= \langle q|1|p\rangle \\
&= \int_{-\infty}^{\infty} dx \langle q|x\rangle \cdot \langle x|p\rangle \\
&= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx e^{i(p-q)x/\hbar} \\
&= \delta(q-p)
\end{aligned}
\tag{58}$$

where the last step uses the Fourier representation for the  $\delta$ -function (see the Appendix of Gasiorowicz).

Next let's look at coordinate space wavefunctions and the way the position and momentum operators are represented in coordinate space. An arbitrary state  $|\psi\rangle$  can be expressed as a superposition of coordinate eigenstates,

$$|\psi\rangle = \int_{-\infty}^{\infty} dx |x\rangle \langle x|\psi\rangle = \int_{-\infty}^{\infty} dx \langle x|\psi\rangle \cdot |x\rangle = \int_{-\infty}^{\infty} dx \psi(x) |x\rangle,
\tag{59}$$

where we have defined  $\langle x|\psi\rangle = \psi(x)$ . Note this is the *standard* definition of the coordinate space wavefunction of the that  $|\psi\rangle$ , since  $\langle x|\psi\rangle$  is the amplitude to find the state  $|\psi\rangle$  at the position  $x$ .

How does  $\hat{X}$  act on  $|\psi\rangle$ ? Using (59) and working slowly step by step we easily find

$$\begin{aligned}\hat{X}|\psi\rangle &= \hat{X} \int_{-\infty}^{\infty} dy |y\rangle \langle y|\psi\rangle \\ &= \int_{-\infty}^{\infty} dy \hat{X}|y\rangle \langle y|\psi\rangle \\ &= \int_{-\infty}^{\infty} dy y |y\rangle \langle y|\psi\rangle \\ &= \int_{-\infty}^{\infty} dy y \psi(y) |y\rangle .\end{aligned}\tag{60}$$

So,

$$\langle x|\hat{X}|\psi\rangle = \int_{-\infty}^{\infty} dy \psi(y) \langle x|y\rangle = \int_{-\infty}^{\infty} dy \psi(y) \delta(x-y) = x \psi(x) ,\tag{61}$$

and we recover the familiar fact that when  $\hat{X}$  acts on the state  $|\psi\rangle$  whose coordinate space wave function is  $\psi(x)$  the resulting state has coordinate space wave function is  $x\psi(x)$ .

How does  $\hat{P}$  act on  $|\psi\rangle$ ? To answer this, let's find the coordinate space wavefunction of the state  $|\Psi\rangle \equiv \hat{P}|\psi\rangle$ . As a first step, we look at a momentum eigenstate,

$$\begin{aligned}\langle x|\hat{P}|p\rangle &= p \langle x|p\rangle \\ &= p \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \\ &= -i\hbar \frac{d}{dx} \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \\ &= -i\hbar \frac{d}{dx} \langle x|p\rangle.\end{aligned}\tag{62}$$

So the coordinate space wave function of the state  $\hat{P}|\psi\rangle$  is  $-i\hbar$  times the derivative of the wavefunction of  $|\psi\rangle$ . Now let's look at an arbitrary state. Define  $|\Psi\rangle \equiv \hat{P}|\psi\rangle$ . Then,

$$\begin{aligned}\langle x|\Psi\rangle &= \langle x|\hat{P}|\psi\rangle \\ &= \int_{-\infty}^{\infty} dp \langle x|\hat{P}|p\rangle \cdot \langle p|\psi\rangle \\ &= \int_{-\infty}^{\infty} dp \left(-i\hbar \frac{d}{dx} \langle x|p\rangle\right) \cdot \langle p|\psi\rangle \\ &= -i\hbar \frac{d}{dx} \langle x| \left[ \int_{-\infty}^{\infty} dp |p\rangle \langle p| \right] |\psi\rangle \\ &= -i\hbar \frac{d}{dx} \psi(x),\end{aligned}\tag{63}$$

where we have liberally used the completeness relation for momentum eigenstates,  $\int dp |p\rangle\langle p| = 1$ . So we conclude

$$\hat{P}|\psi\rangle = |\Psi\rangle \Rightarrow \Psi(x) = -i\hbar \frac{d\psi}{dx}, \quad (64)$$

a standard result from wavemechanics. So we say that the “coordinate space representation of the momentum operator” is  $-i\hbar \frac{d}{dx}$ . If you want an exercise to check your understanding, try to show that the momentum space representation of the position operator is  $i\hbar \frac{d}{dp}$ .

Finally let’s look at the coordinate space matrix elements of the position and momentum operators,  $\langle x|\hat{X}|y\rangle$  and  $\langle x|\hat{P}|y\rangle$ . The first is easy,

$$\begin{aligned} \langle x|\hat{X}|y\rangle &= y\langle x|y\rangle \\ &= y\delta(x-y). \end{aligned} \quad (65)$$

$\hat{X}$  is a “local” operator in position space; it connects states only if they have the same eigenvalue. This is the continuum analog of a diagonal matrix: if we view  $\langle x|\hat{X}|y\rangle \equiv f(x, y)$  as a matrix, only the diagonal ( $x = y$ ) elements are non-zero.

The momentum operator is a little harder. Start with  $\langle x|\hat{P}|\psi\rangle$  from (63)

$$\langle x|\hat{P}|\psi\rangle = -i\hbar \frac{d}{dx}\psi(x) \quad (66)$$

and insert a complete set of position eigenstates after  $\hat{P}$ ,

$$\begin{aligned} \langle x|\hat{P}|\psi\rangle &= \langle x|\hat{P} \int_{-\infty}^{\infty} dy |y\rangle\langle y|\psi\rangle \\ &= \int_{-\infty}^{\infty} dy \langle x|\hat{P}|y\rangle \cdot \langle y|\psi\rangle \\ &= \int_{-\infty}^{\infty} dy \langle x|\hat{P}|y\rangle \psi(y). \end{aligned} \quad (67)$$

Comparing (66) and (67) we conclude

$$\langle x|\hat{P}|y\rangle = -i\hbar \delta(x-y) \frac{d}{dy}, \quad (68)$$

which you can check by substitution. So the momentum operator is almost, but not quite, local in coordinate space. The derivative shows that acting with  $\hat{P}$  involves two infinitesimally nearby points in coordinate space.

Just as an arbitrary matrix is not diagonal, so an arbitrary operator matrix element in coordinate space is not local. If  $\hat{T}$  is some unspecified operator, then acting on an arbitrary state  $|\psi\rangle$ , we find,

$$\begin{aligned} \langle x|\hat{T}|\psi\rangle &\equiv \psi_{\mathcal{T}}(x) \\ &= \int_{-\infty}^{\infty} dy \langle x|\hat{T}|y\rangle \cdot \langle y|\psi\rangle \\ \psi_{\mathcal{T}}(x) &= \int_{-\infty}^{\infty} dy T(x, y) \psi(y). \end{aligned} \quad (69)$$



So the operator  $\hat{\mathcal{T}}$  gets represented by its coordinate space matrix elements  $T(x, y) = \langle x | \hat{\mathcal{T}} | y \rangle$ , which acts as an *integral transform* on the wavefunction  $\psi(x)$ . Only when  $\hat{\mathcal{T}}$  is local (like  $\hat{X}$ ) or nearly local (like  $\hat{P}$ ) does the integral go away leaving a simpler situation.

## 7 Commutators, Measurement and The Uncertainty Principle

*A black cat went past us, and then another that looked just like it.*

---

Neo

In this section, we return from the wilderness of solving differential equations to more formal mathematics. In particular, we want to study the notion of measurement, and **simultaneous measurements of observables** that we alluded to way back in the introduction. There we have been careful to say that we cannot measure with arbitrary accuracy the position *and* momentum of a particle at the same time. Then when we study the Gaussian Wavepacket in section 5.4, we saw that the product of the dispersions of its position and momentum has the minimum value

$$(\Delta x)^2(\Delta p)^2 \geq \frac{\hbar^2}{4}, \quad (298)$$

which we proceed to argue should be interpreted as our inability to measure  $p$  and  $x$  to arbitrary accuracy at the same time. In this section, we will show that non-commuting observables will lead to the Heisenberg Uncertainty Principle – one of the pillars of Quantum Mechanics.

### 7.1 Pure Ensembles and Expectation Values

Postulate 3 tells us that the measurement of an observable  $\hat{O}$  in some state  $\psi = \sum_{n=1}^{\infty} a_n u_n$  yields the eigenvalue  $\lambda_n$  with some probability  $|a_n|^2$ . The state then collapses into  $u_n$ . This is all fine and good in theory, the question is: how do we test for this fact?

The way to do this, is to make *many repeated measurements of identically prepared states*, and plot out a histogram of the results, e.g. we measure  $\lambda_1$  6 times,  $\lambda_2$  32 times,  $\lambda_3$  8 times etc. And then compare this to our theoretical prediction. Of course the more identically prepared states there are, the better our experiment will test the theoretical prediction. Such a set of identically prepared states is called a **Pure Ensemble**.

Given a pure ensemble, and a set of measurements, we can also ask what is the *average* value of all the the measured eigenvalues. In the limit of a very large number of measurements, this is called the **Expectation Value**, which is defined to be

$$\langle O \rangle_{\psi} = \psi(\mathbf{x}) \cdot (\hat{O}\psi(\mathbf{x})) = \int_{\mathbb{R}^3} \psi^{\dagger}(\mathbf{x}) \hat{O}\psi(\mathbf{x}) dV. \quad (299)$$

We can show that this exactly is the average value of the measured eigenvalues

$$\begin{aligned} \int_{\mathbb{R}^3} \psi^{\dagger}(\mathbf{x}) \hat{O}\psi(\mathbf{x}) dV &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \int_{\mathbb{R}^3} a_m^* a_n u_m^*(\mathbf{x}) \hat{O}u_n(\mathbf{x}) dV \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \lambda_n a_m^* a_n u_m^* \int_{\mathbb{R}^3} u_m^*(\mathbf{x}) u_n(\mathbf{x}) dV \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \lambda_n a_m^* a_n \delta_{mn}, \text{ and hence} \\ \langle O \rangle_{\psi} &= \sum_{n=1}^{\infty} \lambda_n |a_n|^2. \end{aligned} \quad (300)$$

It is trivial to show that the expectation value of a Hermitian operator is purely real. Some examples:

- Expectation of  $\hat{x}$ ,

$$\begin{aligned}
\langle \hat{x} \rangle_\psi &= \int dx \psi^\dagger(x) \hat{x} \psi(x) \\
&= \int dx x |\psi(x)|^2 \\
&= \int dx x \rho(x),
\end{aligned} \tag{301}$$

which is the same as the classical notion of finding the expectation value of  $x$  given probability distribution of  $\rho(x)$ .

- Expectation of  $\hat{p}$ ,

$$\begin{aligned}
\langle \hat{p} \rangle_\psi &= -i\hbar \int dx \psi^\dagger(x) \frac{d}{dx} \psi(x) \\
&= -\frac{i\hbar}{4\pi^2} \int dx \int dk \int dk' f^\dagger(k') e^{-ikx} \frac{d}{dx} f(k) e^{ikx} \\
&= -\frac{i\hbar}{4\pi^2} \int dk \int dk' f^\dagger(k') (2\pi\delta(k-k')) f(k) \\
&= \frac{1}{2\pi} \int dk \hbar k |f(k)|^2,
\end{aligned} \tag{302}$$

which is the same as the classical notion of finding the expectation value of  $x$  given probability distribution of  $|f(k)|^2$ , in agreement with Eq. (122).

## 7.2 Commutators and Simultaneous Measurement

What do we mean by “measuring both things at the same time”?

In Classical Mechanics, this simply means that we can set up two different detectors, say  $X$  (for  $x$  measurement) and  $P$  (for  $p$  measurement). To make simultaneous measurements, we press the buttons both at the same time or even with some slight difference in time (to account for experimental error). It doesn't matter which detectors “goes first”, we will get the more or less the same answer.

In Quantum Mechanics, Postulate 3 tells us that the very act of measurement *collapses* the wavefunction, so now it *matters which detector goes first!* Given a wavefunction  $\psi(x)$ , if  $X$  goes first then the following sequence of events occurs

$$\psi(x) \xrightarrow{X} \phi_{x_0}(x) \xrightarrow{P} u_p(x) \tag{303}$$

where  $\phi_{x_0}(x)$  is a highly localized function around the measured value  $x_0$  as discussed previously, and  $u_{p_0}(x)$  is some highly localized function around the measured value  $p_0$ . On the other hand, if  $P$  goes first then

$$\psi(x) \xrightarrow{P} u_{p'_0}(x) \xrightarrow{X} \phi_{x'_0}(x). \tag{304}$$

Since  $\phi_{x_0} \neq \psi(x)$  and  $u_{p'_0}(x) \neq \psi(x)$  in general, the measured pair of values will be **different** – *the first measurement has destroyed some information regarding the second observable!* This is the root reason of why there exist an uncertainty relation in Quantum Mechanics.

We can now ask: *under what conditions will the order of the measurements not matter?* Say if we have two observables,  $\hat{O}_A$  and  $\hat{O}_B$ , then we want

$$\psi(x) \xrightarrow{\hat{O}_A} \phi(x) \xrightarrow{\hat{O}_B} \chi(x), \tag{305}$$

and

$$\psi(x) \xrightarrow{\hat{O}_B} \phi(x) \xrightarrow{\hat{O}_A} \chi(x), \tag{306}$$

to give the same observed eigenvalues of  $\hat{O}_A$  and  $\hat{O}_B$ . By inspection, it is clear that this will occur if  $\phi$  are both eigenfunctions of  $\hat{O}_A$  and  $\hat{O}_B$ , and hence so is  $\chi$ .

To formalize all these words, we will introduce some new mathematics.

**(Definition) Commutator:** The Commutator of two operators  $\hat{O}_A$  and  $\hat{O}_B$  is defined by

$$\boxed{[\hat{O}_A, \hat{O}_B] = \hat{O}_A \hat{O}_B - \hat{O}_B \hat{O}_A.} \quad (307)$$

This definition means that

$$[\hat{O}_A, \hat{O}_B] = -[\hat{O}_B, \hat{O}_A]. \quad (308)$$

We now have two possibilities that describe the situation on measurements above:

- **Commuting Observables and Simultaneous Eigenfunctions:** Suppose now  $\hat{O}_A$  and  $\hat{O}_B$  are two observables. Suppose, further that the wavefunction  $\psi(x)$  is a *simultaneous eigenfunction*  $\hat{O}_A$  and  $\hat{O}_B$  with eigenvalues  $a$  and  $b$

$$\hat{O}_A \psi(x) = a \psi(x), \quad \hat{O}_B \psi(x) = b \psi(x) \quad (309)$$

then

$$[\hat{O}_A, \hat{O}_B] \psi(x) = (\hat{O}_A \hat{O}_B - \hat{O}_B \hat{O}_A) \psi(x) = ab - ba = 0, \quad (310)$$

which is to say, “ $\hat{O}_A$  and  $\hat{O}_B$  commute”. We can write this relation in operator form by dropping  $\psi$

$$\boxed{\text{Commuting Observables} : [\hat{O}_A, \hat{O}_B] = 0.} \quad (311)$$

As have seen in the above example, commuting observables can be measured simultaneously. We call such observables **Compatible Observables** or **Commuting Observables**. Physically, this means that  $\hat{O}_A$  and  $\hat{O}_B$  has definite eigenvalues in  $\psi$ .

Now, let’s state an extremely important theorem.

**Theorem (Simultaneous Basis of Eigenfunctions) :** Suppose  $\hat{O}_A$  and  $\hat{O}_A$  commute, then they share (at least) a basis of simultaneous eigenfunctions.

*Proof:* We will prove this Theorem for the special case where *at least* one of the operator is non-degenerate. Assuming  $\hat{O}_A$  is no-degenerate, so it possess a set of eigenfunctions  $\{\psi_{a_i}\}$  with distinct eigenvalues  $\{a_i\}$ . By the eigenvalue equation

$$\hat{O}_A \psi_{a_i} = a_i \psi_{a_i}, \quad (312)$$

and operating from the left with  $\hat{O}_B$ ,

$$\hat{O}_B \hat{O}_A \psi_{a_i} = a_i \hat{O}_B \psi_{a_i} \quad (313)$$

and using commutativity  $[\hat{O}_A, \hat{O}_B] = 0$ ,

$$\hat{O}_A (\hat{O}_B \psi_{a_i}) = a_i (\hat{O}_B \psi_{a_i}) \quad (314)$$

which is to say that  $\hat{O}_B \psi_{a_i}$  is also an eigenfunction of  $\hat{O}_A$  with eigenvalue  $a_i$ . But since  $\hat{O}_A$  is degenerate,  $\hat{O}_B \psi_{a_i}$  must be the same eigenfunction as  $\psi_{a_i}$  up to a (for the moment possibly complex) number  $\lambda$  (recall that there exist an equivalence class of wavefunctions see Eq. (71)) as  $\psi_{a_i}$ , i.e.

$$\hat{O}_B \psi_{a_i} = \lambda \psi_{a_i}. \quad (315)$$

But this is nothing but an eigenvalue equation for  $\hat{O}_B$  and we identify  $\lambda$  as its eigenvalue, which by Hermiticity is real. Since every eigenfunction of  $\hat{O}_A$  is also an eigenfunction of  $\hat{O}_B$ , it is clear that

$\{\psi_{a_i}\}$  forms a complete basis for *both* operators. In this special case where  $\hat{O}_A$  is non-degenerate, there is only one such basis.  $\square$

The proof for the case where both operators are degenerate is much more involved. Those interested can see Shankar (pg 45).

Since  $\psi_{a_i}$  is also an eigenfunction of both  $\hat{O}_A$  and  $\hat{O}_B$ , and we can also give it a  $b$  label  $\psi_{a,b}$ , and we say that  $\psi_{a,b}$  are **Simultaneous Eigenfunctions** of  $\hat{O}_A$  and  $\hat{O}_B$ .

*Example:* Harking back to section 4.1, recall that  $\hat{\mathbf{p}}$  and  $\hat{H}_{\text{free}}$  share the same Eigenfunctions  $u_{\mathbf{p},E}(\mathbf{x})$  where now we have democratically label the eigenfunction without prejudice to any operator:

$$\hat{H}_{\text{free}}u_{\mathbf{p},E}(\mathbf{x}) = Eu_{\mathbf{p},E}(\mathbf{x}) , \quad \hat{\mathbf{p}}u_{\mathbf{p},E}(\mathbf{x}) = \mathbf{p}u_{\mathbf{p},E}(\mathbf{x}). \quad (316)$$

We will see another case of degeneracy and simultaneous eigenfunctions when we discuss Angular Momentum in section 8.4.

*Example :* Recall the Parity operator in 1 dimension has the action  $\hat{P}\psi(x) = \psi(-x)$ . Now consider a Hamiltonian  $\hat{H} = \hat{p}^2/2m + U(x)$  where the potential  $U(x) = U(-x)$  is reflection invariant. Then  $\hat{P}$  and  $\hat{H}$  commute

$$[\hat{P}, \hat{H}] = 0. \quad (317)$$

*Proof:* Recall from Eq. (187) that for a reflection symmetric potential  $U(x)$ , if  $\chi_E(x)$  is an eigenfunction of  $\hat{H}$  with energy  $E$ , then so is  $\chi_E(-x)$ , then

$$\begin{aligned} [\hat{P}, \hat{H}]\chi_E(x) &= \hat{P}\hat{H}\chi_E(x) - \hat{H}\hat{P}\chi_E(x) \\ &= \hat{P}E\chi_E(x) - \hat{H}\chi_E(-x) \\ &= E(\chi_E(-x) - \chi_E(-x)) = 0 \end{aligned} \quad (318)$$

and by Completeness of the eigenfunctions of  $\hat{H}$ , the proof is complete.  $\square$

**Conservation Laws:** In Classical Mechanics, some observables are *conserved under time evolution* if the potential  $U(\mathbf{x})$  has some symmetry. For example, if  $U(\mathbf{x}) = f(r)$  is spherically symmetric, then we know that the total angular momentum  $\mathbf{L}$  is conserved. In Quantum Mechanics, conservation laws are expressed as the *vanishing of the observable with the Hamiltonian*, i.e. if  $\hat{O}$  commutes with  $\hat{H}$

$$[\hat{O}, \hat{H}] = 0 \quad (319)$$

then the observable is conserved under time evolution. In the above example with Parity, you can see from the many examples in section 5 that if a state has a definite parity, then this parity is conserved under time evolution as long as the Potential is symmetric under reflection.

- **Non-commuting Observables:** The definition for non-commuting observables  $\hat{O}_A$  and  $\hat{O}_B$  is simply

$$\boxed{\text{non - Commuting Observables : } [\hat{O}_A, \hat{O}_B] \neq 0.} \quad (320)$$

In words, we say that “ $\hat{O}_A$  and  $\hat{O}_B$  do not commute”.

As you can easily prove to yourself, non-commuting observables do not share eigenfunctions, hence from the example at the start of this section this means that observations of one will now affect the observations of the other.

An example of this is our favorite pair of observables  $\hat{\mathbf{p}}$  and  $\hat{\mathbf{x}}$ . Acting on some generic state  $\psi(\mathbf{x})$  we find

$$[\hat{x}_i, \hat{x}_j]\psi(\mathbf{x}) = (x_i x_j - x_j x_i) \psi(\mathbf{x}) = 0 \quad (321)$$

while

$$[\hat{p}_i, \hat{p}_j]\psi(\mathbf{x}) = (-i\hbar)^2 \left[ \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} \right] \psi(\mathbf{x}) = 0 \quad (322)$$

using the symmetry of mixed partial derivatives. Finally,

$$\begin{aligned} [\hat{x}_i, \hat{p}_j]\psi(\mathbf{x}) &= \left[ x_i \left( -i\hbar \frac{\partial}{\partial x_j} \right) - \left( -i\hbar \frac{\partial}{\partial x_j} \right) x_i \right] \psi(\mathbf{x}) \\ &= -i\hbar \left[ x_i \frac{\partial \psi}{\partial x_j} - \frac{\partial}{\partial x_j} (x_i \psi) \right] \\ &= -i\hbar \left[ x_i \frac{\partial \psi}{\partial x_j} - \psi \frac{\partial x_i}{\partial x_j} - x_i \frac{\partial \psi}{\partial x_j} \right] \\ &= i\hbar \delta_{ij} \psi(\mathbf{x}). \end{aligned} \quad (323)$$

We obtain the **Canonical Commutator Relationships** for  $\hat{x}_i$  and  $\hat{p}_i$

$$\boxed{[\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij}}, \quad \boxed{[\hat{x}_i, \hat{x}_j] = 0}, \quad \boxed{[\hat{p}_i, \hat{p}_j] = 0} \quad (324)$$

As we will see in the next section, non-commuting observables lead to the Uncertainty Principle.

**\*Canonical Quantization:** In the lectures, we have derived Eq. (324) from our definitions of  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{p}}$ , working in the position basis. However, if we take away the basis, we can *impose* the canonical commutator relations, i.e. specifying Eq. (324) as the starting point for Quantum Mechanics and then deriving the position (or any other) basis operators from there. This is the more usual “modern” view, although our approach of deriving the momentum operator from the properties of translation is, in the view of some, more general.\*

### 7.3 Non-commuting Observables and The Uncertainty Principle

As we told the story at the start of this section, if two observables  $\hat{O}_1$  and  $\hat{O}_2$  do not commute, then the order of the measurements matter. Indeed, since say when the measurement associated with observable  $\hat{O}_1$  is made, the wavefunction collapses into one of its eigenstate, some of the information associated with  $\hat{O}_2$  is “lost” so to speak. In this section, we will quantify this.

**(Definition) Uncertainty Operator:** The uncertainty of a state  $\psi$  with respect to an observable  $\hat{O}$  is defined as

$$\boxed{\Delta \hat{O} = \hat{O} - \langle \hat{O} \rangle_\psi}. \quad (325)$$

This operator has the following properties

- **$\Delta \hat{O}$  is Hermitian.** *Proof:* As  $\hat{O}$  is an observable, it must be Hermitian, and since  $\langle \hat{O} \rangle_\psi$  is just a number,  $\Delta \hat{O}$  must also be Hermitian.  $\square$
- **Dispersion:** The expectation value of  $(\Delta \hat{O})^2$  of a state  $\psi$  is known as the dispersion, and has the following form

$$\langle (\Delta \hat{O})^2 \rangle_\psi = \langle \hat{O}^2 + \langle \hat{O} \rangle_\psi^2 - 2\hat{O}\langle \hat{O} \rangle_\psi \rangle_\psi \quad (326)$$

$$= \langle \hat{O}^2 \rangle_\psi + \langle \hat{O} \rangle_\psi^2 - 2\langle \hat{O} \rangle_\psi \langle \hat{O} \rangle_\psi \quad (327)$$

or

$$\boxed{\langle (\Delta \hat{O})^2 \rangle_\psi = \langle \hat{O}^2 \rangle_\psi - \langle \hat{O} \rangle_\psi^2} \quad (328)$$

i.e. the dispersion of  $\hat{O}$  is the “expectation of the square minus the square of the expectation”, which is consistent with the classical notion of a dispersion of an ensemble.

Furthermore, if  $\chi$  is a normalized eigenfunction of  $\hat{O}$  then  $\langle \Delta \hat{O}^2 \rangle_\chi = 0$ . Proof is by direct application of Eq. (328):

$$\begin{aligned}\langle \hat{O}^2 \rangle_\chi &= \chi \cdot (\hat{O}^2 \chi) = \lambda^2 \text{ and} \\ \langle \hat{O} \rangle_\chi^2 &= (\chi \cdot (\hat{O} \chi)) = \lambda^2\end{aligned}\quad (329)$$

so

$$\langle \hat{O}^2 \rangle_\chi - \langle \hat{O} \rangle_\chi^2 = 0 \quad \square. \quad (330)$$

In other words, if the state  $\chi_A$  is an eigenstate of  $\hat{O}_A$  then the uncertainty is zero and we measure it with probability 1, which is a trivial statement. What is *non-trivial* is that if  $\hat{O}_B$  is another observable which does not commute with  $\hat{O}_A$ , then its uncertainty in any simultaneous measurement on  $\chi_A$  will be infinite! We now state the general form of the **Uncertainty Principle**:

**Uncertainty Principle:** For any given two observables  $\hat{O}_A$  and  $\hat{O}_B$ , then the following uncertainty relation holds for any state  $\psi$

$$\langle \Delta \hat{O}_A^2 \rangle_\psi \langle \Delta \hat{O}_B^2 \rangle_\psi \geq \frac{1}{4} | \langle [\hat{O}_A, \hat{O}_B] \rangle_\psi |^2. \quad (331)$$

In the case when  $\hat{O}_A = \hat{x}$  and  $\hat{O}_B = \hat{p}$ , then using the canonical commutator relation  $[\hat{x}, \hat{p}] = i\hbar$ , we get the original famous **Heisenberg Uncertainty Principle**

$$\langle \Delta \hat{x}^2 \rangle \langle \Delta \hat{p}^2 \rangle \geq \frac{\hbar^2}{4} \quad (332)$$

which you have already seen derived using the Gaussian wavepacket in section 5.4.

We want to prove the Uncertainty Principle Eq. (331) in a snazzy operator way<sup>16</sup>. To do this, we require 2 useful lemmas.

**Lemma 1 (The Schwarz inequality):** For any two normalized states  $\psi$  and  $\phi$ , then<sup>17</sup>

$$(\psi \cdot \psi)(\phi \cdot \phi) \geq |\psi \cdot \phi|^2. \quad (333)$$

*Proof:* For any complex number  $\lambda$  and any two normalized states  $\psi$  and  $\phi$ , we can construct a state

$$\Phi = \psi + \lambda\phi \quad (334)$$

and then

$$\Phi \cdot \Phi \geq 0 \quad \forall \lambda \quad (335)$$

since  $\Phi$  is just another state and its norm must be  $\geq 0$  but  $\leq \infty$  if both  $\psi$  and  $\phi$  are normalizable. If we now set

$$\lambda = -(\phi \cdot \psi)(\phi \cdot \phi) \quad (336)$$

and plug it into Eq. (335), we get Eq. (333).  $\square$

**Lemma 2:** An **anti-Hermitian operator** is defined to be a *linear* operator which obey the relationship

$$\int_{\mathbb{R}^3} f^\dagger(\mathbf{x}) \hat{C} g(\mathbf{x}) dV = - \int_{\mathbb{R}^3} (\hat{C} f(\mathbf{x}))^\dagger g(\mathbf{x}) dV. \quad (337)$$

or more compactly

$$\hat{C} \equiv -\hat{C}^\dagger. \quad (338)$$

<sup>16</sup>See Prof. Nick Dorey's notes for a perhaps more direct way.

<sup>17</sup>This inequality is analogous to the vector space inequality  $|\mathbf{a}|^2 |\mathbf{b}|^2 \geq |\mathbf{a} \cdot \mathbf{b}|^2$  which you might have seen before.

The expectation values of anti-Hermitian operator is purely imaginary. *Proof:* Suppose  $\chi(\mathbf{x})$  is a normalized eigenfunction of  $\hat{C}$  with eigenvalue  $\lambda$ , then taking expectation values of both  $\hat{C}$  and  $\hat{C}^\dagger$

$$\begin{aligned}\langle \hat{C} \rangle_\chi &= \int_{\mathbb{R}^3} \chi^\dagger(\mathbf{x}) \hat{C} \chi(\mathbf{x}) dV = \lambda, \quad \text{and} \\ \langle \hat{C}^\dagger \rangle_\chi &= \int_{\mathbb{R}^3} \chi^\dagger(\mathbf{x}) \hat{C}^\dagger \chi(\mathbf{x}) dV\end{aligned}\tag{339}$$

$$= \int_{\mathbb{R}^3} (\hat{C} \chi(\mathbf{x}))^\dagger \chi(\mathbf{x}) dV = \lambda^*,\tag{340}$$

and using Eq. (337) we see that  $\lambda + \lambda^* = 0$  so  $\lambda \in \mathbb{C}$ , i.e. all its eigenvalues are purely imaginary. Using the Completeness property of linear operators, we can expand any state  $\psi$  in this basis so it follows that the expectation value  $\langle \hat{C} \rangle_\psi \in \mathbb{C}$ .  $\square$

We are now ready to prove Eq. (331).

**Proof (Uncertainty Principle):** Given a state  $\Psi$ , then operating on this state with the uncertainty operators  $\Delta \hat{O}_A$  and  $\Delta \hat{O}_B$  yield

$$\psi = \Delta \hat{O}_A \Psi, \quad \phi = \Delta \hat{O}_B \Psi\tag{341}$$

where  $\psi$  and  $\phi$  are some other states. Using Hermiticity of  $\Delta \hat{O}_A$ , we see that

$$\begin{aligned}\psi \cdot \psi &= \int_{\mathbb{R}^3} (\Delta \hat{O}_A \Psi(\mathbf{x}))^\dagger \Delta \hat{O}_A \Psi(\mathbf{x}) dV \\ &= \int_{\mathbb{R}^3} \Psi(\mathbf{x})^\dagger \Delta \hat{O}_A^2 \Psi(\mathbf{x}) dV = \langle (\Delta \hat{O}_A)^2 \rangle_\Psi\end{aligned}\tag{342}$$

where we have used the Hermiticity of  $\Delta \hat{O}_A$  in the 2nd line. Similarly we can calculate  $\phi \cdot \phi = \langle (\Delta \hat{O}_B)^2 \rangle_\Psi$  and  $\psi \cdot \phi = \langle \Delta \hat{O}_A \Delta \hat{O}_B \rangle_\Psi$ .

Using Lemma 1, we then take the expectation value around the state  $\Psi$  to get

$$\begin{aligned}(\psi \cdot \psi)(\phi \cdot \phi) &\geq |\psi \cdot \phi|^2 \\ \implies \langle (\Delta \hat{O}_A)^2 \rangle_\Psi \langle (\Delta \hat{O}_B)^2 \rangle_\Psi &\geq |\langle \Delta \hat{O}_A \Delta \hat{O}_B \rangle_\Psi|^2.\end{aligned}\tag{343}$$

We are halfway through the proof – our next task is to evaluate the RHS of Eq. (343). First we note that the following identity holds

$$\begin{aligned}\Delta \hat{O}_A \Delta \hat{O}_B &= \frac{1}{2} (\Delta \hat{O}_A \Delta \hat{O}_B + \Delta \hat{O}_A \Delta \hat{O}_B) \\ &= \frac{1}{2} (\Delta \hat{O}_A \Delta \hat{O}_B - \Delta \hat{O}_B \Delta \hat{O}_A + \Delta \hat{O}_B \Delta \hat{O}_A + \Delta \hat{O}_A \Delta \hat{O}_B) \\ &= \frac{1}{2} [\Delta \hat{O}_A, \Delta \hat{O}_B] + \frac{1}{2} (\Delta \hat{O}_A \Delta \hat{O}_B + \Delta \hat{O}_B \Delta \hat{O}_A).\end{aligned}\tag{344}$$

But the commutator  $[\Delta \hat{O}_A, \Delta \hat{O}_B] = [\hat{O}_A, \hat{O}_B]$  is *anti-Hermitian*

$$([\hat{O}_A, \hat{O}_B])^\dagger = (\hat{O}_A \hat{O}_B - \hat{O}_B \hat{O}_A)^\dagger = \hat{O}_B \hat{O}_A - \hat{O}_A \hat{O}_B = -[\hat{O}_A, \hat{O}_B]\tag{345}$$

while the last term on Eq. (344) is Hermitian

$$(\Delta \hat{O}_A \Delta \hat{O}_B + \Delta \hat{O}_B \Delta \hat{O}_A)^\dagger = \Delta \hat{O}_B \Delta \hat{O}_A + \Delta \hat{O}_A \Delta \hat{O}_B.\tag{346}$$

Hence the RHS of Eq. (343) becomes, using Lemma 2 for the expectation value of  $[\hat{O}_A, \hat{O}_B]$ ,

$$\begin{aligned}|\langle \Delta \hat{O}_A \Delta \hat{O}_B \rangle_\Psi|^2 &\geq \underbrace{\left| \frac{1}{2} \langle [\hat{O}_A, \hat{O}_B] \rangle \right|^2}_{\text{Imaginary}} + \underbrace{\left| \frac{1}{2} \langle (\Delta \hat{O}_A \Delta \hat{O}_B + \Delta \hat{O}_B \Delta \hat{O}_A) \rangle \right|^2}_{\text{Real}} \\ &\geq \frac{1}{4} \left| \langle [\hat{O}_A, \hat{O}_B] \rangle \right|^2 + \frac{1}{4} \left| \langle (\Delta \hat{O}_A \Delta \hat{O}_B + \Delta \hat{O}_B \Delta \hat{O}_A) \rangle \right|^2\end{aligned}\tag{347}$$

and since the last term can only make the inequality stronger, the proof is complete.  $\square$



## 7.4 Summary

In this section, we study the notion of simultaneous observations and elaborated on how some observables are inherently incompatible with each other and the measurement of one will destroy information of the other(s). Such incompatibility is encoded in mathematical language as non-commutativity of the operators associated with the observables.

We then show that the Postulates of Quantum Mechanics lead us to the Uncertainty Principle – which is a powerful *consequence* of Postulate 3 (collapse of a wavefunction after a measurement), restricting our ability to extract information out of a wavefunction. How much information is “destroyed” by the collapse is given by the amount of non-commutativity of the observables as indicated by Eq. (331). Returning to  $\hat{x}$  and  $\hat{p}$ , their commutator is  $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$ , i.e. the amount of “lost information” is proportional to the Planck’s Constant  $\hbar$ , which sets the scale of Quantum Mechanics. Since Classically, no information is “lost” in any measurement, the “Classical Limit” of a quantum theory can be recovered by taking the limit  $\hbar \rightarrow 0$ .

This section marks the end of our formal development of Quantum Mechanics.

[Next](#) [Up](#) [Previous](#) [Contents](#)

**Next:** [Postulates of Quantum Mechanics](#) **Up:** [Mathematical Background](#) **Previous:** [Commutators in Quantum Mechanics](#) [Contents](#)

## Linear Vector Spaces in Quantum Mechanics

We have observed that most operators in quantum mechanics are linear operators. This is fortunate because it allows us to represent quantum mechanical operators as matrices and wavefunctions as vectors in some linear vector space. Since computers are particularly good at performing operations common in linear algebra (multiplication of a matrix times a vector, etc.), this is quite advantageous from a practical standpoint.

In an  $n$ -dimensional space we may expand any vector  $\Psi$  as a linear combination of basis vectors

$$\Psi = \sum_{i=1}^n a_i \Psi_i \quad (80)$$

For a general vector space, the coefficients  $a_i$  may be complex; thus one should not be too quick to draw parallels to the expansion of vectors in three-dimensional Euclidean space. The coefficients  $a_i$  are referred to as the "components" of the state vector  $\Psi$ , and for a given basis, the components of a vector specify it completely. The components of the sum of two vectors are the sums of the components. If  $\Psi_a = \sum a_i \Psi_i$  and  $\Psi_b = \sum b_i \Psi_i$  then

$$\Psi_a + \Psi_b = \sum_i (a_i + b_i) \Psi_i \quad (81)$$

and similarly

$$\lambda \Psi_a = \sum_i (\lambda a_i) \Psi_i \quad (82)$$

The *scalar product* of two vectors is a complex number denoted by

$$(\Psi_b, \Psi_a) = (\Psi_a, \Psi_b)^* \quad (83)$$

where we have used the standard linear-algebra notation. If we also require that

$$(\Psi_a, \lambda \Psi_b) = \lambda (\Psi_a, \Psi_b) \quad (84)$$

then it follows that

$$(\lambda\Psi_a, \Psi_b) = \lambda^*(\Psi_a, \Psi_b) \quad (85)$$

We also require that

$$(\Psi_a, \Psi_b + \Psi_c) = (\Psi_a, \Psi_b) + (\Psi_a, \Psi_c) \quad (86)$$

If the scalar product vanishes (and if neither vector in the product is the null vector) then the two vectors are orthogonal.

Generally the basis is chosen to be orthonormal, such that

$$(\hat{\Psi}_i, \hat{\Psi}_j) = \delta_{ij} \quad (87)$$

In this case, we can write the scalar product of two arbitrary vectors as

$$\begin{aligned} (\Psi_a, \Psi_b) &= \left( \sum_i a_i \hat{\Psi}_i, \sum_j b_j \hat{\Psi}_j \right) \\ &= \sum_i \sum_j a_i^* b_j (\hat{\Psi}_i, \hat{\Psi}_j) \\ &= \sum_i a_i^* b_i \end{aligned} \quad (88)$$

This can also be written in vector notation as

$$(\Psi_a, \Psi_b) = (a_1^* a_2^* \cdots a_n^*) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad (89)$$

It is useful at this point to introduce Dirac's bra-ket notation. We define a "bra" as

$$\langle \Psi_a | = (a_1^* a_2^* \cdots a_n^*) \quad (90)$$

and a "ket" as

$$|\Psi_a\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad (91)$$

A bra to the left of a ket implies a scalar product, so

$$\langle\Psi_a|\Psi_b\rangle = (\Psi_a, \Psi_b) \quad (92)$$

Sometimes in superficial treatments of Dirac notation, the symbol  $\langle\Psi_a|\Psi_b\rangle$  is defined alternatively as

$$\langle\Psi_a|\Psi_b\rangle = \int \Psi_a^*(x)\Psi_b(x)dx \quad (93)$$

This is equivalent to the above definition if we make the connections  $a_i = \Psi_a(x)$  and  $b_i = \Psi_b(x)$ . This means that our basis vectors are *every possible value of x*. Since  $\underline{x}$  is continuous, the sum is replaced by an integral (see Szabo and Ostlund [4], exercise 1.17). Often only the subscript of the vector is used to denote a bra or ket; we may have written the above equation as

$$\langle a|b\rangle = \int \Psi_a^*(x)\Psi_b(x)dx \quad (94)$$

Now we turn our attention to matrix representations of operators. An operator  $\hat{A}$  can be characterized by its effect on the basis vectors. The action of  $\hat{A}$  on a basis vector  $\hat{\Psi}_j$  yields some new vector  $\underline{\Psi'_j}$  which can be expanded in terms of the basis vectors so long as we have a complete basis set.

$$\hat{A}\hat{\Psi}_j = \underline{\Psi'_j} = \sum_i^n \hat{\Psi}_i A_{ij} \quad (95)$$

If we know the effect of  $\hat{A}$  on the basis vectors, then we know the effect of  $\hat{A}$  on any arbitrary vector because of the linearity of  $\hat{A}$ .

$$\begin{aligned} \Psi_b &= \hat{A}\Psi_a = \hat{A}\sum_j a_j \hat{\Psi}_j = \sum_j a_j \hat{A}\hat{\Psi}_j = \sum_j \sum_i a_j \hat{\Psi}_i A_{ij} \\ &= \sum_i \hat{\Psi}_i (\sum_j A_{ij} a_j) \end{aligned} \quad (96)$$

or

$$b_i = \sum_j A_{ij} a_j \quad (97)$$

This may be written in matrix notation as

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad (98)$$

We can obtain the coefficients  $A_{ij}$  by taking the inner product of both sides of equation 95 with  $\hat{\Psi}_i$ , yielding

$$\begin{aligned} (\hat{\Psi}_i, \hat{A}\hat{\Psi}_j) &= (\hat{\Psi}_i, \sum_k^n \hat{\Psi}_k A_{kj}) \\ &= \sum_k^n A_{kj} (\hat{\Psi}_i, \hat{\Psi}_k) \\ &= A_{ij} \end{aligned} \quad (99)$$

since  $(\hat{\Psi}_i, \hat{\Psi}_k) = \delta_{ik}$  due to the orthonormality of the basis. In bra-ket notation, we may write

$$A_{ij} = \langle i | \hat{A} | j \rangle \quad (100)$$

where  $i$  and  $j$  denote two basis vectors. This use of bra-ket notation is consistent with its earlier use if we realize that  $\hat{A}|j\rangle$  is just another vector  $|j'\rangle$ .

It is easy to show that for a linear operator  $\hat{A}$ , the inner product  $(\Psi_a, \hat{A}\Psi_b)$  for two general vectors (not necessarily basis vectors)  $\Psi_a$  and  $\Psi_b$  is given by

$$(\Psi_a, \hat{A}\Psi_b) = \sum_i \sum_j a_i^* A_{ij} b_j \quad (101)$$

or in matrix notation

$$(\Psi_a, \hat{A}\Psi_b) = (a_1^* a_2^* \cdots a_n^*) \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad (102)$$

By analogy to equation (93), we may generally write this inner product in the form

$$(\Psi_a, \hat{A}\Psi_b) = \langle a | \hat{A} | b \rangle = \int \Psi_a^*(x) \hat{A}\Psi_b(x) dx \quad (103)$$

Previously, we noted that  $(\Psi_a, \Psi_b) = (\Psi_b, \Psi_a)^*$ , or  $\langle a | b \rangle = \langle b | a \rangle^*$ . Thus we can see also that

$$(\Psi_a, \hat{A}\Psi_b) = (\hat{A}\Psi_b, \Psi_a)^* \quad (104)$$

We now define the *adjoint* of an operator  $\hat{A}$ , denoted by  $\hat{A}^\dagger$ , as that linear operator for which

$$(\Psi_a, \hat{A}\Psi_b) = (\hat{A}^\dagger\Psi_a, \Psi_b) \quad (105)$$

That is, we can make an operator act *backwards* into "bra" space if we take it's adjoint. With this definition, we can further see that

$$(\Psi_a, \hat{A}\Psi_b) = (\hat{A}\Psi_b, \Psi_a)^* = (\Psi_b, \hat{A}^\dagger\Psi_a)^* = (\hat{A}^\dagger\Psi_a, \Psi_b) \quad (106)$$

or, in bra-ket notation,

$$\langle a | \hat{A} | b \rangle = \langle \hat{A} b | a \rangle^* = \langle b | \hat{A}^\dagger | a \rangle^* = \langle \hat{A}^\dagger a | b \rangle \quad (107)$$

If we pick  $\Psi_a = \hat{\Psi}_i$  and  $\Psi_b = \hat{\Psi}_j$  (i.e., if we pick two basis vectors), then we obtain

$$\begin{aligned} (\hat{A}\hat{\Psi}_i, \hat{\Psi}_j) &= (\hat{\Psi}_i, \hat{A}^\dagger\hat{\Psi}_j) \\ (\hat{\Psi}_j, \hat{A}\hat{\Psi}_i)^* &= (\hat{\Psi}_i, \hat{A}^\dagger\hat{\Psi}_j) \\ A_{ji}^* &= \underline{A_{ij}^\dagger} \end{aligned} \quad (108)$$

But this is precisely the condition for the elements of a matrix and its adjoint! Thus the adjoint of the matrix representation of  $\hat{A}$  is the same as the matrix representation of  $\hat{A}^\dagger$ .

This correspondence between operators and their matrix representations goes quite far, although of course the specific matrix representation depends on the choice of basis. For instance, we know from linear algebra that if a matrix and its adjoint are the same, then the matrix is called Hermitian. The same is true of the operators; if

$$\hat{A} = \hat{A}^\dagger \quad (109)$$

then  $\hat{A}$  is a Hermitian operator, and all of the special properties of Hermitian operators apply to  $\hat{A}$  or its matrix representation.

---

[Next](#) [Up](#) [Previous](#) [Contents](#)

**Next:** [Postulates of Quantum Mechanics](#) **Up:** [Mathematical Background](#) **Previous:** [Commutators in Quantum Mechanics](#) [Contents](#)

*David Sherrill 2006-08-15*

## Chapter 10 The Hydrogen Atom

There are many good reasons to address the hydrogen atom beyond its historical significance. Though hydrogen spectra motivated much of the early quantum theory, research involving the hydrogen remains at the cutting edge of science and technology. For instance, transitions in hydrogen are being used in 1997 and 1998 to examine the constancy of the fine structure constant over a cosmological time scale<sup>2</sup>. From the view point of pedagogy, the hydrogen atom merges many of the concepts and techniques previously developed into one package. It is a particle in a box with spherical, soft walls. Finally, the hydrogen atom is one of the precious few realistic systems which can actually be solved analytically.

### The Schrodinger Equation in Spherical Coordinates

In chapter 5, we separated time and position to arrive at the time independent Schrodinger equation which is

$$\mathcal{H}|E_i\rangle = E_i|E_i\rangle, \quad (10-1)$$

where  $E_i$  are eigenvalues and  $|E_i\rangle$  are energy eigenstates. Also in chapter 5, we developed a one dimensional position space representation of the time independent Schrodinger equation, changing the notation such that  $E_i \rightarrow E$ , and  $|E_i\rangle \rightarrow \psi$ . In three dimensions the Schrodinger equation generalizes to

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi = E\psi,$$

where  $\nabla^2$  is the Laplacian operator. Using the Laplacian in spherical coordinates, the Schrodinger equation becomes

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\psi + V(r)\psi = E\psi. \quad (10-2)$$

In spherical coordinates,  $\psi = \psi(r, \theta, \phi)$ , and the plan is to look for a variables separable solution such that  $\psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$ . We will in fact find such solutions where  $Y(\theta, \phi)$  are the spherical harmonic functions and  $R(r)$  is expressible in terms of associated Laguerre functions. Before we do that, interfacing with the previous chapter and arguments of linear algebra may partially explain why we are proceeding in this direction.

### Complete Set of Commuting Observables for Hydrogen

Though we will return to equation (10-2), the Laplacian can be expressed

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{\tan\theta}\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right). \quad (10-3)$$

Compare the terms in parenthesis to equation 11-33. The terms in parenthesis are equal to  $-\mathcal{L}^2/\hbar^2$ , so assuming spherical symmetry, the Laplacian can be written

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{\mathcal{L}^2}{r^2\hbar^2},$$

---

<sup>2</sup> Schwarzschild. "Optical Frequency Measurement is Getting a Lot More Precise," Physics Today 50(10) 19-21 (1997).



and the Schrodinger equation becomes

$$\left[ -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{\mathcal{L}^2}{r^2 \hbar^2} \right) + V(r) \right] \psi = E\psi. \quad (10-4)$$

Assuming spherical symmetry, which we will have because a Coulomb potential will be used for  $V(r)$ , we have complicated the system of chapter 11 by adding a radial variable. Without the radial variable, we have a complete set of commuting observables for the angular momentum operators in  $\mathcal{L}^2$  and  $\mathcal{L}_z$ . Including the radial variable, we need a minimum of one more operator, if that operator commutes with both  $\mathcal{L}^2$  and  $\mathcal{L}_z$ . The total energy operator, the Hamiltonian, may be a reasonable candidate. What is the Hamiltonian here? It is the group of terms within the square brackets. Compare equations (10-1) and (10-4) if you have difficulty visualizing that. In fact,

$$[\mathcal{H}, \mathcal{L}^2] = 0, \quad \text{and} \quad [\mathcal{H}, \mathcal{L}_z] = 0,$$

so the Hamiltonian is a suitable choice. The complete set of commuting observables for the hydrogen atom is  $\mathcal{H}$ ,  $\mathcal{L}^2$ , and  $\mathcal{L}_z$ . We have all the eigenvalue/eigenvector equations, because the time independent Schrodinger equation is the eigenvalue/eigenvector equation for the Hamiltonian operator, *i.e.*, the the eigenvalue/eigenvector equations are

$$\begin{aligned} \mathcal{H}|\psi\rangle &= E_n|\psi\rangle, \\ \mathcal{L}^2|\psi\rangle &= l(l+1)\hbar^2|\psi\rangle, \\ \mathcal{L}_z|\psi\rangle &= m\hbar|\psi\rangle, \end{aligned}$$

where we subscripted the energy eigenvalue with an  $n$  because that is the symbol conventionally used for the energy quantum number (per the particle in a box and SHO). Then the solution to the problem is the eigenstate which satisfies all three, denoted  $|n, l, m\rangle$  in abstract Hilbert space. The representation in position space in spherical coordinates is

$$\langle r, \theta, \phi | n, l, m \rangle = \psi_{nlm}(r, \theta, \phi).$$

**Example 10-1:** Starting with the Laplacian included in equation (10-2), show the Laplacian can be express as equation (10-3).

$$\begin{aligned} \nabla^2 &= \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \\ &= \frac{1}{r^2} \left( 2r \frac{\partial}{\partial r} + r^2 \frac{\partial^2}{\partial r^2} \right) + \frac{1}{r^2 \sin \theta} \left( \cos \theta \frac{\partial}{\partial \theta} + \sin \theta \frac{\partial^2}{\partial \theta^2} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \\ &= \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2 \tan \theta} \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \\ &= \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left( \frac{\partial^2}{\partial \theta^2} + \frac{1}{\tan \theta} \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right), \end{aligned}$$

which is the form of equation (10-3).

**Example 10-2:** Show  $[\mathcal{H}, \mathcal{L}^2] = 0$ .

$$[\mathcal{H}, \mathcal{L}^2] = \mathcal{H} \mathcal{L}^2 - \mathcal{L}^2 \mathcal{H}$$

$$\begin{aligned}
&= \left[ -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{\mathcal{L}^2}{r^2 \hbar^2} \right) + V(r) \right] \mathcal{L}^2 - \mathcal{L}^2 \left[ -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{\mathcal{L}^2}{r^2 \hbar^2} \right) + V(r) \right] \\
&= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^2} \mathcal{L}^2 - \frac{\hbar^2}{2m} \frac{2}{r} \frac{\partial}{\partial r} \mathcal{L}^2 + \frac{\hbar^2}{2m} \frac{\mathcal{L}^4}{r^2 \hbar^2} + \frac{\hbar^2}{2m} V(r) \mathcal{L}^2 \\
&\quad + \frac{\hbar^2}{2m} \mathcal{L}^2 \frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2m} \mathcal{L}^2 \frac{2}{r} \frac{\partial}{\partial r} - \frac{\hbar^2}{2m} \frac{\mathcal{L}^4}{r^2 \hbar^2} - \frac{\hbar^2}{2m} \mathcal{L}^2 V(r) \\
&= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^2} \mathcal{L}^2 - \frac{\hbar^2}{2m} \frac{2}{r} \frac{\partial}{\partial r} \mathcal{L}^2 + \frac{\hbar^2}{2m} V(r) \mathcal{L}^2 + \frac{\hbar^2}{2m} \mathcal{L}^2 \frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2m} \mathcal{L}^2 \frac{2}{r} \frac{\partial}{\partial r} - \frac{\hbar^2}{2m} \mathcal{L}^2 V(r)
\end{aligned}$$

where the third and seventh terms in  $\mathcal{L}^4$  sum to zero. The spherical coordinate representation of  $\mathcal{L}^2$  is

$$\mathcal{L}^2 = -\hbar^2 \left( \frac{\partial^2}{\partial \theta^2} + \frac{1}{\tan \theta} \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right)$$

and has angular dependence only. The partial derivatives with respect to the radial variable act only on terms without radial dependence. Partial derivatives with respect to angular variables do not affect the potential which is a function only of the radial variable. Therefore, the order of the operator products is interchangeable, and

$$[\mathcal{H}, \mathcal{L}^2] = -\frac{\hbar^2}{2m} \mathcal{L}^2 \frac{\partial^2}{\partial r^2} - \frac{\hbar^2}{2m} \mathcal{L}^2 \frac{2}{r} \frac{\partial}{\partial r} + \frac{\hbar^2}{2m} \mathcal{L}^2 V(r) + \frac{\hbar^2}{2m} \mathcal{L}^2 \frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2m} \mathcal{L}^2 \frac{2}{r} \frac{\partial}{\partial r} - \frac{\hbar^2}{2m} \mathcal{L}^2 V(r) = 0.$$

Instead of the verbal argument, we could substitute the angular representation of  $\mathcal{L}^2$ , form the 18 resultant terms, explicitly interchange nine of them, and get the same result.

---

**Example 10-3:** Show  $[\mathcal{H}, \mathcal{L}_z] = 0$ .

$$\begin{aligned}
[\mathcal{H}, \mathcal{L}_z] &= \mathcal{H} \mathcal{L}_z - \mathcal{L}_z \mathcal{H} \\
&= \left[ -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{\mathcal{L}^2}{r^2 \hbar^2} \right) + V(r) \right] \mathcal{L}_z - \mathcal{L}_z \left[ -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{\mathcal{L}^2}{r^2 \hbar^2} \right) + V(r) \right] \\
&= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^2} \mathcal{L}_z - \frac{\hbar^2}{2m} \frac{2}{r} \frac{\partial}{\partial r} \mathcal{L}_z + \frac{\hbar^2}{2m} \frac{\mathcal{L}^2 \mathcal{L}_z}{r^2 \hbar^2} + \frac{\hbar^2}{2m} V(r) \mathcal{L}_z \\
&\quad + \frac{\hbar^2}{2m} \mathcal{L}_z \frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2m} \mathcal{L}_z \frac{2}{r} \frac{\partial}{\partial r} - \frac{\hbar^2}{2m} \frac{\mathcal{L}_z \mathcal{L}^2}{r^2 \hbar^2} - \frac{\hbar^2}{2m} \mathcal{L}_z V(r) \\
&= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^2} \mathcal{L}_z - \frac{\hbar^2}{2m} \frac{2}{r} \frac{\partial}{\partial r} \mathcal{L}_z + \frac{\hbar^2}{2m} V(r) \mathcal{L}_z + \frac{\hbar^2}{2m} \mathcal{L}_z \frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2m} \mathcal{L}_z \frac{2}{r} \frac{\partial}{\partial r} - \frac{\hbar^2}{2m} \mathcal{L}_z V(r)
\end{aligned}$$

where the third and seventh terms in  $\mathcal{L}^2 \mathcal{L}_z$  sum to zero because we already know those two operators commute. The spherical coordinate representation of  $\mathcal{L}_z$  is

$$\mathcal{L}_z = -i\hbar \frac{\partial}{\partial \phi}$$

and has angular dependence only. Again there are no partial derivatives which affect any term of the other operator, or the potential  $V(r)$ , in any of the operator products. Therefore, the order of the operator products is interchangeable, and

$$[\mathcal{H}, \mathcal{L}_z] = -\frac{\hbar^2}{2m} \mathcal{L}_z \frac{\partial^2}{\partial r^2} - \frac{\hbar^2}{2m} \mathcal{L}_z \frac{2}{r} \frac{\partial}{\partial r} + \frac{\hbar^2}{2m} \mathcal{L}_z V(r) + \frac{\hbar^2}{2m} \mathcal{L}_z \frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2m} \mathcal{L}_z \frac{2}{r} \frac{\partial}{\partial r} - \frac{\hbar^2}{2m} \mathcal{L}_z V(r) = 0.$$

---

## Separating Radial and Angular Dependence

In this and the following three sections, we illustrate how the angular momentum and magnetic moment quantum numbers enter the symbology from a calculus based argument. In writing equation (10–2), we have used a representation, so are no longer in abstract Hilbert space. One of the consequences of the process of representation is the topological arguments of linear algebra are obscured. They are still there, simply obscured because the special functions we use are orthogonal, so can be made orthonormal, and complete, just as bras and kets in a dual space are orthonormal and complete. The primary reason to proceed in terms of a position space representation is to attain a position space description. One of the by-products of this chapter may be to convince you that working in the generality of Hilbert space in Dirac notation can be considerably more efficient. Since we used topological arguments to develop angular momentum in the last chapter, and arrive at identical results to those of chapter 11, we rely on connections between the two to establish the meanings of  $l$  and  $m$ . They have the same meanings within these calculus based discussions.

As noted, we assume a variables separable solution to equation (10–2) of the form

$$\psi(r, \theta, \phi) = R(r) Y(\theta, \phi). \quad (10-5)$$

An often asked question is “How do you know you can assume that?” You do not know. You assume it, and if it works, you have found a solution. If it does not work, you need to attempt other methods or techniques. Here, it will work. Using equation (10–5), equation (10–2) can be written

$$\begin{aligned} \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) R(r) Y(\theta, \phi) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) R(r) Y(\theta, \phi) \\ + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} R(r) Y(\theta, \phi) - \frac{2m}{\hbar^2} [V(r) - E] R(r) Y(\theta, \phi) = 0 \\ \Rightarrow Y(\theta, \phi) \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) R(r) + R(r) \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) Y(\theta, \phi) \\ + R(r) \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} Y(\theta, \phi) - \frac{2m}{\hbar^2} [V(r) - E] R(r) Y(\theta, \phi) = 0. \end{aligned}$$

Dividing the equation by  $R(r) Y(\theta, \phi)$ , multiplying by  $r^2$ , and rearranging terms, this becomes

$$\begin{aligned} \left\{ \frac{1}{R(r)} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) R(r) - \frac{2mr^2}{\hbar^2} [V(r) - E] \right\} \\ + \left[ \frac{1}{Y(\theta, \phi) \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) Y(\theta, \phi) + \frac{1}{Y(\theta, \phi) \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} Y(\theta, \phi) \right] = 0. \end{aligned}$$

The two terms in the curly braces depend only on  $r$ , and the two terms in the square brackets depend only upon angles. With the exception of a trivial solution, the only way the sum of the groups can be zero is if each group is equal to the same constant. The constant chosen is known as the **separation constant**. Normally, an arbitrary separation constant, like  $K$ , is selected and then you solve for  $K$  later. In this example, we are instead going to stand on the shoulders of

some of the physicists and mathematicians of the previous 300 years, and make the enlightened choice of  $l(l+1)$  as the separation constant. It should become clear  $l$  is the angular momentum quantum number introduced in chapter 11. Then

$$\frac{1}{R(r)} \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) R(r) - \frac{2mr^2}{\hbar^2} [V(r) - E] = l(l+1) \quad (10-6)$$

which we call the **radial equation**, and

$$\frac{1}{Y(\theta, \phi) \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) Y(\theta, \phi) + \frac{1}{Y(\theta, \phi) \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} Y(\theta, \phi) = -l(l+1), \quad (10-7)$$

which we call the **angular equation**. Notice the signs on the right side are opposite so they do, in fact, sum to zero.

## The Angular Equation

The solutions to equation (10-7) are the spherical harmonic functions, and the  $l$  used in the separation constant is, in fact, the same used as the index  $l$  in the spherical harmonics  $Y_{l,m}(\theta, \phi)$ . In fact, it is the angular momentum quantum number. But where is the index  $m$ ? How is the magnetic moment quantum number introduced? To answer these questions, remember the spherical harmonics are also separable, *i.e.*,  $Y_{l,m}(\theta, \phi) = f_{l,m}(\theta) g_m(\phi)$ . We will use such a solution in the angular equation, without the indices until we see where they originate. Using the solution  $Y(\theta, \phi) = f(\theta) g(\phi)$  in equation (10-7),

$$\begin{aligned} \frac{1}{f(\theta) g(\phi) \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) f(\theta) g(\phi) + \frac{1}{f(\theta) g(\phi) \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} f(\theta) g(\phi) &= -l(l+1) \\ \Rightarrow \frac{1}{f(\theta) \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) f(\theta) + \frac{1}{g(\phi) \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} g(\phi) &= -l(l+1). \end{aligned}$$

Multiplying the equation by  $\sin^2 \theta$  and rearranging,

$$\frac{\sin \theta}{f(\theta)} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) f(\theta) + l(l+1) \sin^2 \theta + \frac{1}{g(\phi)} \frac{\partial^2}{\partial \phi^2} g(\phi) = 0.$$

The first two terms depend only on  $\theta$ , and the last term depends only on  $\phi$ . Again, the only non-trivial solution such that the sum is zero is if the groups of terms each dependent on a single variable is equal to the same constant. Again using an enlightened choice, we pick  $m^2$  as the separation constant, so

$$\frac{\sin \theta}{f(\theta)} \frac{d}{d\theta} \left( \sin \theta \frac{d}{d\theta} \right) f(\theta) + l(l+1) \sin^2 \theta = m^2, \quad (10-8)$$

$$\frac{1}{g(\phi)} \frac{d^2}{d\phi^2} g(\phi) = -m^2, \quad (10-9)$$

and that is how the magnetic moment quantum number is introduced. Again, (10-8) and (10-9) need to sum to zero so the separation constant has opposite signs on the right side in the two equations.

## The Azimuthal Angle Equation

The solution to the azimuthal angle equation, equation (10-9), is

$$g(\phi) = e^{im\phi} \Rightarrow g_m(\phi) = e^{im\phi}, \quad (10-10)$$

where the subscript  $m$  is added to  $g(\phi)$  because it is now clear there are as many solutions as there are allowed values of  $m$ .

---

**Example 10-4:** Show  $g_m(\phi) = e^{im\phi}$  is a solution to equation (10-9).

$$\frac{d^2}{d\phi^2} g_m(\phi) = \frac{d^2}{d\phi^2} e^{im\phi} = \frac{d}{d\phi} (im) e^{im\phi} = (im)^2 e^{im\phi} = -m^2 g_m(\phi).$$

Using this in equation (10-9),

$$\frac{1}{g(\phi)} \frac{d^2}{d\phi^2} g(\phi) = -m^2 \Rightarrow \frac{1}{g(\phi)} \left( -m^2 g_m(\phi) \right) = -m^2 \Rightarrow -m^2 = -m^2,$$

therefore  $g_m(\phi) = e^{im\phi}$  is a solution to equation (10-9).

---

## The Polar Angle Equation

This section is a little more substantial than the last. Equation (10-8) can be written

$$\sin \theta \frac{d}{d\theta} \left( \sin \theta \frac{d}{d\theta} \right) f(\theta) + l(l+1) \sin^2 \theta f(\theta) - m^2 f(\theta) = 0.$$

Evaluating the first term,

$$\begin{aligned} \sin \theta \frac{d}{d\theta} \left( \sin \theta \frac{d}{d\theta} \right) f(\theta) &= \sin \theta \frac{d}{d\theta} \left( \sin \theta \frac{df(\theta)}{d\theta} \right) \\ &= \sin \theta \left( \cos \theta \frac{df(\theta)}{d\theta} + \sin \theta \frac{d^2 f(\theta)}{d\theta^2} \right) \\ &= \sin^2 \theta \frac{d^2 f(\theta)}{d\theta^2} + \sin \theta \cos \theta \frac{df(\theta)}{d\theta}. \end{aligned}$$

Using this, equation (10-8) becomes

$$\sin^2 \theta \frac{d^2 f(\theta)}{d\theta^2} + \sin \theta \cos \theta \frac{df(\theta)}{d\theta} + l(l+1) \sin^2 \theta f(\theta) - m^2 f(\theta) = 0. \quad (10-11)$$

We are going to change variables using  $x = \cos \theta$ , and will comment on this substitution later. We then need the derivatives with respect to  $x$  vice  $\theta$ , so

$$\frac{df(\theta)}{d\theta} = \frac{df(x)}{dx} \frac{dx}{d\theta} = \frac{df(x)}{dx} (-\sin \theta) = -\sin \theta \frac{df(x)}{dx},$$

and

$$\begin{aligned}\frac{d^2 f(\theta)}{d\theta^2} &= \frac{d}{d\theta} \left( -\sin \theta \frac{df(x)}{dx} \right) = -\cos \theta \frac{df(x)}{dx} - \sin \theta \frac{d}{d\theta} \frac{df(x)}{dx} \\ &= -\cos \theta \frac{df(x)}{dx} - \sin \theta \frac{d}{dx} \frac{dx}{d\theta} \frac{df(x)}{dx} = -\cos \theta \frac{df(x)}{dx} - \sin \theta \frac{d}{dx} \left( -\sin \theta \right) \frac{df(x)}{dx} \\ &= -\cos \theta \frac{df(x)}{dx} + \sin^2 \theta \frac{d^2 f(x)}{dx^2}.\end{aligned}$$

Substituting just the derivatives in the equation (10–11),

$$\sin^2 \theta \left( \sin^2 \theta \frac{d^2 f(x)}{dx^2} - \cos \theta \frac{df(x)}{dx} \right) + \sin \theta \cos \theta \left( -\sin \theta \frac{df(x)}{dx} \right) + l(l+1) \sin^2 \theta f(x) - m^2 f(x) = 0,$$

which gives us an equation in both  $\theta$  and  $x$ , which is not formally appropriate. This is, however, an informal text, and it becomes difficult to keep track of the terms if all the substitutions and reductions are done at once. Dividing by  $\sin^2 \theta$ , we get

$$\sin^2 \theta \frac{d^2 f(x)}{dx^2} - \cos \theta \frac{df(x)}{dx} - \cos \theta \frac{df(x)}{dx} + l(l+1) f(x) - \frac{m^2}{\sin^2 \theta} f(x) = 0.$$

The change of variables is complete upon summing the two first derivatives, using  $\cos \theta = x$ , and  $\sin^2 \theta = 1 - \cos^2 \theta = 1 - x^2$ , which is

$$\left(1 - x^2\right) \frac{d^2 f(x)}{dx^2} - 2x \frac{df(x)}{dx} + l(l+1) f(x) - \frac{m^2}{1 - x^2} f(x) = 0.$$

This is the **associated Legendre equation**, which reduces to **Legendre equation** when  $m = 0$ . The function has a single argument so there is no confusion if the derivatives are indicated with primes, and the associated Legendre equation is often written

$$\left(1 - x^2\right) f''(x) - 2x f'(x) + l(l+1) f(x) - \frac{m^2}{1 - x^2} f(x) = 0,$$

and becomes the Legendre equation,

$$\left(1 - x^2\right) f''(x) - 2x f'(x) + l(l+1) f(x) = 0,$$

when  $m = 0$ . The solutions to the associated Legendre equation are the associated Legendre polynomials discussed briefly in the last section of chapter 11. To review that in the current context, associated Legendre polynomials can be generated from Legendre polynomials using

$$P_{l,m}(x) = (-1)^m \sqrt{(1-x^2)^m} \frac{d^m}{dx^m} P_l(x),$$

where the  $P_l(x)$  are Legendre polynomials. Legendre polynomials can be generated using

$$P_l(x) = \frac{(-1)^l}{2^l l!} \frac{d^l}{dx^l} (1-x^2)^l.$$

The use of these generating functions was illustrated in example 11–26 as intermediate results in calculating spherical harmonics.

The first few Legendre polynomials are listed in table 10–1. Our interest in those is to generate associated Legendre functions. The first few associated Legendre polynomials are listed in table 10–2.

$P_0(x) = 1$	$P_3(x) = \frac{1}{2}(5x^3 - 3x)$
$P_1(x) = x$	$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$
$P_2(x) = \frac{1}{2}(3x^2 - 1)$	$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$
Table 10 – 1. The First Six Legendre Polynomials.	

$P_{0,0}(x) = 1$	$P_{2,0}(x) = \frac{1}{2}(3x^2 - 1)$
$P_{1,1}(x) = -\sqrt{1-x^2}$	$P_{3,3}(x) = -15(\sqrt{1-x^2})^3$
$P_{1,0}(x) = x$	$P_{3,2}(x) = 15x(1-x^2)$
$P_{2,2}(x) = 3(1-x^2)$	$P_{3,1}(x) = -\frac{3}{2}(5x^2 - 1)\sqrt{1-x^2}$
$P_{2,1}(x) = -3x\sqrt{1-x^2}$	$P_{3,0}(x) = \frac{1}{2}(5x^3 - 3x)$
Table 10 – 2. The First Few Associated Legendre Polynomials.	

Two comment concerning the tables are appropriate. First, notice  $P_l = P_{l,0}$ . That makes sense. If the Legendre equation is the same as the associated Legendre equation with  $m = 0$ , the solutions to the two equations must be the same when  $m = 0$ . Also, many authors will use a positive sign for all associated Legendre polynomials. This is a different choice of phase. We addressed that following table 11–1 in comments on spherical harmonics. We choose to include a factor of  $(-1)^m$  with the associated Legendre polynomials, and the sign of all spherical harmonics will be positive as a result.

Finally, remember the change of variables  $x = \cos \theta$ . That was done to put the differential equation in a more elementary form. In fact, a dominant use of associated Legendre polynomials is in applications where the argument is  $\cos \theta$ . One example is the generating function for spherical harmonic functions,

$$Y_{l,m}(\theta, \phi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_{l,m}(\cos \theta) e^{im\phi} \quad m \geq 0, \quad (10-10)$$

and

$$Y_{l,-m}(\theta, \phi) = Y_{l,m}^*(\theta, \phi), \quad m < 0,$$

where the  $P_{l,m}(\cos \theta)$  are associated Legendre polynomials. If we need a spherical harmonic with  $m < 0$ , we will calculate the spherical harmonic with  $m = |m|$ , and then calculate the adjoint.

To summarize the last three sections, we separated the angular equation into an azimuthal and a polar portion. The solutions to the azimuthal angle equation are exponentials including the magnetic moment quantum number in the argument. The solutions to the polar angle equation are the associated Legendre polynomials, which are different for each choice of orbital angular momentum and magnetic moment quantum number. Both quantum numbers are introduced into

the respective differential equations as separation constants. Since we assumed a product of the two functions to get solutions to the azimuthal and polar parts, the solutions to the original angular equation (10–7) are the products of the two solutions  $P_{l,m}(\cos \theta) e^{im\phi}$ . These factors are included in equation (10–10). All other factors in equation (10–12) are simply normalization constants. The products  $P_{l,m}(\cos \theta) e^{im\phi}$  are the spherical harmonic functions, the alternating sign and radical just make the orthogonal set orthonormal.

## Associated Laguerre Polynomials and Functions

The azimuthal equation was easy, the polar angle equation a little more substantial, but you will likely perceive the solution to the radial equation as plain, old heavy! There is no easy way to do this. Our approach will be to relate the radial equation to the associated Laguerre equation, for which the associated Laguerre functions are solutions. A popular option to solve the radial equation is a power series solution, for which we will refer you to Griffiths<sup>3</sup>, or Cohen–Tannoudji<sup>4</sup>.

Laguerre polynomials are solutions to the Laguerre equation

$$x L_j''(x) + (1 - x) L_j'(x) + j L_j(x) = 0.$$

The first few Laguerre polynomials are listed in table 10–3.

$L_0(x) = 1$ $L_1(x) = -x + 1$ $L_2(x) = x^2 - 4x + 2$ $L_3(x) = -x^3 + 9x^2 - 18x + 6$ $L_4(x) = x^4 - 16x^3 + 72x^2 - 96x + 24$ $L_5(x) = -x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120$ $L_6(x) = x^6 - 36x^5 + 450x^4 - 2400x^3 + 5400x^2 - 4320x + 720$
--

Table 10 – 3. The First Seven Laguerre Polynomials.

Laguerre polynomials of any order can be calculated using the generating function

$$L_j(x) = e^x \frac{d^j}{dx^j} e^{-x} x^j.$$

The Laguerre polynomials do not form an orthogonal set. The related set of **Laguerre functions**,

$$\phi_j(x) = e^{-x/2} L_j(x) \tag{10 – 13}$$

is orthonormal on the interval  $0 \leq x < \infty$ . The Laguerre functions are not solutions to the Laguerre equation, but are solutions to an equation which is related.

Just as the Legendre equation becomes the associated Legendre equation by adding an appropriate term containing a second index, the associated Laguerre equation is

$$x L_j^{k''}(x) + (1 - x + k) L_j^{k'}(x) + j L_j^k(x) = 0, \tag{10 – 14}$$

<sup>3</sup> Griffiths, *Introduction to Quantum Mechanics* (Prentice Hall, Englewood Cliffs, New Jersey, 1995), pp. 134–141.

<sup>4</sup> Cohen–Tannoudji, Diu, and Laloe, *Quantum Mechanics* (John Wiley & Sons, New York, 1977), pp. 794–797.



which reduces to the Laguerre equation when  $k = 0$ . The first few associated Laguerre polynomials are

$L_0^0(x) = L_0(x)$	$L_0^2(x) = 2$
$L_1^0(x) = L_1(x)$	$L_3^0(x) = L_3(x)$
$L_1^1(x) = -2x + 4$	$L_3^1(x) = -4x^3 + 48x^2 - 144x + 96$
$L_0^1(x) = 1$	$L_2^3(x) = 60x^2 - 600x + 1200$
$L_2^0(x) = L_2(x)$	$L_3^3(x) = -120x^3 + 2160x^2 - 10800x + 14400$
$L_2^1(x) = 3x^2 - 18x + 18$	$L_2^3(x) = -20x^3 + 300x^2 - 1200x + 1200$
$L_2^2(x) = 12x^2 - 96x + 144$	$L_1^3(x) = -24x + 96$
$L_1^2(x) = -6x + 18$	$L_0^3(x) = 6$

Table 10 – 4. Some Associated Laguerre Polynomials.

Notice  $L_j^0 = L_j$ . Also notice the indices are all non-negative, and either index may assume any integral value. We will be interested only in those associated Laguerre polynomials where  $k < j$  for hydrogen atom wave functions.

Associated Laguerre polynomials can be calculated from Laguerre polynomials using the generating function

$$L_j^k(x) = (-1)^k \frac{d^k}{dx^k} L_{j+k}(x).$$

---

**Example 10–5:** Calculate  $L_3^1(x)$  starting with the generating function.

We first need to calculate  $L_4(x)$ , because

$$L_j^k(x) = (-1)^k \frac{d^k}{dx^k} L_{j+k}(x) \Rightarrow L_3^1(x) = (-1)^1 \frac{d^1}{dx^1} L_{3+1}(x) = -\frac{d}{dx} L_4(x).$$

Similarly, if you want to calculate  $L_3^2$ , you need to start with  $L_5$ , and to calculate  $L_4^3$ , you need to start with  $L_7$ . So using the generating function,

$$\begin{aligned} L_4(x) &= e^x \frac{d^4}{dx^4} e^{-x} x^4 \\ &= e^x \frac{d^3}{dx^3} \left( -e^{-x} x^4 + e^{-x} 4x^3 \right) \\ &= e^x \frac{d^2}{dx^2} \left( e^{-x} x^4 - e^{-x} 4x^3 - e^{-x} 4x^3 + e^{-x} 12x^2 \right) = e^x \frac{d^2}{dx^2} \left( e^{-x} x^4 - e^{-x} 8x^3 + e^{-x} 12x^2 \right) \\ &= e^x \frac{d}{dx} \left( -e^{-x} x^4 + e^{-x} 4x^3 + e^{-x} 8x^3 - e^{-x} 24x^2 - e^{-x} 12x^2 + e^{-x} 24x \right) \\ &= e^x \frac{d}{dx} \left( -e^{-x} x^4 + e^{-x} 12x^3 - e^{-x} 36x^2 + e^{-x} 24x \right) \\ &= e^x \left( e^{-x} x^4 - e^{-x} 4x^3 - e^{-x} 12x^3 + e^{-x} 36x^2 + e^{-x} 36x^2 - e^{-x} 72x - e^{-x} 24x + e^{-x} 24 \right) \\ &= e^x e^{-x} \left( x^4 - 16x^3 + 72x^2 - 96x + 24 \right) \\ &= x^4 - 16x^3 + 72x^2 - 96x + 24, \end{aligned}$$

per table 10–4. Then to get  $L_3^1(x)$ ,

$$\begin{aligned} L_3^1 &= -\frac{d}{dx} L_4(x) \\ &= -\frac{d}{dx} (x^4 - 16x^3 + 72x^2 - 96x + 24) \\ &= -(4x^3 - 48x^2 + 144x - 96) \\ &= -4x^3 + 48x^2 - 144x + 96, \end{aligned}$$

per table 10–3.

Associated Laguerre polynomials are not orthogonal but **associated Laguerre functions** of the type

$$\Phi_j^k(x) = e^{-x/2} x^{k/2} L_j^k(x)$$

are orthogonal on the interval  $0 \leq x < \infty$ , so can be made an orthonormal set. Again, the  $\Phi_j^k(x)$  are not solutions to the associated Laguerre equation but are solutions to a related equation.

We are specifically interested in a slightly different associated Laguerre function than the usual first choice indicated above, *i.e.*, we are interested in

$$y_j^k(x) = e^{-x/2} x^{(k+1)/2} L_j^k(x). \quad (10-15)$$

These are also not solutions to the associated Laguerre equation, but they are solutions to

$$y_j^{k''}(x) + \left( -\frac{1}{4} + \frac{2j+k+1}{2x} - \frac{k^2-1}{4x^2} \right) y_j^k(x) = 0. \quad (10-16)$$

The reason for our interest in (10–16) and its solutions (10–15), is that equation (10–16) is a form of the radial equation, so the radial functions  $R(r)$  we seek are  $R_{n,l}(r) = A y_n^l(r)$ , where  $A$  is simply a normalization constant.

**Example 10–6:** Show equation (10–15) satisfies equation (10–16).

Unlike some of the toy problems given as examples, this example is a critical connection...unless you take our word for it, and then you should skip this. We are going to use the result of this example as a direct link to the solution of the radial equation. We are going to simplify the notation to minimize clutter, and will explain as we go.

To attain the second derivative, we need the first derivative, and use the notation

$$y = e^{-x/2} x^{(k+1)/2} v,$$

for equation (10–15) where  $v = L_j^k(x)$ , because the indices do not change and only serve to add clutter, and we can remember the independent variable is  $x$ . The first derivative is

$$\begin{aligned} y' &= -\frac{1}{2} e^{-x/2} x^{(k+1)/2} v + e^{-x/2} \left( \frac{k+1}{2} \right) x^{(k-1)/2} v + e^{-x/2} x^{(k+1)/2} v' \\ &= \left[ -\frac{1}{2} v + \left( \frac{k+1}{2x} \right) v + v' \right] e^{-x/2} x^{(k+1)/2} \\ \Rightarrow \left( e^{x/2} x^{-(k+1)/2} \right) y' &= -\frac{1}{2} v + \frac{k+1}{2x} v + v'. \end{aligned}$$

Notice we adjusted the second term on the right to do the factoring. Using the same adjustment technique, will factor these terms out of the second derivative as we go. These are also factors common to equation (10–15). Since the right side of equation (10–16) is zero, after we substitute the second derivative and the function into (10–16), we will simplify the equation by dividing the equation by common factors, therefore, none of the common factors will enter into the final solution. The exponentials and powers still need to be considered in differentiation, but their inverses will appear on the left and only the terms which have impact will appear on the right. Proceeding....

$$\left(e^{x/2}x^{-(k+1)/2}\right) y'' = \frac{1}{4}v - \frac{1}{2}\frac{k+1}{2x}v - \frac{1}{2}v' - \frac{1}{2}\frac{k+1}{2x}v + \frac{k+1}{2x}\frac{k-1}{2x}v + \frac{k+1}{2x}v' - \frac{1}{2}v' + \frac{k+1}{2x}v' + v''.$$

Substituting the second derivative and the function into equation (10–16),

$$y'' + \left(-\frac{1}{4} + \frac{2j+k+1}{2x} - \frac{k^2-1}{4x^2}\right)y = 0,$$

and dividing by the common factor of  $e^{-x/2}x^{(k+1)/2}$ , the remaining terms are

$$\begin{aligned} & \left(\frac{1}{4}v - \frac{1}{2}\frac{k+1}{2x}v - \frac{1}{2}v' - \frac{1}{2}\frac{k+1}{2x}v + \frac{k+1}{2x}\frac{k-1}{2x}v + \frac{k+1}{2x}v' - \frac{1}{2}v' + \frac{k+1}{2x}v' + v''\right) \\ & \quad + \left(-\frac{1}{4} + \frac{2j+k+1}{2x} - \frac{k^2-1}{4x^2}\right)v = 0 \\ \Rightarrow & \quad v'' + \cancel{\frac{1}{4}}v - \frac{1}{2}\frac{k+1}{2x}v - \frac{1}{2}v' - \frac{1}{2}\frac{k+1}{2x}v + \cancel{\frac{k^2-1}{4x^2}}v + \frac{k+1}{2x}v' - \frac{1}{2}v' + \frac{k+1}{2x}v' \\ & \quad - \cancel{\frac{1}{4}}v + \frac{2j+k+1}{2x}v - \cancel{\frac{k^2-1}{4x^2}}v = 0 \\ \Rightarrow & \quad v'' - \cancel{\frac{k-1}{2x}}v - \frac{1}{2}v' - \cancel{\frac{k-1}{2x}}v + \frac{k+1}{x}v' - \frac{1}{2}v' + \frac{j}{x}v + \cancel{\frac{k-1}{2x}}v = 0 \\ \Rightarrow & \quad v'' - v' + \frac{k+1}{x}v' + \frac{j}{x}v = 0 \\ \Rightarrow & \quad xv'' - xv' + (k+1)v' + jv = 0 \\ \Rightarrow & \quad xv'' + (1-x+k)v' + jv = 0 \end{aligned} \tag{10-17}$$

which is the associated Laguerre equation. Since  $v = L_j^k(x)$ , and the  $L_j^k(x)$  are solutions to the associated Laguerre equation, equation (10–17) is equivalent to

$$xL_j^{k''}(x) + (1-x+k)L_j^{k'}(x) + jL_j^k(x) = 0,$$

which is the associated Laguerre, which we know to be a true statement, so

$$\begin{aligned} y &= e^{-x/2}x^{(k+1)/2}v \\ &= e^{-x/2}x^{(k+1)/2}L_j^k(x) \end{aligned}$$

are solutions to equation (10–16).

---

## The Reduced Mass

Equation (10–2) describes a single particle in a central potential. The hydrogen atom is a two body problem, and the potential is not central but is dependent upon the distance between the nucleus and the electron. Were we able to anchor the nucleus to a stationary location we could designate an origin, equation (10–2) would be an accurate description. This is not possible, but we can reach a similar end by picturing the center of mass being anchored to a fixed location. If we use the **reduced mass** in place of the electron mass,

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

the radial coordinate  $r$  accurately describes the distance between the nucleus and the electron. The effect in equation (10–2) is cosmetic; where there was an  $m$  representing  $m_e$ , it is replaced by  $\mu$ . Because the proton is about 1836 times more massive than the electron, the reduced mass is nearly identically the electron mass. Many authors simply retain the electron mass. Since the center of mass is not actually anchored, a second set of coordinates is required to track the center of mass using this scheme. This consideration and other details of reducing a two particle problem to a one particle problem are adequately covered in numerous texts, including Chohen–Tannoudji<sup>5</sup>, Levine<sup>6</sup>, and many classical mechanics texts.

## Solution of the Radial Equation

The radial equation (10–6) using the reduced mass and the Coulomb potential,  $V(r) = -e^2/r$ , is

$$\begin{aligned} & \frac{1}{R(r)} \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) R(r) - \frac{2\mu r^2}{\hbar^2} \left[ -\frac{e^2}{r} - E \right] - l(l+1) = 0 \\ \Rightarrow & \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) R(r) - \frac{2\mu r^2}{\hbar^2} \left[ -\frac{e^2}{r} - E \right] R(r) - l(l+1) R(r) = 0 \\ \Rightarrow & \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) R(r) + \left[ \frac{2\mu r^2}{\hbar^2} \frac{e^2}{r} + \frac{2\mu r^2}{\hbar^2} E - l(l+1) \right] R(r) = 0. \end{aligned} \quad (10-18)$$

The plan is to get (10–18) into a form comparable to equation (10–16), and we already know the solutions are equation (10–15). We will be able to glean additional information by comparing the equations term by term. The energy levels of the hydrogen atom and the meaning of the indices of the associated Laguerre polynomials, which will be quantum numbers for the hydrogen atom, will come from the comparison of individual terms.

We will make three substitutions to get the last equation into the form of equation (10–16). The first is

$$y(r) = r R(r) \quad \Rightarrow \quad R(r) = \frac{y(r)}{r}. \quad (10-19)$$

---

<sup>5</sup> Cohen–Tannoudji, Diu, and Laloe, *Quantum Mechanics* (John Wiley & Sons, New York, 1977), pp. 784–788.

<sup>6</sup> Levine, *Quantum Chemistry* (Allyn and Bacon, Inc., Boston, Massachusetts, 1983), pp. 101–106.

Making this substitution in the first term and evaluating the derivatives

$$\begin{aligned}
\frac{d}{dr} \left( r^2 \frac{d}{dr} \right) R(r) &= \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) (r^{-1}) y(r) \\
&= \frac{d}{dr} r^2 \left[ (-r^{-2}) y(r) + (r^{-1}) \frac{dy(r)}{dr} \right] \\
&= \frac{d}{dr} \left[ -y(r) + r \frac{dy(r)}{dr} \right] \\
&= -\frac{dy(r)}{dr} + \frac{dy(r)}{dr} + r \frac{d^2 y(r)}{dr^2} \\
&= r \frac{d^2 y(r)}{dr^2}.
\end{aligned}$$

The substitution serves to eliminate the first derivative. We would have both a first and second derivative if we had evaluated the first term using  $R(r)$ . With this and the substitution of equation (10-19), equation (10-18) becomes

$$\begin{aligned}
r \frac{d^2 y(r)}{dr^2} + \left[ \frac{2\mu r e^2}{\hbar^2} + \frac{2\mu r^2}{\hbar^2} E - l(l+1) \right] \frac{y(r)}{r} &= 0 \\
\Rightarrow \frac{d^2 y(r)}{dr^2} + \left[ \frac{2\mu e^2}{r \hbar^2} + \frac{2\mu E}{\hbar^2} - \frac{l(l+1)}{r^2} \right] y(r) &= 0.
\end{aligned}$$

The second substitution is essentially to simplify the notation, and is

$$\left( \frac{\epsilon}{2} \right)^2 = -\frac{2\mu E}{\hbar^2} \tag{10-20}$$

where the negative sign on the right indicates we are looking for bound states, states such that  $E < 0$ , so including the negative sign here lets us have an  $\epsilon$  which is real. The last equation becomes

$$\frac{d^2 y(r)}{dr^2} + \left[ \frac{2\mu e^2}{r \hbar^2} - \frac{\epsilon^2}{4} - \frac{l(l+1)}{r^2} \right] y(r) = 0.$$

The third substitution is a change of variables, and notice it relates radial distance and energy through equation (10-20),

$$x = r\epsilon \Rightarrow r = \frac{x}{\epsilon}, \tag{10-21}$$

$$\Rightarrow dr = \frac{dx}{\epsilon} \Rightarrow \frac{d^2 y(r)}{dr^2} = \frac{d}{dr} \frac{dy(r)}{dr} = \epsilon \frac{d}{dx} \epsilon \frac{dy(x)}{dx} = \epsilon^2 \frac{d^2 y(x)}{dx^2},$$

so our radial equation becomes

$$\begin{aligned}
\epsilon^2 \frac{d^2 y(x)}{dx^2} + \left[ \frac{2\mu e^2 \epsilon}{x \hbar^2} - \frac{\epsilon^2}{4} - \epsilon^2 \frac{l(l+1)}{x^2} \right] y(x) &= 0 \\
\Rightarrow \frac{d^2 y(x)}{dx^2} + \left[ -\frac{1}{4} + \frac{2\mu e^2}{\hbar^2 \epsilon x} - \frac{l(l+1)}{x^2} \right] y(x) &= 0,
\end{aligned} \tag{10-22}$$

and equation (10-22) is equation (10-16) where

$$l(l+1) = \frac{k^2 - 1}{4}, \tag{10-23}$$

and

$$\frac{2\mu e^2}{\hbar^2 \epsilon} = \frac{2j + k + 1}{2}, \quad (10 - 24)$$

Per example 10–6, the solutions are equation (10–15),

$$y_j^k(x) = e^{-x/2} x^{(k+1)/2} L_j^k(x).$$

## Eigenenergies from the Solution of the Radial Equation

Equation (10–23) tells us  $k = 2l + 1$ .

**Example 10–7:** Show  $k = 2l + 1$ .

Equation (10–23) is

$$\begin{aligned} \frac{k^2 - 1}{4} &= l(l + 1) \\ \Rightarrow k^2 &= 4l(l + 1) + 1 \\ &= 4l^2 + 4l + 1 \\ &= (2l + 1)^2 \\ \Rightarrow k &= 2l + 1. \end{aligned}$$

We are going to take what appears to be a slight diversion to evaluate a particular set of factors in equation (10–24),  $\hbar^2/\mu e^2$ , which recurs repeatedly. Going back to the old quantum theory, this is called the Bohr radius, that is

$$a_0 = \frac{\hbar^2}{\mu e^2} = 0.529 \text{ \AA}. \quad (10 - 25)$$

We want to express lengths in terms of the Bohr radius because it is a natural length for the hydrogen atom.

**Example 10–8:** Show  $a_0 = 0.529 \text{ \AA}$ , using both the electron mass and the reduced mass.

This example is intended to illustrate three simple things. First,  $a_0 = 0.529 \text{ \AA}$ , second is to work out the CGS units for  $e^2$ , and then to show the electron mass is a very good approximation to the reduced mass in hydrogen. The electrostatic force in MKS and CGS systems is defined

$$\begin{aligned} \frac{1}{4\pi\epsilon_0} \frac{e_{\text{MKS}}^2}{r^2} &= F = \frac{e_{\text{CGS}}^2}{r^2} \\ \Rightarrow e_{\text{CGS}}^2 &= \frac{e_{\text{MKS}}^2}{4\pi\epsilon_0} = (1.602 \times 10^{-19} \text{ C})^2 (8.988 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2) \\ &= 2.307 \times 10^{-28} \text{ N} \cdot \text{m}^2 = 2.307 \times 10^{-19} \text{ dyne} \cdot \text{cm}^2 \\ &= 2.307 \times 10^{-19} \text{ erg} \cdot \text{cm} = 14.42 \text{ eV} \cdot \text{\AA}. \end{aligned}$$

So

$$a_0 = \frac{\hbar^2}{m_e e^2} = \frac{1}{4\pi^2} \frac{(hc)^2}{m_e c^2 e^2} = \frac{1}{4\pi^2} \frac{(1.240 \times 10^4 \text{ eV} \cdot \text{\AA})^2}{(0.5110 \times 10^6 \text{ eV})(14.42 \text{ eV} \cdot \text{\AA})} = 0.5286 \text{ \AA}.$$

The reduced mass is

$$\mu = \frac{m_p m_e}{m_p + m_e} = \frac{1.673 \times 10^{-24}}{1.673 \times 10^{-24} + 9.110 \times 10^{-28}} m_e = 0.9995 m_e,$$

$$\Rightarrow \frac{\hbar^2}{\mu e^2} = \frac{\hbar^2}{0.9995 m_e e^2} = \frac{0.5286 \text{ \AA}}{0.9995} = 0.5288 \text{ \AA}.$$

There is 0.03% difference between the electron mass and reduced mass values. Many authors simply use the electron mass and it yields a good approximation. The CGS value of  $e^2$  can be mysterious for those who have worked primarily in MKS units. By the way, the square root  $e = 3.797 (eV \cdot \text{\AA})^{1/2}$  can be a convenient way to express the charge on the electron in CGS units.

---

Equation (10–24) gives us the eigenenergies of the hydrogen atom, but requires some development. Since  $k = 2l + 1$ ,

$$\frac{2j + k + 1}{2} = \frac{2j + (2l + 1) + 1}{2} = j + l + 1.$$

From the discussion on associated Laguerre polynomials, the indices  $j$  and  $k$  are non-negative. The sum  $j + l + 1$  can, therefore, assume any integer values of 1 or greater. We are going to rename it  $n$ , or

$$n = j + l + 1. \quad (10 - 26)$$

The new integer index  $n$  is known as the **principal quantum number**. Using the principal quantum number, it follows that the eigenenergies of the hydrogen atom are

$$E_n = -\frac{\hbar^2}{2\mu a_0^2 n^2} = -\frac{13.6 \text{ eV}}{n^2}, \quad (10 - 27)$$

where the quantity  $13.6 \text{ eV}$  is called the Rydberg, usually denoted  $R$  or  $Ry$ . The ground state energy is  $E_0 = -13.6 \text{ eV}$  when  $n = 1$ . It is often convenient to express excited state energies in terms of the ground state energy.

---

**Example 10–9:** Show equation (10–27) follows from equation (10–24).

$$\frac{2\mu e^2}{\hbar^2 \epsilon} = \frac{2j + k + 1}{2}$$

$$\Rightarrow j + l + 1 = n = \frac{2\mu e^2}{\hbar^2 \epsilon}$$

$$\Rightarrow \epsilon = \frac{2\mu e^2}{\hbar^2 n}$$

$$\Rightarrow \epsilon^2 = \frac{4\mu^2 e^4}{\hbar^4 n^2}.$$

Substituting equation (10–20) to eliminate  $\epsilon$  and insert energy,

$$-4 \frac{2\mu E}{\hbar^2} = \frac{4\mu^2 e^4}{\hbar^4 n^2}$$

$$\begin{aligned}\Rightarrow E &= -\frac{\mu^2 e^4 \hbar^2}{2\mu \hbar^4 n^2} = -\left(\frac{\mu e^2}{\hbar^2}\right)^2 \frac{\hbar^2}{2\mu n^2} \\ \Rightarrow E_n &= -\frac{\hbar^2}{2\mu a_0^2 n^2}.\end{aligned}$$

Inserting numerical values,

$$\begin{aligned}E_n &= -\frac{\hbar^2}{2\mu a_0^2 n^2} = -\frac{1}{4\pi^2} \frac{(hc)^2}{2(\mu c^2) a_0^2 n^2} \\ &= -\frac{1}{4\pi^2} \frac{(1.24 \times 10^4 \text{ eV} \cdot \text{\AA})^2}{2(0.511 \times 10^6 \text{ eV})(0.529 \text{ \AA})^2 n^2} = -\frac{13.6 \text{ eV}}{n^2},\end{aligned}$$

so eigenenergies do follow from the solution of the radial equation.

---



## Radial Wave Functions from the Solution of the Radial Equation

Again per example 10–6, the solutions of the radial equation are equation (10–15),

$$y_j^k(x) = e^{-x/2} x^{(k+1)/2} L_j^k(x).$$

Since

$$k = 2l + 1, \quad (k + 1)/2 = ((2l + 1) + 1)/2 = l + 1, \quad \text{and } j + l + 1 = n \Rightarrow j = n - l - 1,$$

we can now express the indices in terms of the quantum numbers  $n$  and  $l$ , meaning

$$y_n^l(x) = e^{-x/2} x^{l+1} L_{n-l-1}^{2l+1}(x),$$

so the solution starts to assume some of the traits of the hydrogen atom. The independent variable is still  $x = \epsilon r$ . We need an independent variable  $r$ , or at least in terms of  $r$ , to be consistent with the spherical coordinate system. Using (10–20) and (10–27), we can solve for  $\epsilon$  in terms of the Bohr radius and the principal quantum number,

$$\begin{aligned} \left(\frac{\epsilon}{2}\right)^2 &= -\frac{2\mu E}{\hbar^2} = -\frac{2\mu}{\hbar^2} \left(-\frac{\hbar^2}{2\mu a_0^2 n^2}\right) = \frac{1}{a_0^2 n^2} \\ \Rightarrow \epsilon^2 &= \frac{4}{a_0^2 n^2} \Rightarrow \epsilon = \frac{2}{a_0 n}, \\ \Rightarrow x &= \frac{2r}{na_0}, \end{aligned}$$

and this is in terms of the desired independent variable, or

$$y_n^l(r) = e^{-r/na_0} \left(\frac{2r}{na_0}\right)^{l+1} L_{n-l-1}^{2l+1}\left(\frac{2r}{na_0}\right).$$

Finally we can relate the associated Laguerre function to the radial function we set out to find through equation (10–19). Remember  $y(r) = r R(r)$ , so

$$\begin{aligned} r R_{n,l}(r) &= e^{-r/na_0} \left(\frac{2r}{na_0}\right)^{l+1} L_{n-l-1}^{2l+1}\left(\frac{2r}{na_0}\right) \\ \Rightarrow R_{n,l}(r) &= A e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{na_0}\right), \end{aligned}$$

where we have added a normalization constant which has absorbed the factor  $2/na_0$  from the power term when we cancel the factor of  $r$ . This still needs to be normalized. We want the radial functions to be individually normalized so the product of the radial wave functions and spherical harmonics, the hydrogen wave function, is normalized. The normalized radial wave functions are

$$R_{n,l}(r) = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{na_0}\right). \quad (10-28)$$

The first few normalized radial wave functions are listed in table 10–5.

$R_{1,0}(r) = 2a_0^{-3/2} e^{-r/a_0}$ $R_{2,0}(r) = \frac{1}{\sqrt{2}} a_0^{-3/2} \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0}$ $R_{2,1}(r) = \frac{1}{\sqrt{24}} a_0^{-3/2} \frac{r}{a_0} e^{-r/2a_0}$ $R_{3,0}(r) = \frac{2}{\sqrt{27}} a_0^{-3/2} \left(1 - \frac{2r}{3a_0} + \frac{2r^2}{27a_0^2}\right) e^{-r/3a_0}$ $R_{3,1}(r) = \frac{8}{27\sqrt{6}} a_0^{-3/2} \left(1 - \frac{r}{6a_0}\right) \frac{r}{a_0} e^{-r/3a_0}$ $R_{3,2}(r) = \frac{4}{81\sqrt{30}} a_0^{-3/2} \frac{r^2}{a_0^2} e^{-r/3a_0}$
<p>Table 10 – 5. The First Six Radial Wave Functions for Hydrogen.</p>

**Example 10–10:** Show equation (10–28) is properly normalized.

We want to normalize

$$R_{n,l}(r) = A e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right).$$

The normalization condition for a function of the radial variable in spherical coordinates is

$$\langle \psi(r) | \psi(r) \rangle = 1 = \int_0^\infty (R_{n,l}(r))^* R_{n,l}(r) r^2 dr$$

where the factor of  $r^2$  is the radial portion of the volume element in spherical coordinates. Using the radial wave functions, this is

$$\begin{aligned} 1 &= \int_0^\infty \left( A e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) \right)^* A e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) r^2 dr \\ &= \int_0^\infty A^* A, e^{-2r/na_0} \left(\frac{2r}{na_0}\right)^{2l} (r^2) L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) dr \\ &= |A|^2 \left(\frac{na_0}{2}\right)^3 \int_0^\infty e^{-2r/na_0} \left(\frac{2r}{na_0}\right)^{2l+2} L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) d\left(\frac{2r}{na_0}\right) \end{aligned} \quad (10-29)$$

where the cubic factor preceding the integral comes from the fact we have grouped the factor of  $r^2$  with the other powers of  $r$ , and also have expressed the differential in terms of the argument of the associated Laguerre polynomial. We are going to borrow two relations from Morse and Feshbach<sup>7</sup>, the first of which is

$$\int_0^\infty z^a e^{-z} L_b^a(z) L_b^a(z) dz = \delta_{b,c} \frac{[\Gamma(a+b+1)]^3}{\Gamma(b+1)}. \quad (10-30)$$

<sup>7</sup> Morse and Feshbach, *Methods of Theoretical Physics* (McGraw–Hill, New York, 1953), pp. 784–785

There are some things to notice about equation (10-30). The Kronecker delta  $\delta_{b,c}$  reflects the orthogonality of the associated Laguerre polynomials. The integral is zero unless the lower indices of the two associated Laguerre polynomials are identical. Also, we expect all indices to be integers, so expect the gamma functions to become factorials. Comparing (10-29) and (10-30), the argument of the associated Laguerre polynomial, the base of the power term, and the differential are all the same in both equations. The upper index of the associated Laguerre polynomials and the power term are the same in (10-30), but differ in (10-29). We can adjust (10-29) using a recursion relation from Morse and Feshbach,

$$z L_b^a(z) = (a + 2b + 1) L_b^a(z) - \frac{b+1}{a+b+1} L_{b+1}^a(z) - (a+b)^2 L_{b-1}^a(z).$$

If we take one factor of the power term and group it with the first associated Laguerre polynomial, we get

$$1 = |A|^2 \left(\frac{na_0}{2}\right)^3 \int_0^\infty e^{-2r/na_0} \left(\frac{2r}{na_0}\right)^{2l+1} \left\{ \left(\frac{2r}{na_0}\right) L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) \right\} L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) d\left(\frac{2r}{na_0}\right),$$

which makes the exponent of the power term match the upper index of the associated Laguerre polynomial, and we can evaluate the term in braces using the recursion relation. For  $a = 2l + 1$  and  $b = c = n - l - 1$ , this is

$$\begin{aligned} \left(\frac{2r}{na_0}\right) L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) &= (2l+1+2n-2l-2+1) L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) \\ &\quad - \frac{n-l-1+1}{2l+1+n-l-1+1} L_{n-l}^{2l+1} \left(\frac{2r}{na_0}\right) - (2l+1+n-l-1)^2 L_{n-l-2}^{2l+1} \left(\frac{2r}{na_0}\right) \\ &= 2n L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) - \frac{n-l}{n+l+1} L_{n-l}^{2l+1} \left(\frac{2r}{na_0}\right) - (n+l)^2 L_{n-l-2}^{2l+1} \left(\frac{2r}{na_0}\right). \end{aligned}$$

We are going to ignore all but the first term. We could substitute all three terms into the integral, distribute other factors, break the integral into three integrals, and evaluate each one using equation (10-30). But the Kronecker delta tells us integrals with unequal lower associated Laguerre polynomial indices will be zero. The only non-vanishing integral will come from the first term, so

$$\begin{aligned} 1 &= |A|^2 \left(\frac{na_0}{2}\right)^3 \int_0^\infty e^{-2r/na_0} \left(\frac{2r}{na_0}\right)^{2l+1} \left\{ 2n L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) \right\} L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) d\left(\frac{2r}{na_0}\right), \\ &= |A|^2 2n \left(\frac{na_0}{2}\right)^3 \int_0^\infty e^{-2r/na_0} \left(\frac{2r}{na_0}\right)^{2l+1} L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right) d\left(\frac{2r}{na_0}\right), \end{aligned}$$

and we are ready to evaluate this integral using equation (10-30) where

$$z \rightarrow \left(\frac{2r}{na_0}\right), \quad a \rightarrow 2l+1, \quad b=c \rightarrow n-l-1,$$

and for integers  $\Gamma(j) = (j-1)!$ . Then

$$\begin{aligned} 1 &= |A|^2 2n \left(\frac{na_0}{2}\right)^3 \frac{[\Gamma(2l+1+n-l-1+1)]^3}{\Gamma(n-l-1+1)} \\ &= |A|^2 2n \left(\frac{na_0}{2}\right)^3 \frac{[\Gamma(n+l+1)]^3}{\Gamma(n-l)} \\ &= |A|^2 2n \left(\frac{na_0}{2}\right)^3 \frac{[(n+l)!]^3}{(n-l-1)!} \end{aligned}$$

$$\Rightarrow A = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}}$$

$$\Rightarrow R_{n,l}(r) = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0}\right).$$

**Example 10–11:** Calculate  $R_{3,1}(r)$  from equation (10–29).

For  $n = 3$ ,  $l = 1$ ,

$$\begin{aligned} R_{3,1}(r) &= \sqrt{\left(\frac{2}{3a_0}\right)^3 \frac{(3-1-1)!}{2(3)[(3+1)!]^3}} e^{-r/3a_0} \left(\frac{2r}{3a_0}\right)^1 L_{3-1-1}^{2(1)+1} \left(\frac{2r}{3a_0}\right) \\ &= \frac{2}{3} \sqrt{\left(\frac{2^3}{3^3 a_0^3}\right) \frac{1}{2 \cdot 3 [(4 \cdot 3 \cdot 2)]^3}} e^{-r/3a_0} \left(\frac{r}{a_0}\right) L_1^3 \left(\frac{2r}{3a_0}\right) \\ &= \frac{2}{3} \sqrt{\frac{2^2}{3^4 [2^6 \cdot 3^3 \cdot 2^3]}} a_0^{-3/2} e^{-r/3a_0} \left(\frac{r}{a_0}\right) \left(-24 \left(\frac{2r}{3a_0}\right) + 96\right) \\ &= 96 \frac{2}{3} \sqrt{\frac{2^2}{2^9 \cdot 3^7}} a_0^{-3/2} \left(-\left(\frac{r}{6a_0}\right) + 1\right) \frac{r}{a_0} e^{-r/3a_0}, \end{aligned}$$

and because  $96 = 2^5 \cdot 3$ ,

$$\begin{aligned} R_{3,1}(r) &= \frac{2^6}{2^3 \cdot 3^3} \sqrt{\frac{1}{2 \cdot 3}} a_0^{-3/2} \left(1 - \frac{r}{6a_0}\right) \frac{r}{a_0} e^{-r/3a_0} \\ &= \frac{8}{27\sqrt{6}} a_0^{-3/2} \left(1 - \frac{r}{6a_0}\right) \frac{r}{a_0} e^{-r/3a_0}, \end{aligned}$$

in agreement with table 10–5.

## Graphs of Radial Wave Functions

The radial wave functions listed in table 10–5 are plotted in figure 10–1. The radial wave functions are all individually normalized, so the probability associated with radial wave functions is

$$\langle \psi(r) | \psi(r) \rangle = \langle R_{n,l}(r) | R_{n,l}(r) \rangle = \int_0^\infty R_{n,l}^*(r) R_{n,l}(r) r^2 dr = \int_0^\infty R_{n,l}^2(r) r^2 dr,$$

where the factor of  $r^2$  in the integrals is from the volume element in spherical coordinates. The functions  $R_{n,l}^2(r) r^2$  are the radial probability densities for the hydrogen atom. The radial probability densities are plotted abreast their corresponding radial wave function in figure 10–2.

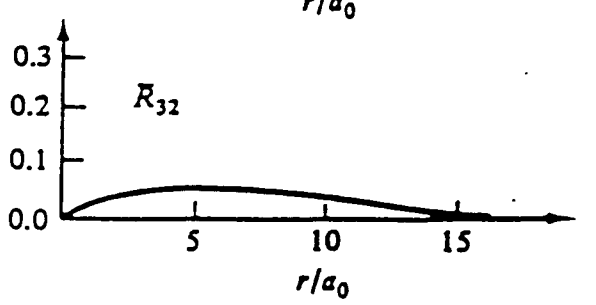
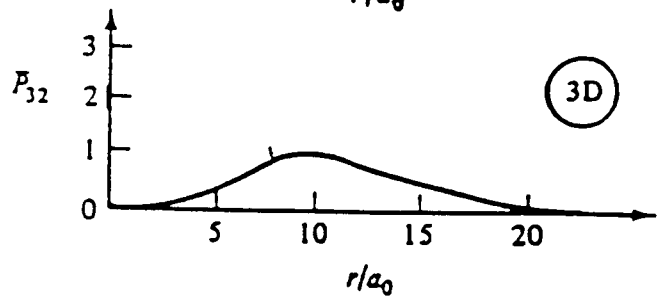
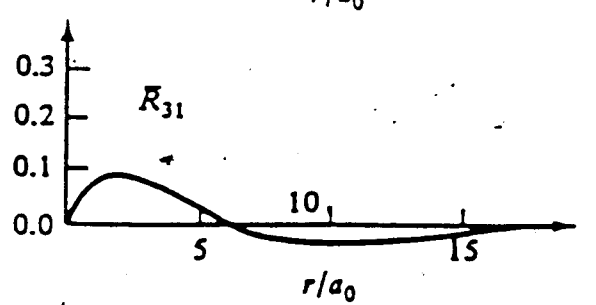
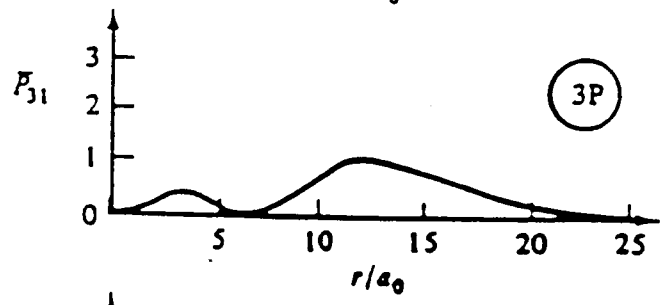
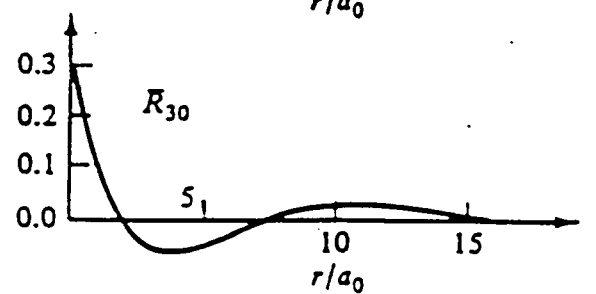
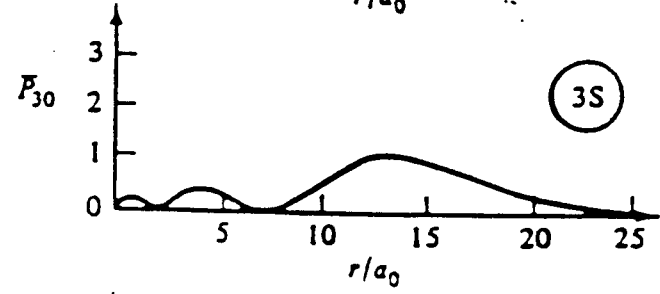
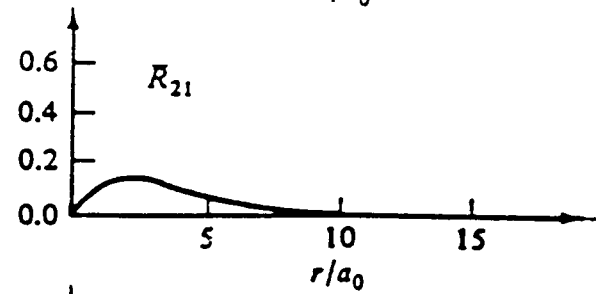
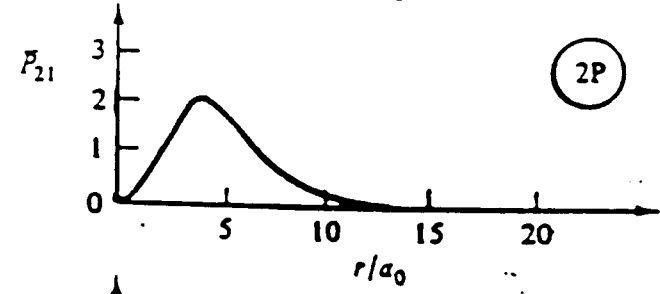
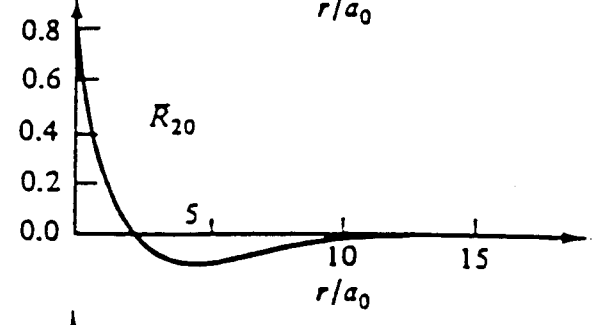
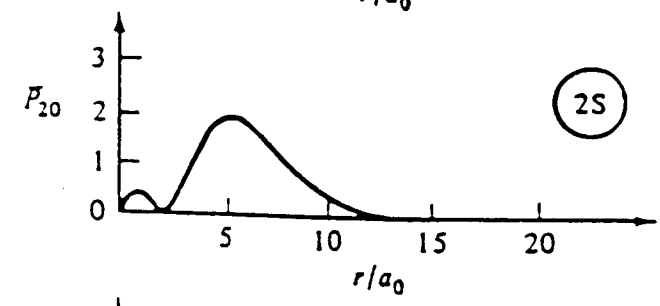
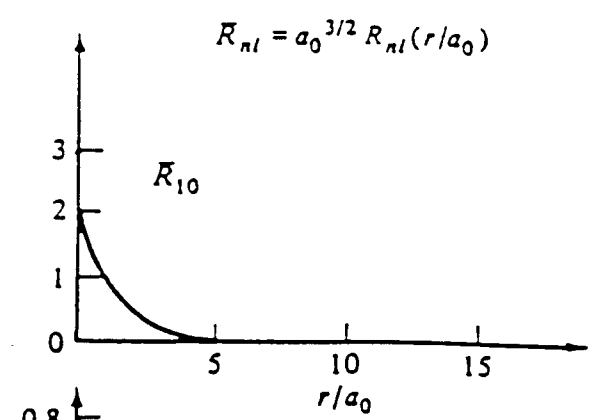
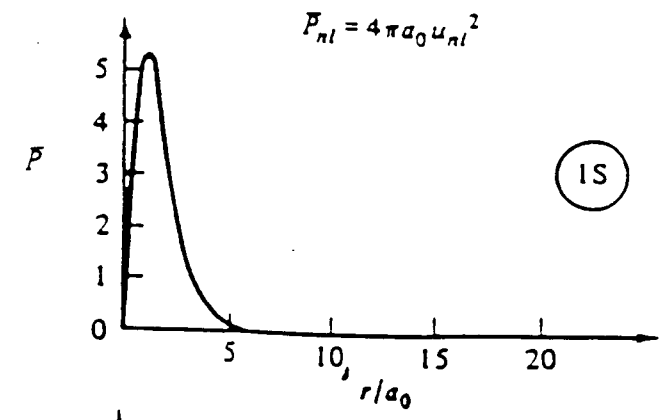


FIGURE 12-5.

FIG. 12-4

Figure 10 – 1. Radial Wave Functions.

Figure 10 – 2. Radial Probability Densities.

## Hydrogen Wave Functions

We have all three parts. The product of the azimuthal, polar, and radial wave functions are the hydrogen wave functions. The product of the azimuthal and polar wave functions are the spherical harmonics, so the hydrogen wave functions are the product of the radial wave functions and spherical harmonics,

$$\psi_{nlm}(r, \theta, \phi) = R_{n,l}(r)Y_{l,m}(\theta, \phi).$$

You can construct the desired hydrogen wave function by selecting the appropriate radial wave function from table 10–5, or calculating it from equation 10–28, and multiplying it by the appropriate spherical harmonic from table 11–1, or calculating it from equation 10–10, *i.e.*,

$$\psi_{nlm}(r, \theta, \phi) = R_{n,l}(r)Y_{l,m}(\theta, \phi).$$

The equation

$$\langle r, \theta, \phi | n, l, m \rangle = R_{n,l}(r)Y_{l,m}(\theta, \phi)$$

denotes representation from Hilbert space.

---

**Example 10–10:** What is  $\psi_{3,1,-1}(r, \theta, \phi)$ ?

From tables 10–5 and 11–1,

$$R_{3,1}(r) = \frac{8}{27\sqrt{6}}a_0^{-3/2} \left(1 - \frac{r}{6a_0}\right) \frac{r}{a_0} e^{-r/3a_0}, \quad Y_{1,-1}(\theta, \phi) = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\phi},$$

so

$$\begin{aligned} \psi_{3,1,-1}(r, \theta, \phi) &= R_{3,1}(r) Y_{1,-1}(\theta, \phi) \\ &= \frac{8}{27\sqrt{6}}a_0^{-3/2} \left(1 - \frac{r}{6a_0}\right) \frac{r}{a_0} e^{-r/3a_0} \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\phi} \\ &= \frac{8\sqrt{3}}{27\sqrt{6}\sqrt{8\pi}}a_0^{-3/2} \left(1 - \frac{r}{6a_0}\right) \frac{r}{a_0} e^{-r/3a_0} \sin \theta e^{-i\phi} \\ &= \frac{2}{27\sqrt{\pi}}a_0^{-3/2} \left(1 - \frac{r}{6a_0}\right) \frac{r}{a_0} \sin \theta e^{-r/3a_0-i\phi}. \end{aligned}$$


---

## Graphics of Hydrogen Wave Functions

The probability of finding the electron in a hydrogen atom in a specific volume is

$$\int_{\Delta V} |\psi_{nlm}(r, \theta, \phi)|^2 dV = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} \int_{\phi_1}^{\phi_2} |\psi_{nlm}(r, \theta, \phi)|^2 r^2 \sin \theta dr d\theta d\phi.$$

The probability of finding the electron of the hydrogen atom anywhere in space is non-zero, however, the probability is vanishingly small for most locations. The locations where the probability densities are highest are depicted in figure 10–3.

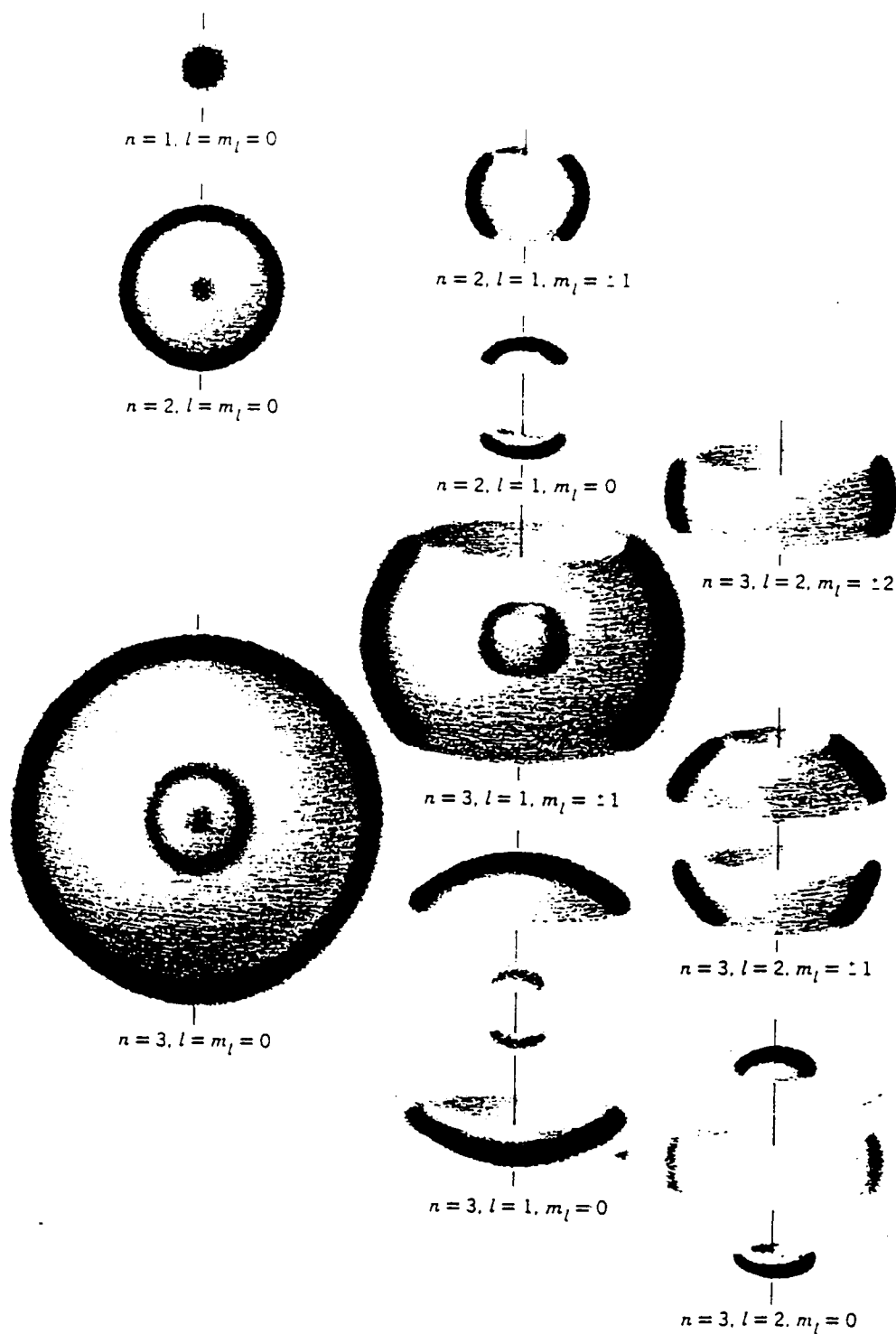


FIGURE 7-10

An artist's conception of the three-dimensional appearance of several one-electron atom probability density functions. For each of the drawings a line represents the  $z$  axis. If all the probability densities for a given  $n$  and  $l$  are combined, the result is spherically symmetrical.

Figure 10 – 3. Graphics for Significant Probability Densities by Quantum Number.

The graphics of figure 10–3 are illustrative and are a culmination of the calculations of this chapter. Additional calculations demonstrate the similarity of simpler systems to this realistic system. In many regards, the techniques developed earlier with simpler systems were examined so we could arrive at the hydrogen atom and figure 10–3.

## Quantum Numbers and Chemistry Terminology

Another depiction explains common chemistry terminology in terms of quantum numbers. The ground state of hydrogen is  $n = 1$ ,  $l = 0$  and  $m = 0$ . Common chemistry designations for the ground state are the K–shell and the  $1s$  subshell. Remember the K–shell may house two electrons in many electron atoms. This means the one electron in the hydrogen atom can exist in one of two possible states, though there is only one possible combination of the three quantum numbers  $n$ ,  $l$ , and  $m$ . There is, however, a fourth quantum number associated with spin. For each set of the three quantum numbers  $n$ ,  $l$ ,  $m$ , there is a possibility of spin up or spin down, and thus two possible states associated with the quantum numbers  $n = 1$ ,  $l = 0$ , and  $m = 0$ , therefore two states associated with the K–shell. We will examine spin in the next chapter. The point of the moment is the designation K–shell and quantum number  $n = 1$  have the same meaning. The quantum number  $n = 2$  is the same as the chemical designation of L–shell. The L–shell has subshells  $2s$  and  $2p$ , where  $s$  corresponds to  $l = 0$  and  $p$  corresponds to  $l = 1$ , which are possible quantum numbers for  $n = 2$ . For  $l = 1$ ,  $m = -1, 0$ , or  $1$  so there are three states possible in the  $p$  subshell. Each of the four L–shell states may have spin up or spin down, so there are eight possible eigenstates in the L–shell. The M–shell corresponds to  $n = 3$ , the N–shell corresponds to  $n = 4$ , and the O–shell corresponds to  $n = 5$ . As the quantum number  $n$  becomes larger, the quantum number  $l$  can take on a greater range of values. The quantum number  $l = 0$  corresponds to the  $s$ –subshell, the quantum number  $l = 1$  corresponds to the  $p$ –subshell, the quantum number  $l = 2$  corresponds to the  $d$ –subshell, the quantum number  $l = 3$  corresponds to the  $f$ –subshell, and the quantum number  $l = 4$  corresponds to the  $g$ –subshell. Each subshell has  $2l + 1$  states. The total states in each shell is  $n^2$ . Figure 10–4 summarizes the commentary in this discussion.

		shell					total	
							$n^2$	
	O	$n = 5$	$\underline{5s}_1$	$\underline{5p}_3$	$\underline{5d}_5$	$\underline{5f}_7$	$\underline{5g}_9$	25
	N	$n = 4$	$\underline{4s}_1$	$\underline{4p}_3$	$\underline{4d}_5$	$\underline{4f}_7$		16
↑	M	$n = 3$	$\underline{3s}_1$	$\underline{3p}_3$	$\underline{3d}_5$			9
E	L	$n = 2$	$\underline{2s}_1$	$\underline{2p}_3$				4
	K	$n = 1$	$\underline{1s}_1$					1
			$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$	

Figure 10 – 4. Quantum Number, Chemical Designation Correspondence.

The subscripts on the underbars of the subshells indicate the number of states  $2l + 1$  in that subshell.

One final comment. Since the energy, to this point, depends only on  $n$ , the quantum numbers  $l$  and  $m$  having no effect on energy, we have an  $n^2$ –fold degeneracy in energy using this picture.



A magnetic field removes this degeneracy...the same magnetic field that will define the  $z$ -axis and provide orientation to the scheme, will remove this degeneracy. This was explored by Pieter Zeeman, so is known as the Zeeman effect. Pragmatically, a field is necessary if the orientation is to make sense.

## Possibilities and Probabilities

In addition to being of intrinsic interest because it is a real system, the electron in a hydrogen atom is a particle in a fancy box. Comparable to a particle in the infinite square well or harmonic oscillator, the electron is confined. Also comparable to other systems, the state vector describing the electron in a hydrogen atom is a linear combination of eigenstates. We assume a limited linear combination to provide some sample calculations which should be reminiscent of calculations for other systems. The examples in that follow all refer to the  $t = 0$  state vector

$$\Psi(r, \theta, \phi) = 2\psi_{1,0,0} + \psi_{2,1,0} \quad (10-31)$$

**Example 10-13:** Normalize equation (10-31).

The state function is composed of 2 parts ground state and 1 part of the given excited state, so we can write

$$\Psi(r, \theta, \phi) = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

so the normalization condition can be written

$$\begin{aligned} 1 &= A^* (2, 1) A \begin{pmatrix} 2 \\ 1 \end{pmatrix} = |A|^2 (4 + 1) \\ &\Rightarrow A = \frac{1}{\sqrt{5}} \\ &\Rightarrow \Psi(r, \theta, \phi) = \frac{2}{\sqrt{5}} \psi_{1,0,0} + \frac{1}{\sqrt{5}} \psi_{2,1,0} \end{aligned}$$

is the properly normalized wave function or state vector.

**Example 10-14:** Normalize equation (10-31) using the orthonormality of the spherical harmonics.

This example demonstrates a useful technique using spherical harmonics as well as the reason the radial wave functions are individually normalized.

$$\begin{aligned} \Psi(r, \theta, \phi) &= 2\psi_{1,0,0} + \psi_{2,1,0} \\ &= 2R_{1,0}Y_{0,0} + R_{2,1}Y_{1,0}, \end{aligned}$$

and the normalization condition is

$$\begin{aligned} 1 &= \int_V A^* (2R_{1,0}Y_{0,0} + R_{2,1}Y_{1,0}) A (2R_{1,0}Y_{0,0} + R_{2,1}Y_{1,0}) dV \\ &= |A|^2 \int_V (4R_{1,0}^2 Y_{0,0}^2 + 4R_{2,1}R_{1,0}Y_{1,0}R_{1,0}Y_{0,0} + R_{2,1}^2 Y_{1,0}^2) dV \\ &= |A|^2 \left( \int_V 4R_{1,0}^2 Y_{0,0}^2 dV + \int_V 4R_{2,1}R_{1,0}Y_{1,0}Y_{0,0} dV + \int_V R_{2,1}^2 Y_{1,0}^2 dV \right) \end{aligned}$$

$$\begin{aligned}
&= |A|^2 \left( 4 \int_0^\infty R_{1,0}^2 r^2 dr \int_\Omega Y_{0,0}^2 d\Omega + 4 \int_0^\infty R_{2,1} R_{1,0} r^2 dr \int_\Omega Y_{1,0} Y_{0,0} d\Omega + \int_0^\infty R_{2,1}^2 r^2 dr \int_\Omega Y_{1,0}^2 d\Omega \right) \\
&= |A|^2 \left( 4(1)(1) + 4 \int_0^\infty R_{2,1} R_{1,0} r^2 dr (0) + (1)(1) \right) \tag{10-32}
\end{aligned}$$

where  $\int Y_{1,0} Y_{0,0} d\Omega$  is zero because the spherical harmonics are orthonormal, so

$$\begin{aligned}
1 &= |A|^2 (4 + 0 + 1) \Rightarrow A = \frac{1}{\sqrt{5}} \\
\Rightarrow \Psi(r, \theta, \phi) &= \frac{2}{\sqrt{5}} \psi_{1,0,0} + \frac{1}{\sqrt{5}} \psi_{2,1,0}.
\end{aligned}$$

The orthonormality of spherical harmonics means

$$\int Y_{i,j} Y_{k,l} d\Omega = \delta_{i,k} \delta_{j,l},$$

We have previously done this calculation using  $\langle i|j \rangle = \delta_{i,j}$ , which is the same statement for systems in a Hilbert space. If integrated over solid angle, the indices must be identical or the integral is zero, as illustrated in equation (10-32). The radial wave functions are not orthonormal, but they are normalized. This means

$$\int_0^\infty R_{i,j} R_{i,j} r^2 dr = 1.$$

Note the matching indices. If the indices match, the integral of their product and  $r^2$  over all possible values of radius is 1. If the indices do not match, you will often have the good fortune of a multiplication by zero by virtue of the integral of the spherical harmonics as in equation (10-32). In the circumstance the spherical harmonics are the same, the integral of radial functions will vanish. The bottom line is the hydrogen wave functions are orthonormal. The radial wave functions themselves, however, are normalized but they are not orthonormal.

A third method to calculate the normalization constant is to use the representations of the hydrogen wave functions and do the integrations explicitly. That method is much longer and more difficult than either of the previous two examples, and will yield the same result.

**Example 10-15:** What are the possible results of individual measurements of energy, angular momentum, and the  $z$ -component of angular momentum?

The possible results are the eigenvalues. For

$$\Psi(r, \theta, \phi) = \frac{2}{\sqrt{5}} \psi_{1,0,0} + \frac{1}{\sqrt{5}} \psi_{2,1,0}$$

we could measure a state with a principal quantum number of  $n = 1$  or  $n = 2$ . The corresponding eigenenergies are

$$E_n = -\frac{13.6 \text{ eV}}{n^2} \Rightarrow E_1 = -13.6 \text{ eV}, \text{ or } -\frac{13.6 \text{ eV}}{4} = -3.4 \text{ eV}$$

so are the possible results of a measurement of energy. For a measurement of angular momentum, we could measure a state with orbital angular momentum quantum number of  $l = 0$  or  $l = 1$ , so anticipate a measurement of

$$\sqrt{0(0+1)}\hbar = 0 \quad \text{or} \quad \sqrt{1(1+1)}\hbar = \sqrt{2}\hbar$$

because the eigenvalues of the square of angular momentum,  $\mathcal{L}^2$ , are  $l(l+1)\hbar$ . The eigenvalues of the  $z$ -component of angular momentum are  $m\hbar$ , and both eigenstates of the state vector have  $m = 0$ , so the only possibility of a measurement of the  $z$ -component of angular momentum is  $0\hbar = 0$ .

**Example 10–16:** What are the probabilities of the possible results of individual measurements of energy, angular momentum, and the  $z$ -component of angular momentum?

The probabilities are  $|\langle \psi | \Psi \rangle|^2$ , so the probability of measuring  $E_0 = -13.6 \text{ eV}$ , corresponding to measuring the eigenstate with  $n = 1$ , and the probability of measuring orbital angular momentum of  $0$ , corresponding to measuring the eigenstate with  $l = 0$ , is

$$\begin{aligned} P(E = -13.6 \text{ eV}) = P(\sqrt{\mathcal{L}^2} = 0) &= \left| \langle \psi_{1,0,0} \left| \frac{2}{\sqrt{5}}\psi_{1,0,0} + \frac{1}{\sqrt{5}}\psi_{2,1,0} \right. \right|^2 \\ &= \left| \frac{2}{\sqrt{5}} \langle \psi_{1,0,0} | \psi_{1,0,0} \rangle + \frac{1}{\sqrt{5}} \langle \psi_{1,0,0} | \psi_{2,1,0} \rangle \right|^2 \\ &= \left| \frac{2}{\sqrt{5}}(1) + 0 \right|^2 = \frac{4}{5} \end{aligned}$$

where we have used the orthonormality of eigenstates, regardless of whether we represent the eigenstates as functions of position and use the orthonormality of the spherical harmonics, or use a more general argument like  $\langle i | j \rangle = \delta_{ij}$ . Similarly, the probability of measuring  $E_1 = -3.4 \text{ eV}$ , corresponding to measurement of the eigenstate with  $n = 2$ , and the probability of measuring orbital angular momentum of  $\sqrt{2}\hbar$ , corresponding to measuring the eigenstate with  $l = 1$ , is

$$\begin{aligned} P(E = -3.4 \text{ eV}) = P(\sqrt{\mathcal{L}^2} = \sqrt{2}\hbar) &= \left| \langle \psi_{2,1,0} \left| \frac{2}{\sqrt{5}}\psi_{1,0,0} + \frac{1}{\sqrt{5}}\psi_{2,1,0} \right. \right|^2 \\ &= \left| \frac{2}{\sqrt{5}} \langle \psi_{2,1,0} | \psi_{1,0,0} \rangle + \frac{1}{\sqrt{5}} \langle \psi_{2,1,0} | \psi_{2,1,0} \rangle \right|^2 \\ &= \left| 0 + \frac{1}{\sqrt{5}}(1) \right|^2 = \frac{1}{5}. \end{aligned}$$

Notice the sum of the probabilities is  $1$ ,  $P(E = -13.6 \text{ eV}) + P(E = -3.4 \text{ eV}) = 4/5 + 1/5 = 1$ , for example. Notice also, we use the normalized state vector, and the bra of the eigenstate serves only to identify the eigenstate is being used for the calculation.

Lastly, there is but one possibility for  $z$ -component of angular momentum, corresponding to  $m = 0$  in both eigenstates, so we can conclude

$$P(\mathcal{L}_z = 0) = 1,$$

without further calculation. Just to illustrate a calculation for which two eigenstates have the same eigenvalue,

$$\begin{aligned}
 P(\mathcal{L}_z = 0) &= \left| \langle \psi_{1,0,0} \left| \frac{2}{\sqrt{5}} \psi_{1,0,0} + \frac{1}{\sqrt{5}} \psi_{2,1,0} \right. \right|^2 + \left| \langle \psi_{2,1,0} \left| \frac{2}{\sqrt{5}} \psi_{1,0,0} + \frac{1}{\sqrt{5}} \psi_{2,1,0} \right. \right|^2 \\
 &= \left| \frac{2}{\sqrt{5}} \langle \psi_{1,0,0} | \psi_{1,0,0} \rangle \right|^2 + \left| \frac{1}{\sqrt{5}} \langle \psi_{2,1,0} | \psi_{2,1,0} \rangle \right|^2 \\
 &= \left| \frac{2}{\sqrt{5}} (1) \right|^2 + \left| \frac{1}{\sqrt{5}} (1) \right|^2 \\
 &= \frac{4}{5} + \frac{1}{5} = 1.
 \end{aligned}$$

There are two additional teaching points in exercise 10–16. If we have a more general linear combination of eigenstates, say

$$\Psi = \sum_i c_i \psi_i = c_1 \psi_1 + c_2 \psi_2 + c_3 \psi_3 + \dots$$

where the sum may terminate at some value  $n$  or be an infinite sum as indicated above. The  $c_i$  are **expansion coefficients** or more simply, the  $c_i$  are individual normalization constants for individual eigenstates. Generalizing from a two eigenstate system to a many eigenstate system, the probability of measuring a property associated with one specific eigenstate, say eigenstate  $\psi_j$ , is

$$P = | \langle \psi_j | \Psi \rangle |^2 = | \langle \psi_j | c_1 \psi_1 + c_2 \psi_2 + c_3 \psi_3 + \dots \rangle |^2 = | c_j \langle \psi_j | \psi_j \rangle |^2 = | c_j |^2,$$

where the orthonormality condition  $\langle i | j \rangle = \delta_{ij}$  tells us  $\langle \psi_j | \psi_j \rangle = 1$ , and cross terms with unlike indices are zero. Knowing this, the calculation for any property associated with  $\psi_{1,0,0}$  for instance is

$$P(\psi_{1,0,0}) = \left| \frac{2}{\sqrt{5}} \right|^2 = \frac{4}{5}$$

and the remainder of the symbology in the calculation is unnecessary.

The second teaching point concerns the calculation of  $P(\mathcal{L}_z = 0)$ . Why do we use

$$P(\mathcal{L}_z = 0) = \left| \langle \psi_{1,0,0} \left| \frac{2}{\sqrt{5}} \psi_{1,0,0} + \frac{1}{\sqrt{5}} \psi_{2,1,0} \right. \right|^2 + \left| \langle \psi_{2,1,0} \left| \frac{2}{\sqrt{5}} \psi_{1,0,0} + \frac{1}{\sqrt{5}} \psi_{2,1,0} \right. \right|^2,$$

and not

$$\left| \langle \psi_{1,0,0} + \psi_{2,1,0} \left| \frac{2}{\sqrt{5}} \psi_{1,0,0} + \frac{1}{\sqrt{5}} \psi_{2,1,0} \right. \right|^2?$$

The answer, referring to equations (1–2) and (1–3), is a measurement will place a linear combination into one specific eigenstate. We have a final state which is distinctly one of two eigenstates, and we can ascertain from the measurement which one. That tells us  $|\psi|^2 = |\psi_a|^2 + |\psi_b|^2$ . If we could not uniquely distinguish a final state, a situation we will encounter in the circumstance of degeneracy,  $|\psi|^2 = |\psi_a + \psi_b|^2$ . One reason a complete set of commuting observables is important is to provide a capability to uniquely distinguish final states, thus avoiding the indeterminacy associated with

$|\psi|^2 = |\psi_a + \psi_b|^2$ , where we do not know what portion of the probability originates with which eigenstate.

## Expectation Value, Uncertainty and Time Dependence

---

**Example 10–17:** What are the expectation values of energy, orbital angular momentum, and the  $z$ -component of angular momentum?

$$\langle E \rangle = \sum_i P(E_i) E_i = \frac{4}{5}(-13.6 \text{ eV}) + \frac{1}{5}(-3.4 \text{ eV}) = -10.88 \text{ eV} - 0.68 \text{ eV} = -11.56 \text{ eV}.$$

$$\langle \sqrt{\mathcal{L}^2} \rangle = \sum_i P(\alpha_i) \alpha_i = \frac{4}{5}(0) + \frac{1}{5}(\sqrt{2}\hbar) = \frac{\sqrt{2}\hbar}{5}.$$

The expectation value of the  $z$ -component of angular momentum is 0, since that is the only possibility.

---

**Example 10–18:** What are the uncertainties of energy, orbital angular momentum, and the  $z$ -component of angular momentum?

$$\begin{aligned} \Delta E &= \sqrt{\sum_i P(E_i) (E_i - \langle E \rangle)^2} \\ &= \left[ \frac{4}{5}(-13.6 \text{ eV} - -11.56 \text{ eV})^2 + \frac{1}{5}(-3.4 \text{ eV} - -11.56 \text{ eV})^2 \right]^{1/2} \\ &= \left[ \frac{4}{5}(-2.04)^2 + \frac{1}{5}(8.16)^2 \right]^{1/2} \text{ eV} = \left[ \frac{4}{5}(4.16) + \frac{1}{5}(66.59) \right]^{1/2} \text{ eV} \\ &= [3.33 + 13.32]^{1/2} \text{ eV} = [16.65]^{1/2} \text{ eV} \\ &= 4.08 \text{ eV}. \end{aligned}$$

$$\begin{aligned} \Delta \sqrt{\mathcal{L}^2} &= \left[ \frac{4}{5} \left( 0 - \frac{\sqrt{2}\hbar}{5} \right)^2 + \frac{1}{5} \left( \sqrt{2}\hbar - \frac{\sqrt{2}\hbar}{5} \right)^2 \right]^{1/2} = \left[ \frac{4}{5} \left( -\frac{\sqrt{2}}{5} \right)^2 + \frac{1}{5} \left( \frac{4\sqrt{2}}{5} \right)^2 \right]^{1/2} \hbar \\ &= \left[ \frac{4}{5} \left( \frac{2}{5^2} \right) + \frac{1}{5} \left( \frac{16 \cdot 2}{5^2} \right) \right]^{1/2} \hbar = \left[ \frac{8}{5^3} + \frac{32}{5^3} \right]^{1/2} \hbar = \left[ \frac{40}{5^3} \right]^{1/2} \hbar \\ &= \frac{2\sqrt{2}}{5} \hbar. \end{aligned}$$

The uncertainty in the  $z$ -component of angular momentum is 0, since there is only one possibility.

---

**Example 10–19:** What is the time dependent state vector?

$$\begin{aligned} |\Psi(t)\rangle &= \sum |j\rangle \langle j|\psi(0)\rangle e^{-iE_j t/\hbar} \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1, 0) \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} e^{-iE_1 t/\hbar} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0, 1) \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} e^{-iE_2 t/\hbar} \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{1}{\sqrt{5}} (2) e^{-iE_1 t/\hbar} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{1}{\sqrt{5}} (1) e^{-iE_2 t/\hbar} \\ &= \frac{2}{\sqrt{5}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-iE_1 t/\hbar} + \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-iE_2 t/\hbar}, \end{aligned}$$

and using eigenenergies in terms of fundamental constants, *i.e.*,

$$E_n = -\frac{\hbar^2}{2\mu a_0^2 n^2},$$

this becomes

$$|\Psi(t)\rangle = \frac{2}{\sqrt{5}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{it\hbar/2\mu a_0^2} + \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{it\hbar/8\mu a_0^2}.$$

---